

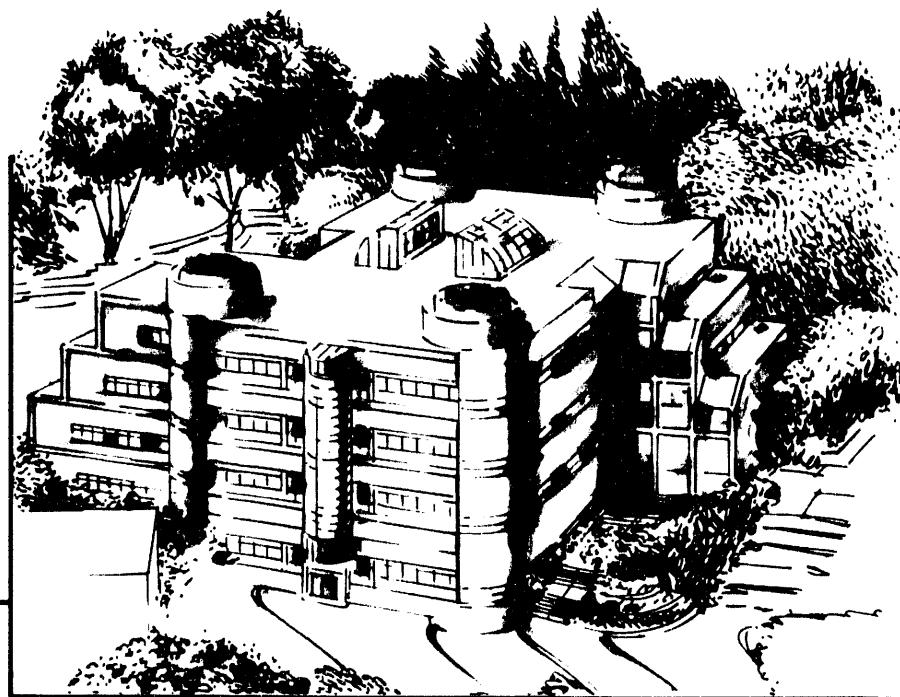
Center for Advanced Materials

CAM

A Low Temperature Scanning Force Microscope for Biological Samples

M.G.L. Gustafsson
(Ph.D. Thesis)

May 1993



Materials and Chemical Sciences Division
Lawrence Berkeley Laboratory • University of California
ONE CYCLOTRON ROAD, BERKELEY, CA 94720 • (415) 486-4755

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. Neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

Lawrence Berkeley Laboratory is an equal opportunity employer.

A Low Temperature Scanning Force Microscope for Biological Samples

**Mats Gustaf Lennart Gustafsson
(Ph.D. Thesis)**

**Department of Physics, University of California,
Berkeley, California 94720**

and

**Center for Advanced Materials
Materials Sciences Division
Lawrence Berkeley Laboratory
Berkeley, California 94720**

May 1993

This work was supported by the California Competitive Technology Program and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under contract number DE-AC03-76SF00098.

MASTER
DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

A Low Temperature Scanning Force Microscope for Biological Samples

Copyright © 1993

by

Mats Gustaf Lennart Gustafsson

**The U.S. Department of Energy has the right to use this document
for any purpose whatsoever including the right to reproduce
all or any part thereof**

Abstract

The scanning force microscope (SFM) is capable of resolving individual atoms on simple, crystalline samples. Much could be learned if biological samples could be imaged with such quality. However, when SFM's have been applied to biology, the observed resolution has typically been at least an order of magnitude worse.

Part of the reason for this difference may be that biological samples are much more deformable, and undergo larger thermal motion, than the crystalline samples. It has been suggested that this situation might improve at lower temperatures, and that the relevant temperature threshold is around 180 K.

I have constructed a SFM which is capable of operating at 143 K, and which is designed expressly for biological applications.

In conjunction with discussing the design of this SFM system, I describe two other contributions to SFM technology: a new method of fabricating tips, and new designs of SFM springs that significantly lower the noise level.

We have used this system to image several biological samples — including collagen, ferritin, RNA and purple membrane — at 143 K and at room temperature.

When comparing images from the two temperatures, we have not observed any dramatic difference in resolution. There are several possible interpretations of this:

- (i) An even lower temperature may be necessary.

- (ii) The resolution may be limited by tip quality, not yet by sample deformation.
- (iii) The prediction of greater sample rigidity at low temperature may be wrong.

I will argue that (ii) is the case, making it impossible to pass judgement on (i) or (iii) at this point.

Regardless of the question of resolution, the low temperature SFM allows the study of new categories of samples, such as those prepared by freeze-fracture. Furthermore, it has the potential of measuring the temperature dependence of the mechanical properties of single molecules. Such measurements could yield new information about the energetics of molecular stability and binding.

We have also explored the use of the SFM as a sample manipulation tool. In particular, we have used it to cut single collagen molecules into segments with a precision of 10 nm or better.

A Low Temperature Scanning Force Microscope for Biological Samples

Contents

I	Introduction	1
	Overview of Thesis	1
	Scanning probe microscopies	2
	SFM detection methods	12
	State of microscopy for biology	16
	The place of SFM	22
	Why a <u>Low Temperature</u> SFM?	23
II	Noise Optimization for the Beam Deflection Method	29
	Background	29
	Geometry	32
	Optimization of Focusing	35
	Comparison to Interferometry	42
	Spring Design	45
III	Microfabrication of Springs	51
	Background: Fractured Tips	51
	Fabrication	54
	Evaluation of Tips	60
	Possible improvements	65
IV	Noise Measurement	71

V Design Concerns for Biological Applications	77
VI The Microscope System	81
SFM Platform	81
Cooler and Dry Chamber	89
Electronics	92
Experimental Procedure	101
VII Calibration	105
Hysteresis of PZT	105
XY Calibration	107
Relative Calibration	112
Z Calibration	114
VIII Images	119
Mica	120
Ferritin	120
Purple Membrane	127
Collagen	142
RNA	153
IX Cold vs. Warm	158
X Sample Modification	174
Removal of Single Molecules	174
Splitting / Removal of Membranes	177
Surgery on Single Molecules	182
XI Future Possibilities	187
Freeze Fracture	187
Elasticity Measurements	196
XII Summary	192

Acknowledgements

I would like to take this opportunity to thank a few of the many individuals who have helped me over the years:

- My advisor, John Clarke, for many useful discussions and, not the least, for finding the funding for this project.
- My collaborator Mark Shattuck, who has done much of the sample preparation and characterization, including all the collagen work, and whose unstoppable optimism and energy have been invaluable.
- Knute Fisher, whose experience and resources in the field of microscopy have been the basis for this entire effort, and Kathleen Yanagimoto, for her careful, consistent and patient sample preparation and TEM work.
- Eric Ganz and Ruth Ellen Thomson, for sharing their hardwon wisdom about STM, and all other members of the Clarke group for teaching me the tricks of the trade.
- Marc Hertlein, for helping me over the threshold to Macintosh programming, and writing the framework of our acquisition software.
- Art Veis, for supplying the lathyritic collagen, and much collagen wisdom, and Yuri Lyubchenko, for sharing his RNA and DNA samples.
- All members of the Berkeley Sensor and Actuator Center for teaching me most of what I know about microfabrication, and the staff of the Microfabrication Laboratory, for maintaining a high quality facility where it is easy and even fun to implement it.

- Past and present members of the Physics department machine shop, and particularly Al Daft, who patiently introduced me to the subtleties of his art.
- Leverne Garner, John Davis and the rest of the Physics electronics shop, for always having one more way to beat the digital noise pickup.
- John Stewart, for patient machining help during the summers he spent in our lab, and Tom Swain, for building a tunneling readout SFM with me in the early days of this project.
- Sue Whitfield, for her contributions as volunteer lab tech and friend.
- All my housemates at 2133 Ward Street over the years, for their friendship, support and good food.
- David Hammer and Rob Zako, for the brainstorming sessions that helped me survive the first years of Physics graduate school, and for the backpacking and bicycling adventures that have helped me survive the rest.

Most of all, I would like to thank Julie Sohl for her immense moral support, without which I would have pushed this machine off a cliff a long time ago.

[This work was supported in part by the Director, Office of Energy Research, Office of Basic Energy Science, Materials Science Division of the U. S. Department of Energy, under contract number DE-AC03-76SF00098.]

Chapter I: Introduction

Overview of Thesis

The subject of this thesis is a scanning force microscope (SFM) for the study of biological samples at low temperature. The thesis can be thought of as having four parts: (1) an introduction (chapter I), (2) a description of new designs of SFM springs and tips (chapters II-IV), (3) a description of our low temperature SFM (chapters V-VII), and (4) a presentation and discussion of SFM image data acquired with this system (chapters VIII-XI).

In the introduction, after introducing the concept and origin of SFM and other scanned probe microscopies, and the role of SFM as an addition to the existing set of microscopy tools available for biological applications, I will present some arguments why a low temperature SFM could be expected to be advantageous.

The second section discusses the detection system and tips. (I mention it as a separate section because the improvements described there apply to SFM in general, not specifically to the low temperature SFM.) After considering the various possible detection schemes, I will digress into a calculation of the optimum noise level of the method we used, the "optical beam deflection" method, and draw some conclusions about how best to design the micro-springs which are the heart of an SFM. I will then describe the process I used to manufacture such springs, including my unique "controlled

fracture" technique for supplying them with extremely sharp tips, and compare this to other tip making methods in use.

The next section constitutes a description of our SFM system hardware, electronics and software, and of how its design was determined by the requirements of biological samples.

In the "Data" section I describe our experiences using this SFM to image several different biological systems, show some examples from the more than 1500 images we have acquired with this system, both at room temperature and at 143 K, and discuss what conclusions can be drawn about possible advantages of low temperature SFM. I will also describe experiments in which we have explored the use of the SFM not only as an imaging tool, but as a sample modification tool.

Finally, I will discuss possible future directions, including the exciting possibility of using low temperature SFM's to measure directly the mechanical properties of single biomolecules, and the temperature dependence of these properties.

Scanning probe microscopies

In the last decade a growing family of microscopic techniques have appeared, which is generically referred to as scanning probe microscopies¹ (SPM). These techniques are all based on the same basic idea (Fig. 1-1): A sharp tip is brought into proximity of a sample surface, and some short-range tip-sample interaction is monitored. The intensity of this interaction depends on — and is thus a measure of —the distance from the tip to the sample. The measured intensity is used to feed back to the tip position (in the "Z"

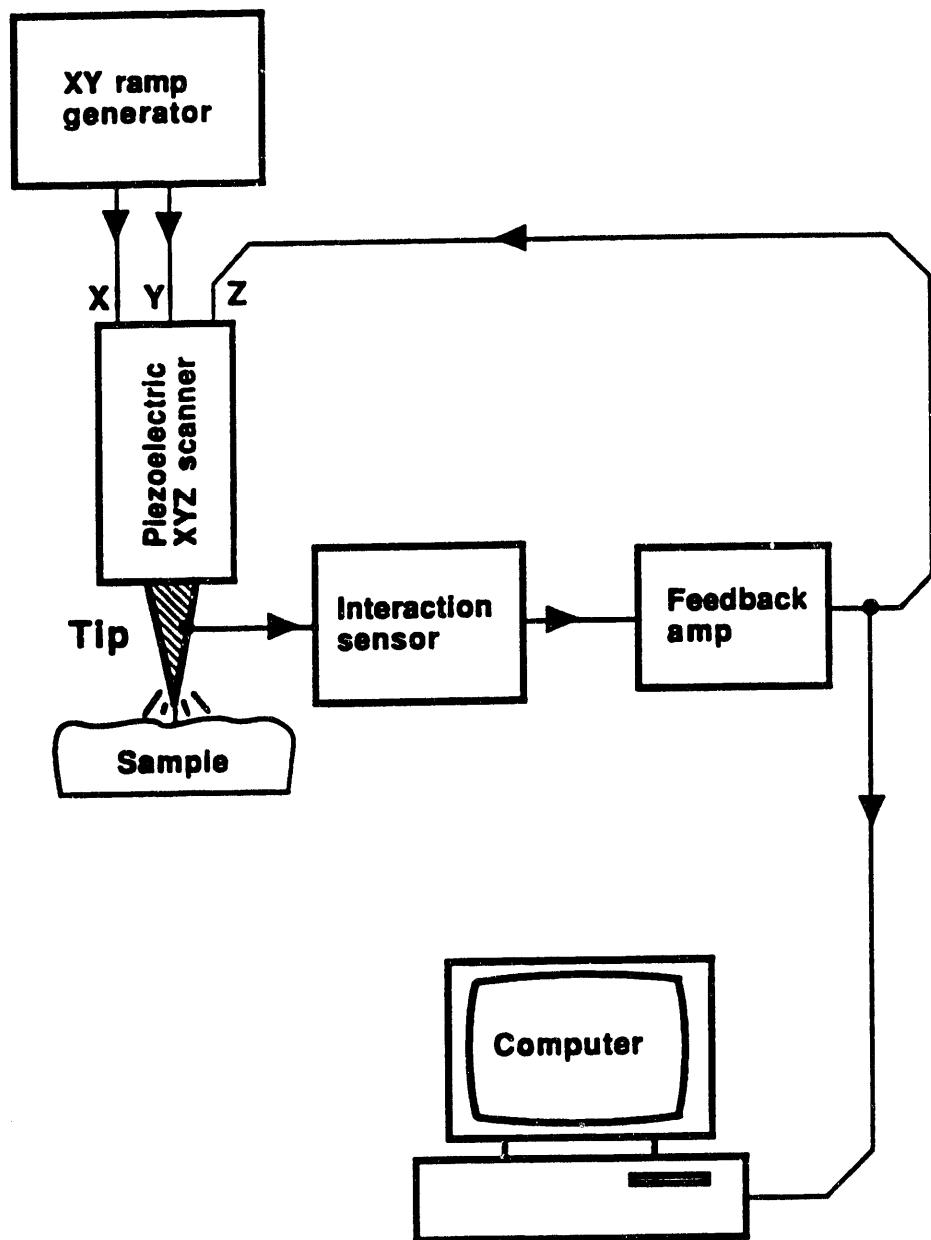


Figure 1-1. Schematic of a generic scanned probe microscope.

direction, perpendicular to the sample surface), so as to regulate the tip-sample distance at a constant value. The tip is then scanned laterally (parallel to the sample surface, in the "XY plane") in a raster pattern. As the tip encounters the hills and valleys of the sample surface, the feedback regulation will cause it to be withdrawn or advanced in order to maintain a constant tip-sample distance. This feedback signal thus contains information about the sample topography, and is recorded and used to generate three-dimensional images of the sample surface. (In some cases it is more convenient to scan and/or feedback regulate the position of the sample, instead of the tip. This is entirely equivalent.) The Z resolution of an SPM technique depends on how strongly the interaction in question varies with the tip-sample distance, and the precision with which it can be measured. In many cases the Z resolution is better than typical interatomic distances. The XY resolution depends on the sharpness of the tip, and also on the range of the interaction, since the latter determines how large a portion of the tip is active.

STM

The different microscopies differ only in the nature of the interaction which is monitored. In the original technique, scanning tunneling microscopy² (STM), a bias voltage is applied between the tip and the sample, and the interaction in question is the tunneling current that can occur if the distance between them is below a nm or so. (In typical operation, the bias voltage is on the order of 100 mV and the tunneling current 0.1-1 nA.) Since the tunneling current depends extremely strongly on distance, changing by about a factor

of 10 for a distance change of 0.1 nm, atomic resolution is relatively easy to achieve, and STM is a powerful technique. However, it is obviously limited to electrically conducting samples, and is therefore of limited use for biology.

OTHER SPM'S

While STM and SFM (to be described in the next paragraph) have now developed into commercial technologies, other scanning probe microscopies are still in the exploratory stage. Examples are

- Scanning thermal microscopy³, where one measures the rate of heat transfer between the sample and a tip which contains a microscopic temperature sensor.
- Scanning ion conductance microscopy⁴, where the tip is hollow and operated in an aqueous solution, and one measures the change in ion conductance between the inside and the outside when the opening at the apex of the tip is partially blocked by the sample. For biological purposes, SICM has the advantage of being a non-contact technique that can be performed in a natural aqueous environment, and that processes related to ion transport can potentially be studied. However, the resolution at present⁵ is no better than 100 nm, limited by tip technology.
- Various techniques based on the detection of photons emitted in an STM-like situation,⁶ or, conversely,
- the detection of the effect of incoming photons on the tunneling current⁷.

SFM

The technique that is the subject of this thesis, scanning force microscopy (SFM, also called atomic force microscopy, AFM) was actually the second SPM technique to be invented,⁸ in 1985. In SFM, the interaction one measures is a mechanical force between tip and sample. The tip is placed on a weak spring, typically a cantilever, and one detects the force by measuring the deflections of this spring. One could object that nothing is gained by the presence of the tip and spring, since the position of the spring must be measured with sub-atomic precision if such precision is desired in the image. However, it is a fundamentally easier task to measure the position of a macroscopic object than of a sub-nanometer spot on the sample.

FORCES

SFM is itself not one technique but a family of techniques, since there are several interactions which can manifest themselves as a force. For special purposes the SFM can be made sensitive to magnetic⁹ or electrostatic¹⁰ forces. If one's interest is simply the sample topography, the forces of interest are the long range, attractive van der Waals interaction, and the short range exchange repulsion. The latter can equivalently be thought of as the repulsive part of the van der Waals force, or, more down to earth, as the same mechanical contact force that keeps your foot from sinking through the floor. If one operates in the attractive regime the forces used can be very low, down to a few pN, but at the expense of a large tip-sample distance, which by necessity means a correspondingly low lateral resolution. Because of the short range nature of the repulsive interaction, the resolution in the repulsive regime can be much

higher — in fact, atomic resolution has been achieved on a number of samples¹¹ — but the forces involved are larger, typically 0.1-1 nN. From now on, I will refer to the repulsive mode unless stated otherwise.

If an SFM is operated in air, a third effect occurs. As anyone knows who has tried to pump out a vacuum chamber, all surfaces exposed to air are covered with a thin film consisting mainly of water. If an SFM tip is brought close to a sample, at one point their water layers will touch and merge, and the resulting water surface takes on an hourglass shape (Fig. 1-2). The associated surface tension creates a force which pulls the tip toward the sample until it makes mechanical contact and a repulsive force of the same magnitude develops at the tip. Thus, even if the total, measured force is zero, there is a large, repulsive force applied at the tip. The magnitude of this force is typically around 100 nN, much larger than the force levels at which one would like to work, and at which one could work if this "meniscus effect" did not exist. The meniscus force can be partly compensated by pulling on the tip (i. e. regulating the total force at a negative value), but not fully, since in practice the meniscus force is not totally constant as the tip is scanned, and if it is at any point overcompensated for, the tip will pull away from the surface and the scan will have to be aborted. The force with which the tip has to be pulled back before it leaves the surface will be referred to as the "snap-out" force. A large snap-out is a sign of a severe meniscus effect, or a contaminated tip.

As was first pointed out by Paul Hansma,¹² the meniscus effect can be circumvented by operating in a liquid, where no air/water

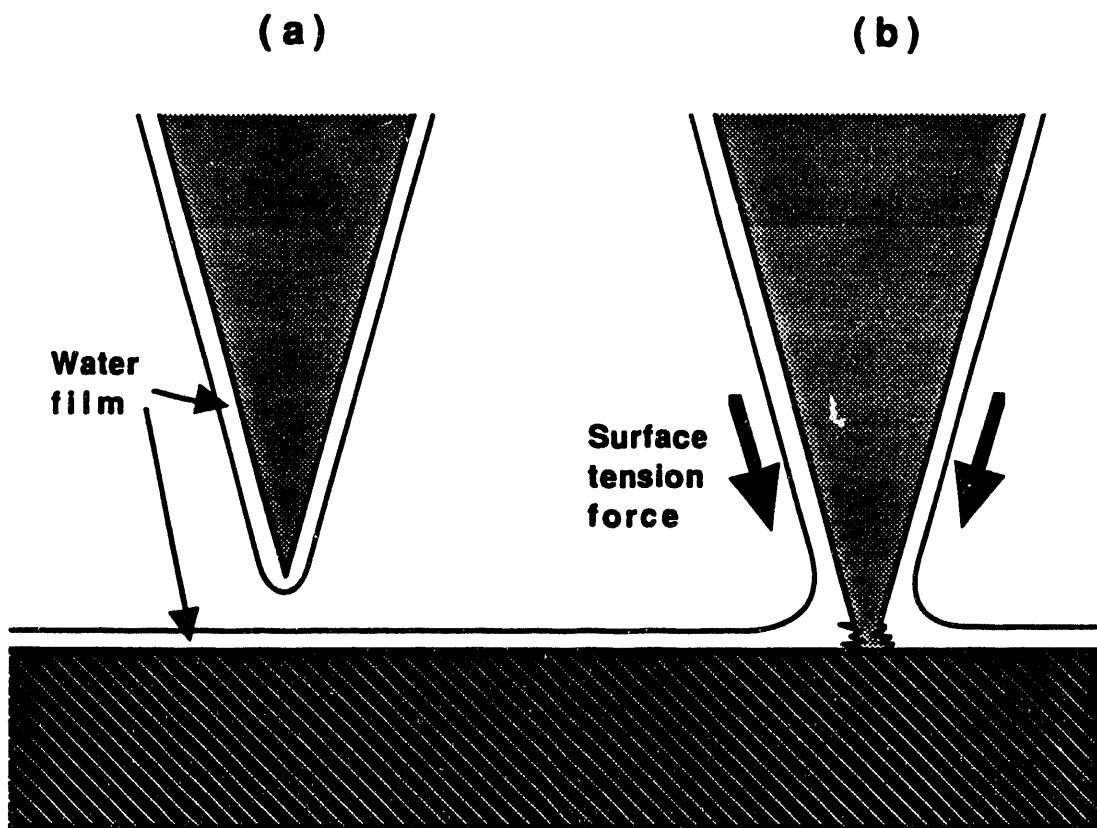


Figure 1-2. The meniscus effect. (a) If exposed to air, both the tip and the sample will be covered with a water film. (b) When the tip is brought near the sample, these films touch and merge. The surface tension of the combined film creates a strong attractive force.

interface, with associated surface tension, can exist. When scanning in liquids one can typically use forces in the 0.1-1 nN range (30 pN has been reported¹³), while in air forces of 50-100 nN are necessary.

Another mechanism contributes to the advantage, in the sense of being able to work at low force, of working in liquids. The total force measured in the SFM is the sum of all forces acting on the tip. In the absence of other effects, this means the integral of the van der Waals force over the volume of the tip. Since the range of the attractive part of the van der Waals force is relatively large, there will be a sizeable attractive contribution to this integral from regions of the tip which are far from the sample. At zero total force, there must be an equal, repulsive force acting at the apex. As discussed for the meniscus force, one can compensate for this force by operating at an attractive total force, but only to the extent it is constant over the scan area. The effect of immersing the tip and sample in liquid is to partially compensate for the van der Waals interaction. This can be seen most obviously by considering the case when the van der Waals interactions of the liquid and the sample with the tip are exactly equivalent. The tip would not be able to tell the difference between the liquid and the sample, and the effective force on the tip would be zero independently of the distance to the sample (as long as it was out of range of the exchange repulsion). With the appropriate choice of liquids, it should be possible to eliminate totally the attractive van der Waals force, or even to make the effective force repulsive. Thus it may be that even operation of an SFM in vacuum will not allow forces as low as are possible in liquids.

IMAGING MODES

All scanned probe microscopies, including SFM, have several possible imaging modes. The one outlined above, where the feedback signal is the basis for the image, can be called the "feedback," "closed loop," or "topographic" mode. It is also referred to in the context of STM as the "constant current" mode, in SFM as the "constant force" mode, and so on.

The second mode involves interrupting the feedback loop and using the, now varying, value of the tip-sample interaction directly to form the image. This mode is known by many names, all of which are confusing. I will refer to it in this thesis as the "open loop" mode. Its main advantage is that one can scan much faster, since the feedback response time is no longer a limitation. The strong disadvantage is that it can only be used if the sample is flat enough for the tip to stay within the usable range of the sample even without feedback. (I. e. if the tip does not get too far away for the interaction to be measurable, and does not get so close as to risk tip or sample damage.) For STM, this range is less than a nm, while in SFM it depends on the force constant of the spring and the fragility of the sample, but can be several tens of nm in some cases.

A third possible mode is the "error signal" mode, where the feedback loop is in effect, but the interaction intensity is used for the image anyway. Since the interaction is in principle held constant by the feedback loop, what is displayed in the image is only a residual error, hence the name. This has the effect of differentiating the image in the direction of the scan, creating an effect not unlike a lightsourced 3D image (Fig. 1-3). This image does not contain

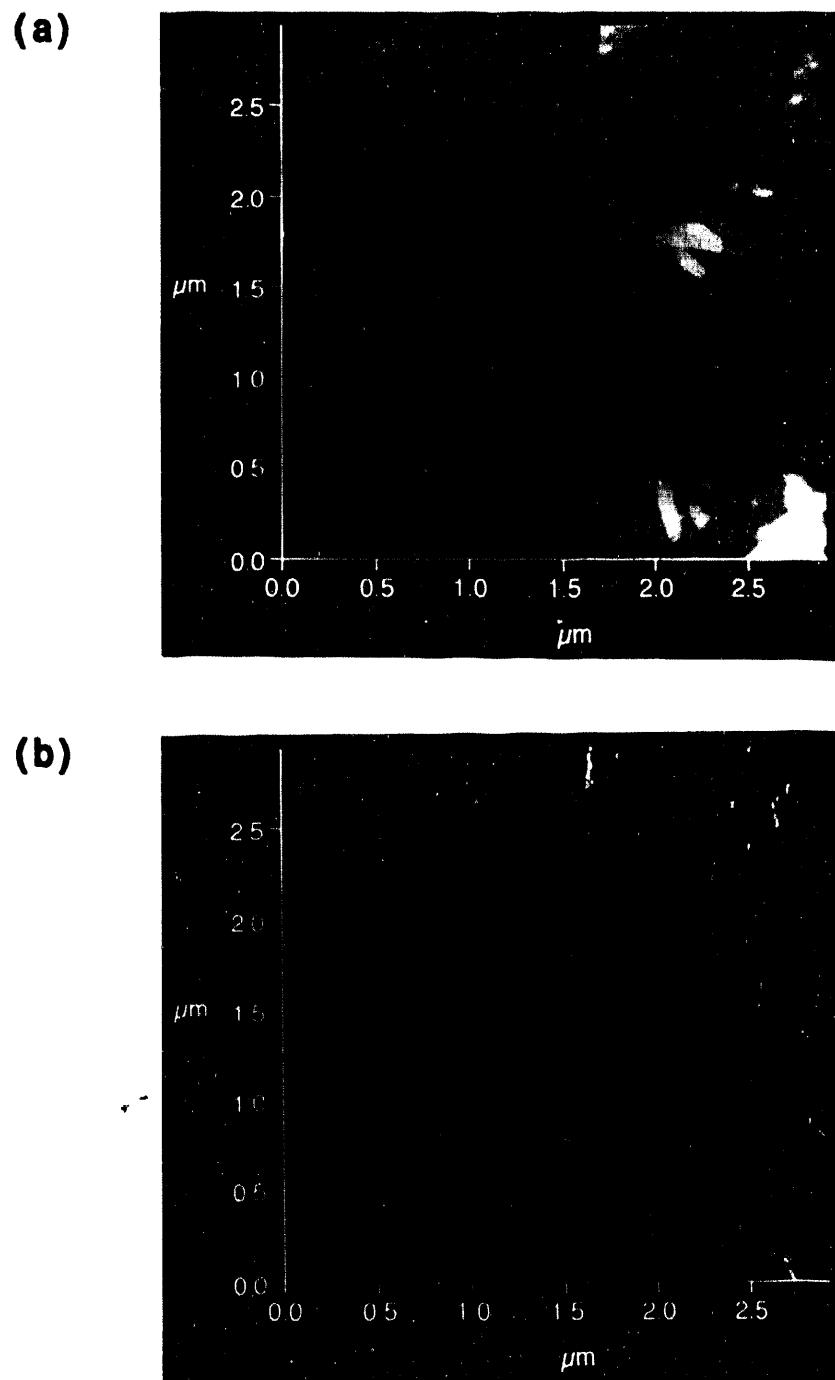


Figure 1-3. The same sample imaged in (a) the feedback mode and (b) the error signal mode. The sample consists of patches of cell membrane ("purple membrane," described in more detail in chapter VIII), in places overlapping or folded in double. An image in the open loop mode would look very similar to (a).

quantitative height information, but makes it possible to view simultaneously the surface details of parts of the sample which are at widely different heights. (Of course, if the dynamic range is sufficient, one can create a similar effect after the fact by differentiating an image that was acquired in the feedback or open loop modes.)

SFM Detection methods

When I began working on SFM's in 1987, SFM as a technology was still in its infancy, and no consensus existed as to what the most suitable methods were for detecting deflection of the SFM spring. Before embarking on the system design, we considered the pros and cons of a number of different detection schemes. I will briefly describe the various methods in this section.

TUNNELING

The original paper on SFM¹⁴ used a tunneling tip, much like an STM tip, on the back side of the spring to detect the spring deflection [Fig. 1-4 (a)]. This is in theory a very sensitive method, and we based our first SFM designs on this concept. However, in practice it is very difficult to ensure that the tunneling gap is free of contamination. If any contamination is present, the tunneling tip will have to apply a force to penetrate it, and this force can be as large as or greater than the forces one is trying to measure. The phenomenon of forces caused by contamination in the tunneling gap had been studied earlier in the context of STM, and the problem is severe not only in air, but even in ultra high vacuum.¹⁵ This problem made tunneling readout SFM's unreliable, and the technique was all but abandoned

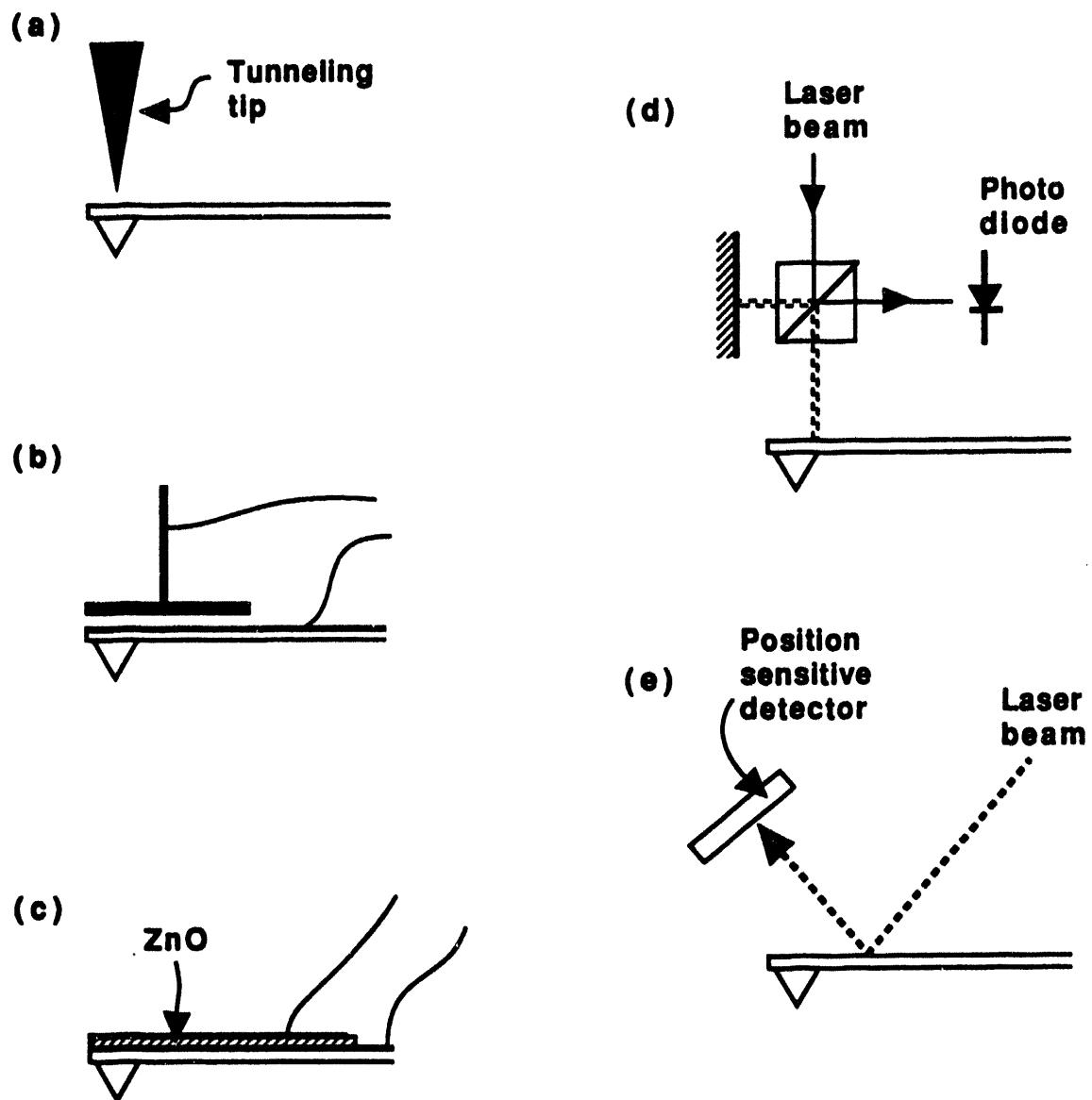


Figure 1-4. SFM detection methods: (a) tunneling, (b) capacitance, (c) piezoelectric cantilever, (d) interferometry, (e) optical beam deflection.

soon after other, more reliable methods appeared. It is now used only in special applications such as helium temperature UHV SFM.¹⁶

CAPACITANCE

Another possibility which we considered at length is capacitive readout. In this scheme, one senses the capacitance between the back side of the spring and a nearby plate [Fig. 1-4 (b)]. There are two different ways to arrive at a usable geometry: (i) One can bring the counter electrode into close proximity of the spring under feedback control, in a way analogous to the tip-sample approach in an STM, or (ii) one can microfabricate the whole parallel-plate assembly. In case (i), the counter electrode can in principle be brought arbitrarily close to the spring, limited only by surface roughness and contamination. The effective area, however, is small, since in practice the counter electrode has to be convex in both directions (two flat plates would have to be aligned to a prohibitive precision to get so close without touching anywhere). In case (ii), the plate area can be much larger, but the gain is not as large as one would think, since the plates separation is also increased, and the sensitivity scales with the square of the separation.

Capacitances can be measured with exquisite precision.¹⁷ However, the capacitance changes involved here are small enough to challenge the limits of this precision. As a numerical example, a parallel plate capacitor with an area of $100 \times 100 \mu\text{m}$ and a spacing of $1 \mu\text{m}$ has a capacitance of 88 fF . In an SFM, one would like to be able to measure changes of one part in 10^5 of this distance (10 pm , a typical atomic corrugation on the surface of a crystalline sample) in a bandwidth of about 1 kHz . This would require a measurement

sensitivity of $0.027 \text{ aF}/\sqrt{\text{Hz}}$, a nontrivial but not impossible goal — an electronics system designed specifically for capacitive SFM has been published¹⁸ (several years after the time we considered this method) which has a noise level of $0.014 \text{ aF}/\sqrt{\text{Hz}}$.

The capacitive readout method has very attractive aspects, particularly version (ii) above, where the whole readout system is self-contained on a single chip. However, if such an SFM is operated in air, the spring becomes strongly overdamped because of the thin, effectively trapped air film between the electrodes.¹⁹ This can be somewhat alleviated by perforating one the spring with a large number of closely spaced vent holes. However, if one also wants to be able to operate in a liquid (which, as we have seen earlier, is advantageous because of the meniscus effect in air), the damping problem is likely to be extreme. Thus the capacitance method seems most suited for vacuum applications.

PIEZOELECTRIC

If a piezoelectric layer such as zinc oxide²⁰ is included in the SFM spring when it is fabricated, one can use the charge (or voltage) induced in this layer to detect the spring deflection [Fig. 1-4 (c)]. However, this technique does not work at zero frequency, because of dissipation of the charge through unavoidable leakage conductances. The method does share with the microfabricated version of the capacitive scheme the attractive property of being self-contained on a chip, and my estimates indicated that the signal was sufficiently strong to be usable in the attractive force regime (where AC detection can be used). However, our main interest was in the

repulsive force mode of operation, where DC sensitivity is necessary.

INTERFEROMETRY

The basic idea of this method is to construct an interferometer in such a way that one of its two laser beams is reflected from the SFM spring, so that any deflection of the spring causes a detectable change in the interference pattern [Fig. 1-4 (d)]. Several variations on this concept have been developed, including homodyne,²¹ heterodyne²² and all-fiberoptic²³ interferometers. These methods can reach very low noise levels, particularly in the AC mode.

BEAM DEFLECTION

In this method, one reflects a laser beam from the back side of the spring, and detects any changes in the *angle* of the reflected beam using a position sensitive detector [Fig. 1-4 (e)].

The beam deflection and interferometric methods are now dominating the SFM field, by virtue of high reliability and low noise. Since they will be described and analyzed in chapter II, I will not court repetitiousness by going into any more detail here.

State of microscopy for biology

A full review of all the microscopic techniques now available to the biologist is obviously out of the scope of this thesis. However, I would like to give a little background as to what is there, to create a context in which to judge the abilities of the SFM.

OPTICAL

Optical microscopy has recently seen a revival of interest, because of the advent of confocal techniques²⁴ and fluorescent

labeling.²⁵ Optical microscopy has strong advantages in that samples can be imaged in a totally native state while the processes of life are taking place, and in that the interior, not just the surface, of the sample can be viewed. However, it is obviously not a high resolution technique, its resolution being limited by diffraction to several hundred nm.

SEM

Higher resolution can be achieved with scanning electron microscopy (SEM). This is a purely surface sensitive technique, and requires the sample to be exposed to vacuum, and usually that it be coated with a conducting film to avoid charging effects. The resolution is limited by the conducting film, the spot size, and the diffusion length of the secondary electrons, but can reach 6-10 nm.

TEM

The highest resolution conventional technique available is transmission electron microscopy (TEM), which can achieve atomic resolution on inorganic crystals²⁶. As in the case of SEM, the sample must be able to survive exposure to vacuum. While TEM it is not strictly a surface sensitive technique, a TEM sample must be very thin, usually 100 nm or less. To achieve contrast in TEM, biological samples can be stained with heavy elements (using compounds such as uranyl acetate or phosphotungstic acid). High resolution can be achieved Alternatively, a thin film containing a contrasting element can be deposited on the sample, embedded in a supporting carbon film, and viewed in the TEM after the sample itself has been removed by etching. Both methods are capable of visualizing single protein molecules. Figure 1-5 (a) is a TEM micrograph of a ferritin

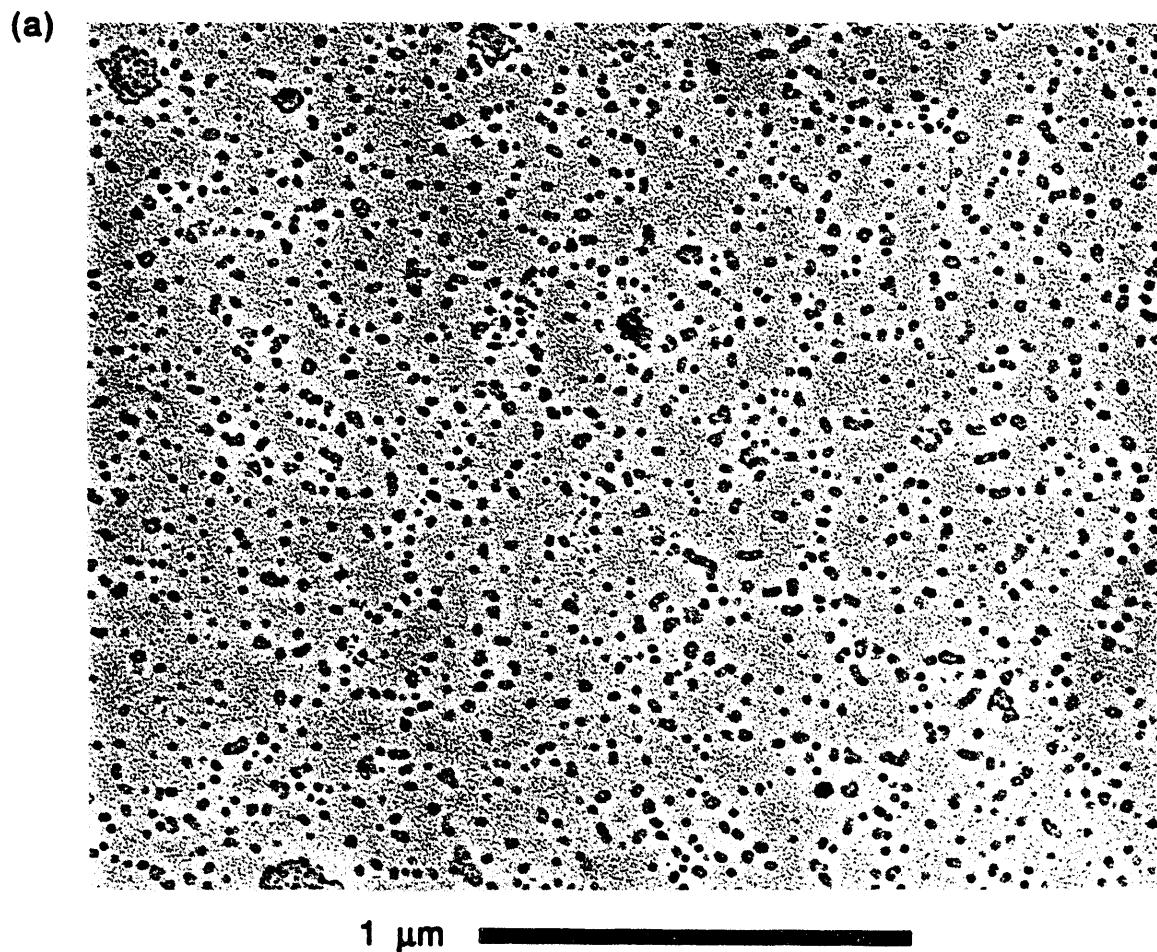


Figure 1-5. (a) A TEM micrograph of Ferritin molecules, negatively stained with phosphotungstic acid. (b) The principle of negative stain.

sample (ferritin is a spherical protein molecule which will be described in more detail in chapter VIII) that has been "negatively stained" with phosphotungstic acid. Negative staining involves letting a thin layer of stain solution dry with the sample. Surface tension leads to a concentration of stain at edges of sample features [Fig 1-5(b)], creating a dark outline. The ferritin molecules thus appear as dark circles in Fig. 1-5 (a).

Some molecules that are too small or narrow to show up with negative stain can be seen with the replication method. In Fig. 1-6(a), single molecules of collagen (a fibrous protein which is also described in chapter VIII) have been "shadowed" with a platinum-carbon mixture, i. e. this mixture was evaporated onto the sample at a grazing angle, which makes the Pt-C grains appear preferentially on the "upwind" side of sample features [Fig. 1-6 (b)]. In this case, the sample was rotated in its own plane during evaporation, so that all sides of the molecules were reached by the Pt-C ("rotary shadowing").

One realizes from the previous paragraphs that while the intrinsic resolution of the TEM is very high, a substantial amount of preparation has to be applied for a biological specimen to be visible, and that what is seen may be somewhat distantly related to the sample in its native state. In the shadowing method, the sample itself has actually been removed, and one is looking at a distribution of metal grains.

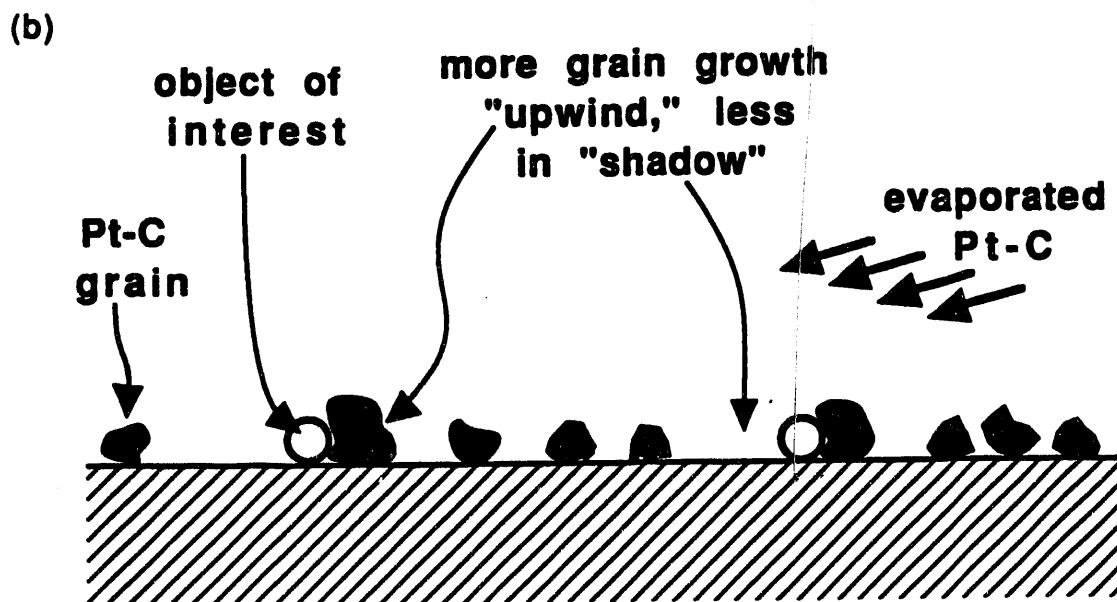
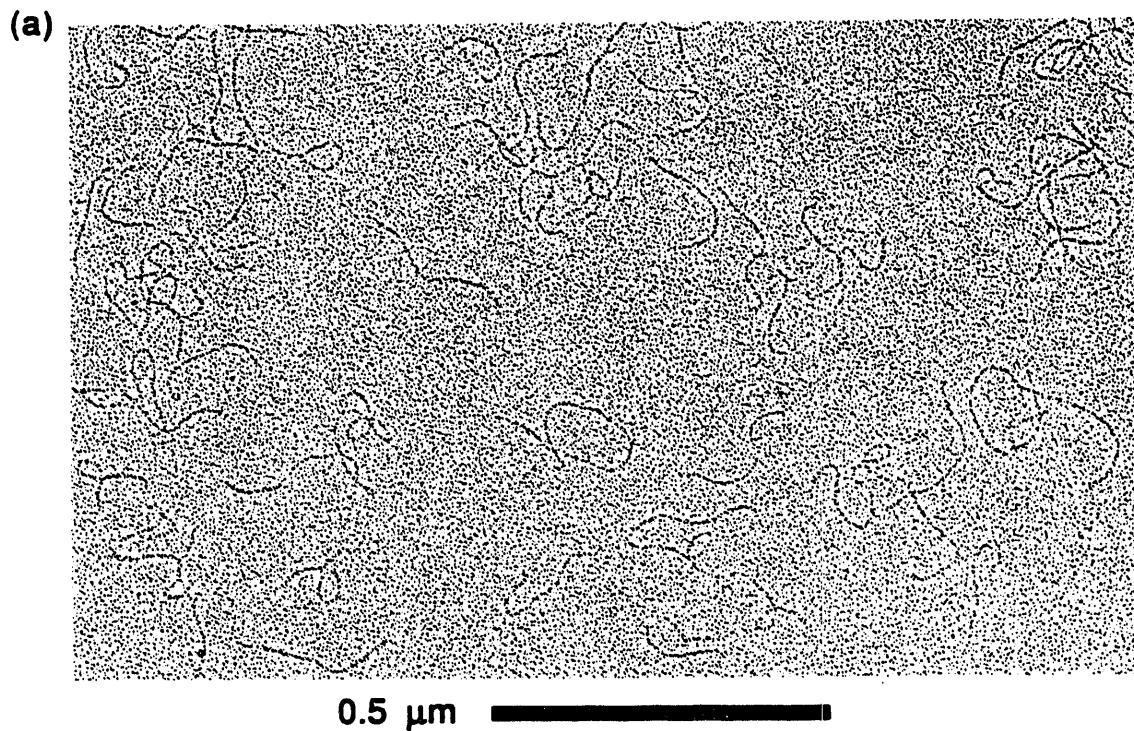


Figure 1-6. (a) A TEM micrograph of collagen monomers, rotary shadowed with a platinum-carbon mixture. (b) Shadowing involves evaporating the Pt-C at a low angle, which makes grains appear preferentially on the "upwind" side of features. In this case, the sample was rotated during evaporation, so that all sides were covered.

TEM samples cannot contain water (since they are exposed to vacuum, either in the TEM or in the evaporator, if the shadowing method is used), but simple air drying is often traumatic and can cause severe artifacts. Several methods, such as chemical fixation,²⁷ are used to alleviate this problem. One approach which circumvents the problem entirely is to image (or shadow) the sample at cryogenic temperature, where the vapor pressure of ice is low enough that the sample can be exposed to vacuum without being dehydrated. Indeed, fully hydrated samples such as whole cells in aqueous buffer can be frozen in their native configuration, and details of the cellular interior can be exposed by fracturing the frozen sample ("freeze-fracture"), after which some of the water ice may be sublimated off ("etching") to expose cellular features in relief. These very useful techniques are based on the premise that the sample can be frozen into its native state. However, if ice crystals are allowed to form, these will distort or totally disrupt the sample structure. To avoid these distortions, one must minimize the size of such ice crystals, ideally to the point where the water turns into an amorphous solid without crystallizing at all ("vitrification"). The only parameter available to achieve this is the freezing rate, and accordingly much effort has been devoted to developing "rapid freezing" technology. Typical methods²⁸ involve rapidly plunging thin samples into non-boiling cryogenic liquids (such as isopentane, propane or Freon-22) or liquid jets, or slamming samples against a cold copper block.

I will refer to these cryomicroscopy methods repeatedly later on, which is the reason I have discussed them here at some length.

SOFT X-RAYS

An intriguing idea that is just now being realized is to use soft X-rays in the equivalent of optical microscope.²⁹ This has the potential to combine the advantages of high resolution, bulk sensitivity, and the ability to image samples in their native environment. However, the resolution is so far limited by the precision with which the "lenses" (zone plates made by electron beam lithography) can be fabricated. The best resolution achieved so far³⁰ is about 50 nm. Another concern is X-ray damage to the sample.

The place of SFM

What is the role of SFM in this context? The answer to this question is still being worked out. SFM's suitable for biological work have only been commercially available a few years, and new uses are steadily appearing. Truly reproducible images of biomolecules were published for the first time as recently as 1992.³¹

In my view, there are three niches where SFM introduces new opportunities. The first is its ability to image uncoated, untreated samples in their native, aqueous environment. In particular, this means that dynamic processes can be studied, as well as static structure. Even whole, live cells can be scanned.³² While SFM is not as totally uninvasive as optical microscopy — since the tip is touching the sample — its resolution is incomparably higher.

The second niche is the very high resolution end. SFM is at an advantage over electron microscopies in that it does not require the samples to be coated or stained (though it does share the drawback

of imaging only the surface of the sample). Like the TEM, the SFM can achieve atomic resolution on easy samples^{33,34}. The ultimate resolution possible of SFM on biological samples has yet to be determined, and will quite likely continue to be improved as advances are made in tip fabrication and sample preparation. If there is a fundamental limit, this limit is likely to be set by sample deformation, which is the subject of the next section.

The third niche is not strictly a microscopy, but it is a truly unique ability of SFM: Since the SFM tip is touching the sample, it can be used not only as an imaging tool, but as a sample manipulation and force measurement tool on the nm scale. Examples of this will be described in chapters X and XI.

Why a Low Temperature SFM?

Since the invention of the SFM, it has been hoped that its remarkable, often atomic, resolution on hard, crystalline surfaces^{33,34} could be reproduced on biological samples. So far, however, the resolution achieved on samples containing large biomolecules has been more modest³⁵. Several factors may contribute to this:

- (1) The three-dimensional nature of these samples puts more stringent demands on the sharpness of the scanning tip than relatively flat crystalline samples do.
- (2) The lateral forces applied by the tip may move the sample or even dislodge it from the substrate.
- (3) The forces applied by the scanning tip may excessively deform the sample.

(4) Thermal motion at room temperature can be larger than the features of interest and thus obscure them.³⁶

While factors (1) and (2) can be (and are) attacked through the continuing development of improved tips and new substrate preparation protocols, respectively, the only possibility of improving factors (3) and (4) (at a given tip force) is through lowering the sample temperature. Therefore, it has been an important goal to develop a low temperature SFM for biological samples. In addition, such an instrument opens up several new possibilities:

Standard procedures for preparing samples for SFM involves adsorbing them firmly to some form of substrate, a process that is likely to distort their conformation. However, one can lock a biological sample into its native shape while it is still in its natural, aqueous environment, by freezing it rapidly enough to vitrify the water²⁸ (after which the sample may be freeze-fractured, and/or some of the water may be sublimated off³⁷). Such rapid-freezing is an established and successful technique in electron microscopy, and the low temperature SFM will allow its use in force microscopy.

Considerable attention^{38,39} has lately been given to exploiting the unique nature of SFM, in that it makes direct, controlled mechanical contact with the sample, turning the SFM into a probe of mechanical properties of the sample on a molecular size scale. A low temperature SFM will add to these investigations the new dimension of variable temperature over a wide range. Such variable

temperature measurements could shed light on the energetics of molecular stability and binding.

Low temperature SFM's have been developed by several groups,^{40,41} for other purposes, but attempts to build an instrument suited for biology have been hampered by vibrational noise from bubbling cryogens, and by condensation of contaminants⁴².

References:

- ¹H. K. Wickramasinghe, AIP Conference Proceedings, **241**, 9 (1991).
- ²G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel, Phys. rev. Lett. **49**, 57 (1982).
- ³M. Nonnenmacher and H. K. Wickramasinghe, Appl. Phys. Lett. **61**, 168 (1992).
- ⁴P.K. Hansma, B. Drake, O. Marti, S. A. C. Gould, and C. B. Prater, Science **243**, 641 (1989).
- ⁵C. B. Prater, P. K. Hansma, M. Tortonese, and C. F. Quate, Rev. Sci. Instr. **62**, 2634 (1991).
- ⁶J. K. Gimzewski, J. K. Sass, R. R. Schlitter, and J. Schott, Europhys. Lett. **8**, 435 (1989).
- ⁷M. Volcker, W. Krieger, H. Walther, Phys. Rev. Lett. **66**, 1717(1991).
- ⁸G Binnig, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
- ⁹H. J. Mamin, D. Rugar, J. E. Stern, B. D. Terris, and S. E. Lambert, Appl. Phys. Lett. **53**, 1563 (1988).

10 J. E. Stern, B. D. Terris, H. J. Mamin, and D. Rugar, *Appl. Phys. Lett.* **53**, 2717 (1988).

11 G. Binnig, Ch. Gerber, E. Stoll, T. R. Albrecht, and C. F. Quate, *Europhys. Lett.* **3**, 1281 (1987).

12 A. L. Weisenhorn, P.K. Hansma, T. R. Albrecht, and C. F. Quate, *Appl. Phys. Lett.* **54**, 2651 (1989).

13 A. L. Weisenhorn, F.-J. Schmitt, W. Knoll, and P. K. Hansma, P.K., *Ultramicroscopy* **42-44.B**, 1125 (1992).

14 G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).

15 H. J. Mamin, E. Ganz, D. W. Abraham, R. E. Thomson, and J. Clarke, *Phys. Rev. B* **34**, 9015 (1986).

16 F. J. Giessibl, C. Gerber, and G. Binnig, *J. Vac. Sci. Technol. B* **9**, 984 (1991).

17 See for example H. L. Taitly, "Transducers in Mechanical and Electronic design," Dekker, New York, 1986.

18 G. Neubauer, S. R. Cohen, G. M. McClelland, D. Horne, and C. M. Mate, *Rev. Sci. Instr.* **61**, 2296 (1990).

19 W. S. Griffin, H. H. Richardson, and S. Yamanami, *J. Basic Engineering, Trans. ASME* **88**, 451 (1966).

20 E. S. Kim, R. S. Muller, *IEEE Electron Device Lett.* **EDL-8**, 467 (1987).

21 R. Erlandsson, G. M. McClelland, C. M. Mate, and S. Chiang, *J. Vac. Sci. Technol. A6*, 266 (1988).

22Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **52**, 1103 (1988).

23D. Rugar, H. J. Mamin, and P. Guethner, *Appl. Phys. Lett.* **55**, 2588 (1989).

24G. J. Brakenhoff, H. T. M. van der Voort, E. A. van Spronsen, and N. Nanninga, *Scanning Microsc.* **2**, 33 (1988).

25D. L. Taylor and E. D. Salmon, *Meth. Cell Biol.* **29**, 207 (1989)

26Atomic resolution TEM

27J. J. Bozzola, "Electron microscopy: principles and techniques for biologists," Jones and Bartlett, Boston, 1992.

28A. W. Robards and U. B. Sleytr, "Low Temperature Methods in Biological Electron Microscopy," Elsevier, Amsterdam, 1985.

29L. B. Da Silva, J. E. Trebes, S. Mrowka, T. W. Barbee, Jr., J. Brase, J. A. Koch, R. A. London, B. J. Macgowan, D. L. Matthews, D. Minyard, G. Stone, T. Yorkey, E. Anderson, D. T. Attwood, and D. Kern, *Optics Letters* **17**, 754 (1992).

30L. B. Da Silva, J. E. Trebes, R. Balhorn, S. Mrowka, E. Anderson, D. T. Attwood, T. W. Barbee Jr., J. Brase, M. Corzett, and J. Gray, *Science* **258**, 269 (1992).

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

32E. Henderson, P. G. Haydon, and D. S. Sakaguchi, *Science* **257**, 1944 (1992).

33G Binnig, Ch. Gerber, E. Stoll, T. R. Albrecht, and C. F. Quate, *Europhys. Lett.* **3**, 1281 (1987).

34 T. R. Albrecht and C. F. Quate, *J. Appl. Phys.* **62**, 2599 (1987)

35 B.L. Blackford, M. H. Jericho, and P. J. Mulhern, *Scanning Microscopy*, **5** 907 (1991).

36 J. A. McCammon and S. C. Harvey, *Dynamics of proteins and nucleic acids* (Cambridge University Press, Cambridge, U.K., 1987).

37 D. Southworth, K. Fisher, and D. Branton, "Principles of Freeze-fracturing and Etching", in D. Glick and R. M. Rosenbaum, eds, "Techniques of Biochemical and Biophysical Morphology", vol. 2, Wiley Interscience, New York 1975, p. 247.

38 P. Maivald, H. J. Butt, S. A. C. Gould, C. B. Prater, B. Drake, J. A. Gurley, V. B. Elings, and P. K. Hansma, *Nanotechnology* **2**, 103 (1991).

39 N. J. Tao, S. M. Lindsay, and S. Lees, *Biophys. J.* **63**, 1165 (1992).

40 M. D. Kirk, T. R. Albrecht, and C. F. Quate, *Rev. Sci. Instr.* **59**, 833 (1988).

41 F. J. Giessibl, C. Gerber, and G. Binnig, *J. Vac. Sci. Technol. B* **9** 984 (1991).

42 C. B. Prater, M. R. Wilson, J. Garnæs, J. Massie, V. B. Elings, and P. Hansma, *J. Vac. Sci. Technol. B* **9**, 989 (1991).

Chapter II: Noise Optimization for the Beam Deflection Method

Background

Since its invention in 1985,¹ the SFM has rapidly found a wide variety of applications. In the repulsive mode, where the tip is in physical contact with the sample, atomic resolution has been achieved on crystalline surfaces.² In non-contact modes of operation the SFM has been used to measure many different tip-sample interactions, including the van der Waals force,³ electrostatic forces⁴ — with single carrier resolution⁵ — and magnetic forces⁶, including those generated by electron spin resonance.⁷ Many factors affect the resolution of SFM's, two of the most important being the noise level of the system used to detect the tip motion, which determines the precision with which weak forces can be detected, and the sharpness of the tip, which often is the limiting factor for the lateral resolution, especially in repulsive mode work on surfaces which are not atomically flat. I will discuss both of these issues — noise sensitivity in this chapter, and tip making techniques in chapter III.

In the vision of its inventors,⁸ the SFM employed a tunneling tip on the back side of the spring to detect its deflection. Since then, many different detection methods have been developed. Today, the most common are the optical beam deflection method^{9,10}, preferred in most cases for its simplicity and stability, and various interferometric techniques,^{3,11-14} which can achieve lower levels of shot noise and are often preferred for noise-critical applications

such as non-contact force studies. I will show that it is possible to improve the noise level of the beam deflection method substantially by a slight redesign of the spring geometry, so that this trade-off between simplicity and low noise may be unnecessary.

Figure 2-1 shows the geometry of an optical beam deflection SFM. The SFM tip is located at the end of a spring, typically a cantilever spring, so that the tip-sample forces cause deflections of this spring. A laser beam is reflected from a mirror area on the back side of the spring onto a two-segment photo diode. At the rest position of the spring, equal parts of the laser power hit the two diode segments. Any deflection of the spring causes an angular deflection of the outgoing beam, which moves laser power from one segment onto the other, giving rise to a differential output signal. This signal is thus a measure of the tip motion.

Already Amer⁹, in first describing the optical beam deflection technique, showed that its dominant noise contribution, the shot noise of the laser beam, is proportional to the ratio of the effective length of the spring to the length of the mirror, and this fact has since been pointed out again¹⁵. However, SFM springs commonly in use today, including those commercially available, are not designed with this parameter in mind. In this chapter, I will describe SFM springs which have been designed specifically to optimize this parameter. Noise measurements on the resulting springs are discussed in chapter IV.

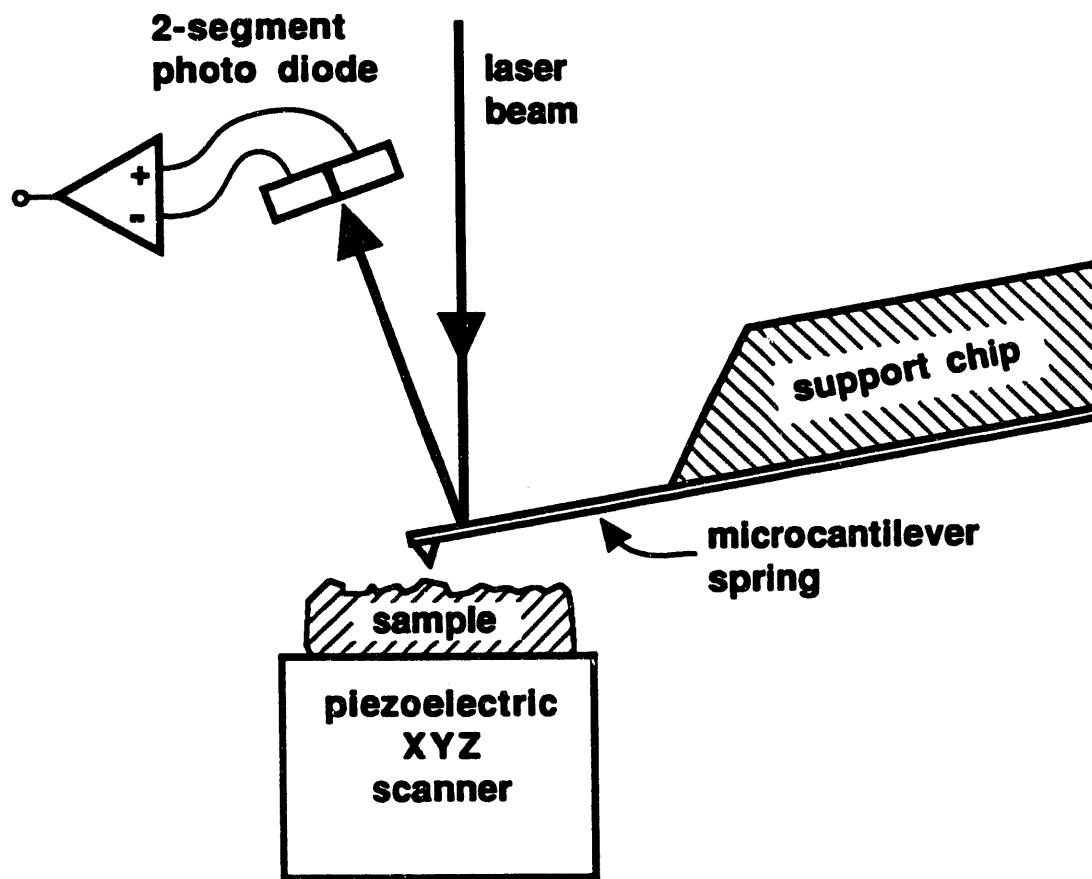


Figure 2-1. Principle of an "optical beam deflection" SFM. The SFM tip is located at the end of a spring, typically a cantilever spring, so that the tip-sample forces cause deflections of this spring. A laser beam is reflected off a mirror area on the back side of the spring, onto a two-segment photo diode. Any deflection of the spring causes an angular deflection of the outgoing beam, creating a differential output signal from the photo diodes.

Geometry

In normal operation of an optical beam deflection SFM, the dominant source of noise on the spring deflection signal is photon shot noise: the random fluctuations in the number of photons hitting the two diode segments. If in a given time interval $n/2$ photons are expected to impinge on each segment, the actual number will have a standard deviation of $(n/2)^{1/2}$, uncorrelated between the segments, so the difference has a standard deviation of $n^{1/2}$. The resultant noise on the differential diode current has a white spectrum, with the rms noise density¹⁶

$$N_i = (2eSP_r)^{1/2} , \quad (A/\sqrt{Hz}) \quad (1)$$

where P_r is the total laser power reflected from the spring, S is the responsivity of the photo diodes (in Amperes/Watt) and e is the electronic charge. What we want to minimize, however, is not the detector current noise N_i , but the tip position noise N_z , which is given by

$$N_z = N_i/(di/dz) . \quad (2)$$

To understand what determines (di/dz) , consider Fig. 2-2. If the tip is pushed back a distance dz , the angle of the mirror part of the spring will change by an amount $d\alpha = dz/r$, where r , defined by this relation, is the "effective length" of the spring. (For example, r equals the physical length L for an idealized spring which is hinged at the base and otherwise rigid, while the value of r for a

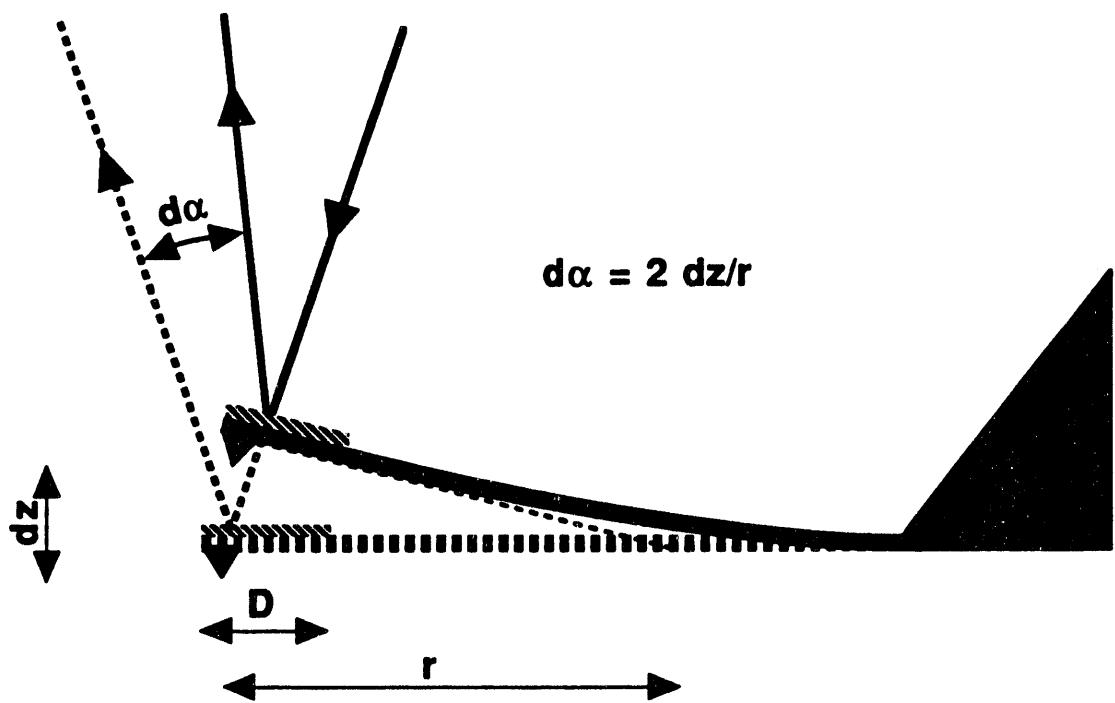


Figure 2-2. Geometry for noise optimization. When the tip is pushed back a distance dz , the angle of the mirror part of the spring changes by an amount $d\alpha$, and the angle of the reflected laser beam by $2d\alpha$. The quantity $r = dz/d\alpha$ is referred to as the "effective length" of the spring. The length D of the mirror sets the lower limit to the angular width of the reflected beam by diffraction, and the ratio D/r can be shown to determine the sensitivity of the method.

rectangular spring with a small mirror area at its end is $2L/3$). The reflected laser beam will deflect by an angle $2d\alpha$, moving an amount $dP = 2d\alpha\phi(0)$ of laser power from one diode segment to the other, where $\phi(0)$ is the power density per unit angle at the center of the reflected beam. A high $\phi(0)$ means a narrow beam — we can define an effective angular width δ of the beam by $\phi(0) = P_r / \delta$. The corresponding differential current signal is

$$di = 2S dP = 4S d\alpha \phi(0) = \frac{4S}{r} \phi(0) dz = \frac{4SP_r}{r\delta} dz. \quad (3)$$

Thus, to maximize the signal, we want a short spring (low r) and a narrow reflected beam (small δ) (and, of course, high beam power P_r). Unfortunately, these wishes are in opposition - a short spring necessitates a small mirror, which makes the angular width of the reflected beam large because of diffraction. For example, if the usable mirror area of the spring has a length D in the spring direction, the narrowest possible reflected beam is the one created by a uniform illumination, which has an angular profile

$$\phi(\alpha) = P_r \frac{D}{\lambda} \text{sinc}^2\left(\frac{\pi D}{\lambda} \alpha\right), \quad (4)$$

[where $\text{sinc}(x) = \sin(x)/x$ and λ is the laser wavelength] in the far field limit. (The far field is the relevant regime: it is favorable to work there, and easy to do so. In any SFM using microfabricated springs, one does not have to go further than a couple of tens of mm

from the spring to be in the far field regime.) In particular, $\phi(0) = P_r \cdot D / \lambda$, so the signal becomes [using (3)]

$$dI = \frac{4S}{r} \phi(0) dz = \frac{4S P_r}{\lambda} \left(\frac{D}{r} \right) dz. \quad (5)$$

Thus, and this is the main point of this example, the relevant geometry parameter in designing an SFM spring for low shot noise is the ratio D/r of mirror length to effective length, i. e. a *large fraction of the spring length should be used for the mirror*. Typical SFM springs being used now have only a small, triangular area near the tip available for use as a mirror.

A second point to note is that the distance between the spring and the detector does not enter at all. This is quite natural since the appropriate regime is the far field, where the reflected beam is conical and all relevant quantities are angular. Unfortunately, this point is often obscured by the usage of the term "optical lever" for the beam deflection method, which implies that the "leverage" gained by a large ratio of detector distance to spring length would somehow lower the noise, which as we see is not the case.

Optimization of Focusing

In order to derive a usable expression for the noise level, we have to be a little more careful. Expression (4) for the signal is somewhat optimistic because of the assumption of uniform illumination. It is impossible to generate an incident laser beam with perfectly uniform, sharp-edged profile, so uniform illumination cannot be

achieved without wasting a large part of the incident laser power "over the edge" of the mirror area. This loss of light would decrease the signal [by diminishing P_r in (4), which is the power reflected off the mirror, not the total power P of the incident beam]. A more realistic assumption is that we are given a laser beam of fixed power P and with a gaussian profile (the light emerging from a single-transverse-mode laser or single-mode optical fiber is typically gaussian to a good approximation). One then has to determine how to focus this beam onto the mirror so as to maximize the signal-to-noise ratio. To approach this optimization problem, we first need some background from gaussian optics.

A gaussian beam of power P has a radial intensity profile¹⁷

$$I(\rho) = I_0 \exp\left(-\frac{2\rho^2}{w^2}\right), \quad (6)$$

where

$$I_0 = 2P/\pi w^2$$

and ρ is the radial distance from the center of the beam. The quantity w , the radius at which the intensity has decreased by a factor e^2 from its value at the center, will be referred to as the " e^2 radius" of the beam. This e^2 radius varies along the beam path: for a beam which is focused "to a point", $w(z)$ is given by

$$w(z) = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{1/2}, \quad (7)$$

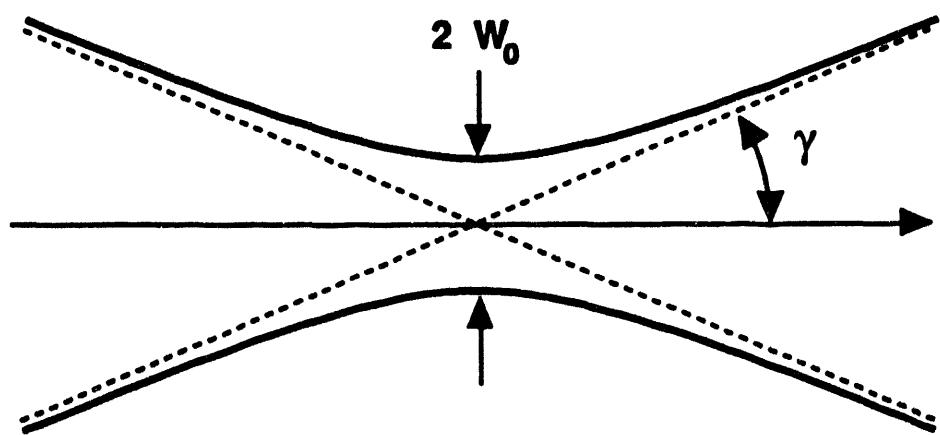


Figure 2-3. Gaussian optics. When a gaussian laser beam is focused "to a point", the actual minimum e^2 radius w_0 reached at the waist is related to the far field cone angle γ by $w_0 = \lambda/\pi\gamma$.

i. e. it decreases (linearly in the region far from the focal point), reaches a minimum w_0 at the beam "neck" and then increases again (also linearly in the far field) (Fig. 2-3). In the far field on both sides, the beam is conical, and thus can be described by an *angular profile*. This angular profile is also gaussian, given by an expression similar to (5) with the place of w taken by an angular e^2 half-width γ . The neck radius w_0 is inversely proportional to γ ; explicitly

$$w_0 = \lambda/\pi \gamma . \quad (8)$$

This cone angle γ , and thus w_0 , can be chosen freely by matching the focal length of the focusing lens appropriately to the radius of the source beam.

If we now introduce our finite-size mirror (of length D) at the neck of the beam, it will truncate the tails of the beam profile, effectively acting as an aperture. If w_0 is chosen much larger than D , the uniform illumination limit [(4), (5)] can be approached, but at the cost of losing most of the laser power. If it is chosen much smaller than D , the full incident power is reflected, but the angular width of the reflected beam, now given by γ , becomes unnecessarily large. We want to find the optimum value of w_0 so as to minimize the shot noise referred to the spring motion, N_z . Using (1), (2) and (3), we get

$$N_z = N_i \left(\frac{di}{dz} \right)^{-1} = (2eSP_r)^{1/2} \left(\frac{4S}{r} \phi(0) \right)^{-1} \propto \frac{P_r^{1/2}}{\phi(0)} . \quad (9)$$

We will first calculate the two-dimensional angular power profile $\phi(\alpha, \beta)$ (the angles being measured from the center of the beam in the x and y direction, respectively) in the far field, generated by our truncated gaussian beam profile at the mirror; then $\phi(\alpha)$ and P_r are found as

$$\phi(\alpha) = \int_{-\infty}^{\infty} \phi(\alpha, \beta) d\beta \quad (10)$$

and

$$P_r = \int_{-\infty}^{\infty} \phi(\alpha) d\alpha \quad (11)$$

The electric field strength at a point in the far field at angles α , β and distance R from the mirror is related to the field strength at the mirror by the Fraunhofer diffraction integral¹⁸

$$E(\alpha, \beta, R) = \frac{i}{\lambda R} \iint_{\text{mirror}} E(x, y) e^{\frac{2\pi i}{\lambda} (\alpha x + \beta y)} dx dy$$

so that the power density per unit solid angle in the far field is

$$\phi(\alpha, \beta) = \frac{1}{\lambda^2} \left| \iint_{\text{mirror}} \sqrt{I(x, y)} e^{\frac{2\pi i}{\lambda} (\alpha x + \beta y)} dx dy \right|^2, \quad (12)$$

where $I(x, y)$ is the power density per unit area at the mirror. Using (6) with $w = w_0$, and assuming a square mirror, we find

$$\phi(\alpha, \beta) = \frac{I_0}{\lambda^2} \left| \int_{-D/2}^{D/2} \int_{-D/2}^{D/2} e^{-(x^2+y^2)/w_0^2} e^{\frac{2\pi i}{\lambda}(\alpha x + \beta y)} dx dy \right|^2. \quad (13)$$

What we want to minimize is the z noise, given by (9),

$$N_z \propto \frac{P_f^{1/2}}{\phi(0)} = \frac{\left[\iint \phi(\alpha, \beta) d\alpha d\beta \right]^{1/2}}{\int \phi(0, \beta) d\beta}. \quad (14)$$

Luckily, $\phi(\alpha, \beta)$ factorizes into separate x and y integrals:

$$\phi(\alpha, \beta) = \frac{I_0}{\lambda^2} f(\alpha) f(\beta), \quad (15)$$

where $f(\xi) = \left| \int_{-D/2}^{D/2} e^{-x^2/w_0^2} e^{(2\pi i/\lambda) \xi x} dx \right|^2. \quad (15a)$

This factorization greatly simplifies (14):

$$\frac{P_f^{1/2}}{\phi(0)} = \frac{\left[\iint \phi(\alpha, \beta) d\alpha d\beta \right]^{1/2}}{\int \phi(0, \beta) d\beta} = \frac{\lambda}{I_0^{1/2}} \frac{\left[\int f(\alpha) d\alpha \int f(\beta) d\beta \right]^{1/2}}{f(0) \int f(\beta) d\beta} = \frac{\lambda}{f(0) I_0^{1/2}}, \quad (16)$$

since the two integrals in the square root sign are identical (under the assumption of a square mirror).

Thus all we need to do is maximize $f(0) I_0^{1/2}$. From (16) and (6), we find

$$I(0)I_0^{1/2} = \left(\frac{2P}{\pi}\right)^{1/2} \frac{1}{w_0} \left(\int_{-D/2}^{D/2} e^{-x^2/w_0^2} dx \right)^2. \quad (17)$$

Introducing the dimensionless variables $a = D/2w_0$ and $s = x/w_0$, (17) becomes

$$I(0)I_0^{1/2} = \left(\frac{2P}{\pi}\right)^{1/2} \frac{D}{2} \left[\frac{1}{a} \left(\int_{-a}^a e^{-s^2} ds \right)^2 \right]. \quad (18)$$

The quantity in square brackets is a real function $h(a)$ of a , and can be maximized numerically: it reaches a maximum value of $h_{\text{opt}}=2.231$ at $a_{\text{opt}}=0.990$. Thus the optimal e^2 radius of the incoming beam is $w_0 = D/2a_{\text{opt}} = 0.505 D$. Essentially, the e^2 diameter of the beam waist should equal the mirror size.

Now we can find the minimum z noise by substituting h_{opt} back through (18), (16) and (9).

$$N_{z_{\text{opt}}} = r \left(\frac{e}{8S}\right)^{1/2} \frac{\lambda}{2.23 (D/2)(2P/\pi)^{1/2}} = 0.397 \frac{\lambda r}{D} \left(\frac{e}{SP}\right)^{1/2}. \quad (19)$$

This result was derived for the case of a square mirror and circular incident beam. There is another case which is of interest, namely when the beam is focused much more tightly — or, equivalently, the mirror is much wider — in the y than in the x direction. This can for instance be the case if the highly elliptical

beam from a laser diode is used directly. In this case, the y integrals in the above expressions are carried to infinity, and one finds through a similar calculation that the optimum e^2 radius is 0.567 D, and the numerical factor in (19) is improved slightly to 0.394.

Another implicit assumption we have made so far is that the width of the slit in the photo diode is negligible. Interestingly, this is not the ideal situation. With a finite slit width s , the signal is proportional to the beam intensity $I(s/2)$ at the edges of the slit, while the noise is proportional to the square root of the total power falling on the diodes, i.e. to

$$\left[2 \int_{s/2}^{\infty} I(x) dx \right]^{1/2}.$$

The beam intensity $I(s/2)$ decreases only quadratically with s near the center of the beam, while the above integral decreases linearly with s . Thus the noise benefit from the diminished total power falling on the diodes is greater than the signal loss due to the decreased power at the edges of the slit. In fact, the optimum width of the slit is quite large. For a gaussian beam of e^2 radius W at the detector, the optimum slit width is 0.612 W , and the noise at this point is lower than for a zero-width slit by a factor of 0.887.

Comparison to Interferometry

We now want to compare the noise level, (19), of the beam deflection method to that of fiber interferometry. The geometry of

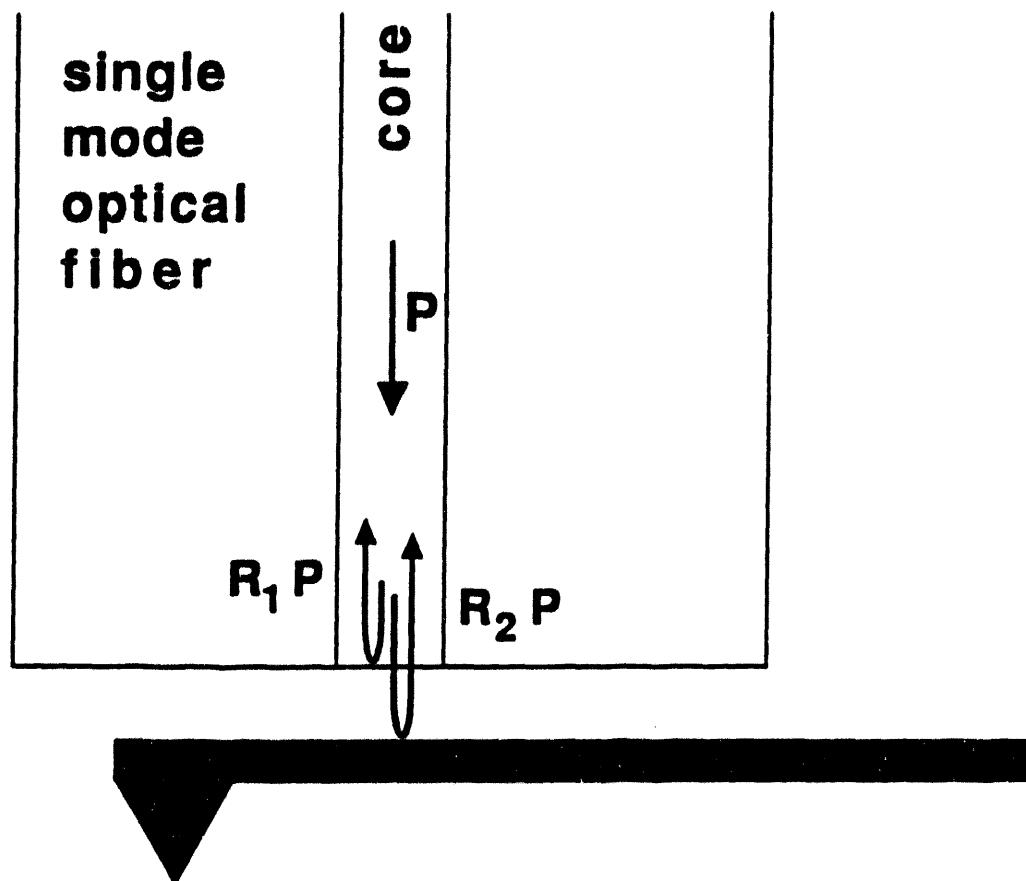


Figure 2-4. Geometry of fiber interferometric SFM. A single mode optical fiber is brought into close proximity to the SFM spring. A fraction $R_1 P$ of the incident laser power is reflected at the fiber endface. The remaining light exits the fiber and is reflected off the spring, and some amount $R_2 P$ reenters the fiber core. Since there is only one mode available in the fiber, the resulting reflected power depends on the interference between these two components, which in turn depends on the fiber-spring distance.

a fiber interferometric SFM is shown in Fig. 2-4. The end of a single-mode optical fiber is brought into close proximity to the reflective back side of the SFM spring. Some fraction $R_1 P$ of the incoming light power P is reflected at the fiber end face, while some amount $R_2 P$ of the remainder reenters the fiber core after being reflected off the spring. Since there is only one back-propagating mode available in the fiber, the resultant reflected power is determined by the interference of these two components. This depends on the phase difference between them, which is a measure of the spring position.

R_1 is determined by the index of refraction of the glass, and is typically¹⁹ about 0.04, while R_2 varies depending of the distance between the fiber and the spring. The total power returning along the fiber is

$$P_r = [R_1 + R_2 + 2(R_1 R_2)^{1/2} \sin(4\pi z/\lambda)] P . \quad (20)$$

The shot noise N_i in the photo detector is still given by (1), while the signal is now

$$di/dz = S (dP_r/dz) = (8\pi SP/\lambda) (R_1 R_2)^{1/2} , \quad (21)$$

assuming that the fiber position has been adjusted to make use of the steepest point of the interference pattern. At this point, the sine in (20) vanishes, so $P_r = (R_1 + R_2)P$ and the z noise is

$$N_z = N_i (di/dz)^{-1} = [2eS (R_1 + R_2)P]^{1/2} [(8\pi SP/\lambda) (R_1 R_2)^{1/2}]^{-1}$$

$$= \frac{\lambda}{\pi} \left(\frac{e}{32SP} \frac{R_1+R_2}{R_1R_2} \right)^{1/2} = \frac{\lambda}{\pi} \left[\frac{e}{32SP} \frac{1}{R_1} \left(1 + \frac{R_1}{R_2} \right) \right]^{1/2} . \quad (22)$$

We see that R_2 should be chosen as large as possible, but that it has little influence as long as $R_2 \gg R_1$. (Note that this is different from the condition $R_2 \approx R_1$ implied to be optimal in the original paper on this technique.¹⁹⁾ For uniqueness, we will consider the maximum possible value of R_2 , which (ignoring multiple reflections) is $(1-R_1)^2$. Then

$$N_z = 0.287 \lambda \left(\frac{e}{SP} \right)^{1/2} \quad (23)$$

for $R_1 = 0.04$. Comparing (23) with (19) we see that

$$\frac{N_z(\text{beam deflection})}{N_z(\text{interferometry})} = 1.38 \frac{r}{D} , \quad (24)$$

so that if the ratio D/r of the mirror length to the effective length can approach unity, the noise level of the beam deflection method can become comparable to that of fiber interferometry. In fact, in the next section we will show springs for which the effective D/r exceeds 1.5, making the beam deflection method potentially quieter.

Spring Design

A typical modern SFM spring²⁰ is microfabricated from a thin film of silicon or silicon nitride, is 0.6-2 μm thick, 100-200 μm long, and has a V shape as shown in Fig. 2-5 (a). The film forms a

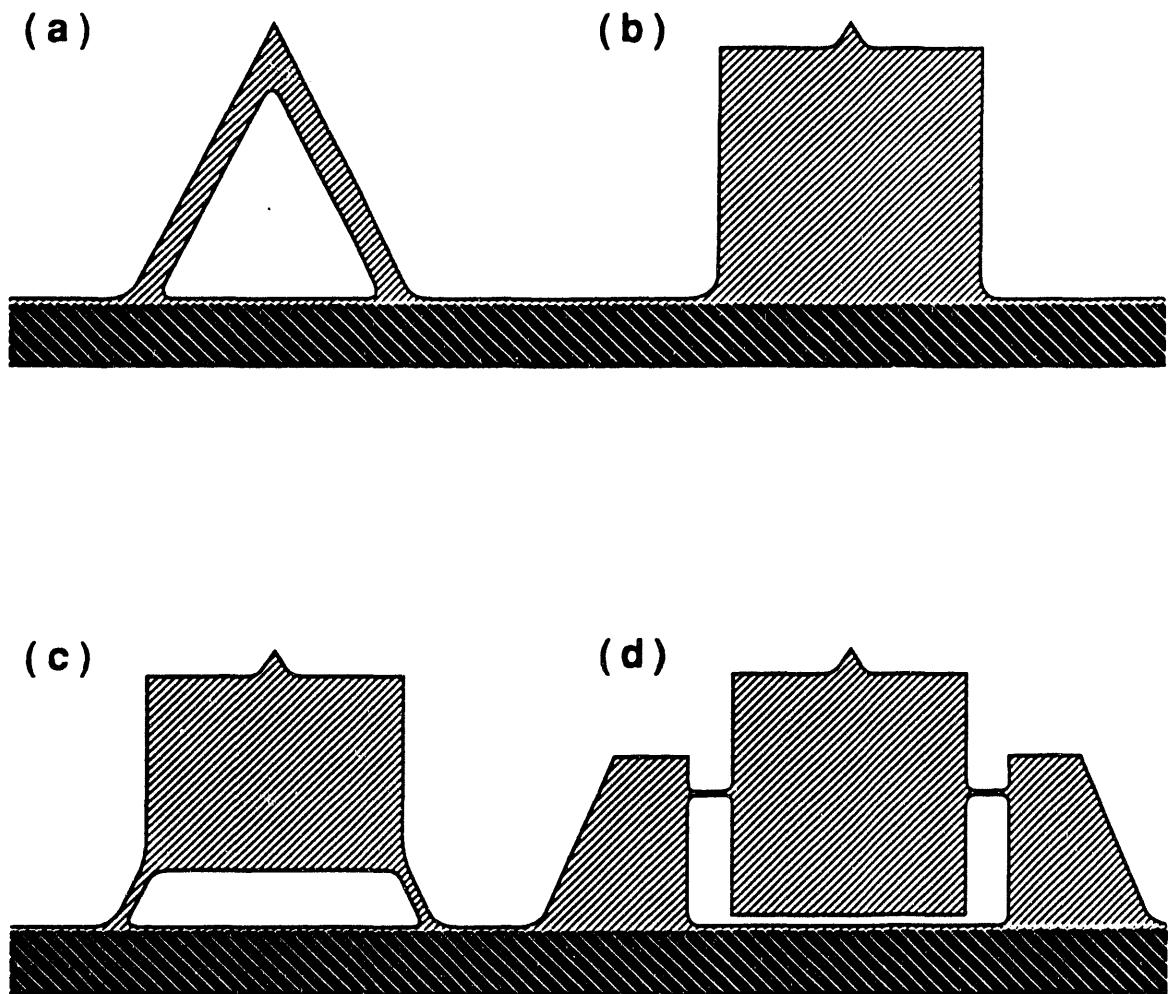


Figure 2-5. Possible spring designs. (a) The common "V" design. (b) An attempt to maximize the D/r ratio (see Fig. 2-2). Apart from increasing the force constant excessively, this idea does not work as well as expected because the mirror bends when the spring deflects. (c) Here the mirror area is relatively rigid, and separated from the base by weak "legs". The D/r is still a considerable improvement over (a), and the force constant can be adjusted by altering the leg dimensions. (d) With a torsion spring design, the D/r can exceed unity.

cantilever spring extending from the edge of a silicon or glass chip. The resulting force constant is in the range 0.05-10 N/m depending on application. The V shape increases the lateral stiffness. When used for the optical beam deflection method, the back side of the spring is metallized, and the small, triangular area near the tip is used as the mirror.

This shape can be adapted to increase the D/r ratio, and thus improve the shot noise. The most obvious approach is to extend the mirror all the way from the tip to the base, forming a solid rectangle [Fig. 2-5(b)]. Apart from increasing the force constant considerably, which may be undesirable, this idea fails to achieve the expected D/r of unity, because the mirror is allowed to bend. The noise expression (19) was derived assuming a rigid mirror, and a bending mirror yields a larger noise — effectively, the parts of the mirror which are near the base do not contribute to the signal since their angle changes very little when the spring is deflected. Instead, it is more effective to separate a more or less rigid "mirror plate" from the support by flexible "legs", where most of the bending takes place [Fig. 2-5(c)]. This also allows the spring constant to be adjusted by altering the leg width.

It has been accepted since the invention of the SFM that the appropriate spring geometry is a cantilever — so accepted, in fact, that SFM springs are uniformly referred to as "cantilevers". However, there are other possible geometries that should be considered, some of which may be superior for certain applications. One such option is the torsion spring, Fig. 2-5(d). In this design, the effective length is the distance from the tip to the line between the

suspension wires, so the mirror size can actually exceed the effective length, and by close to a factor of two.

An unrelated advantage of the torsion spring concept is that the spring can be suspended symmetrically with respect to its center of mass, so that lateral vibrations of the support do not couple to the main, torsional mode of the spring. The lateral modes of the suspension do not affect the signal, and can be made much stiffer than the main mode.

A possible disadvantage of torsion springs is that their lateral force constants, while much stiffer than the main mode, are still softer than those of V-shaped cantilevers. This makes unwanted stick-slip behavior more likely in repulsive-mode work on "sticky" samples. Though this is not the shot noise limited regime for which these springs are intended, we have successfully tested torsion springs on biological samples in the repulsive mode.

References:

- 1 G Binnig, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
- 2 G Binnig, Ch. Gerber, E. Stoll, T. R. Albrecht, and C. F. Quate, Europhys. Lett. **3**, 1281 (1987).
- 3 Y. Martin, C. C. Williams, and H. K. Wickramasinghe, J. Appl. Phys. **61**, 4723 (1987).
- 4 Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, Appl. Phys. Lett. **52**, 1103 (1988).
- 5 C. Scönenberger and S. F. Alvarado, Phys. Rev. Lett. **65**, 3162 (1990).

6 J. J. Sáenz, N. Garcia, P. Grütter, E. Meyer, H. Heinzelmann, R. Weisendanger, L. Rosenthaler, H. R. Hidber, and H.-J. Günterrod, J. Appl. Phys. **62**, 4293 (1987).

7 D. Rugar, C. S. Yannoni and J. A. Sidles, Nature **360**, 563 (1992).

8 G Binnig, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

9 G. Meyer and N. M. Amer, Appl. Phys. Lett. **53**, 1045 (and 2400, erratum) (1988).

10 S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, and J. Gurley, J. Appl. Phys. **65** 164 (1989).

11 R. Erlandsson, G. M. McClelland, C. M. Mate, and S. Chiang, J. Vac. Sci. Technol. **A6**, 266 (1988).

12 D. Rugar, H. J. Mamin, R. Erlandsson, J. E. Stern, and B. D. Terris, Rev. Sci. Instr. **59**, 2337 (1988).

13 D. Rugar, H. J. Mamin, and P. Guethner, Appl. Phys. Lett. **55**, 2588 (1989).

14 C. Schönenberger and S. F. Alvarado, Rev. Sci. Instr. **60**, 3131 (1989).

15 C. A. J. Putman, B. G. de Groot, N. F. van Hulst, J. Greve, J. Appl. Phys. **72**, 6 (1992).

16 P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., Cambridge University Press, Cambridge, U.K., 1989, p. 432.

17 R. Guenter, *Modern Optics*, John Wiley & sons, New York, 1990, p. 341.

18 R. Guenter, *Modern Optics*, John Wiley & sons, New York, 1990, p. 349.

19 D. Rugar, H. J. Mamin, R. Erlandsson, J. E. Stern, and B. D. Terris,
Rev. Sci. Instr. **59**, 2337 (1988).

20 T. R. Albrecht and C. F. Quate, J. Appl. Phys. **62**, 2599 (1987)

Chapter III: Microfabrication of Springs

Background: Fractured Tips

This chapter describes the practical fabrication of cantilever and torsion springs following the designs described in the previous chapter. In addition to the enlarged mirror area discussed there, these springs incorporate another new feature: a unique process for creating very sharp scan tips, which we call "controlled fracture." Before describing this technique, I will give some brief motivation.

While it is clearly desirable in SFM to have as "sharp" a tip as possible, the appropriate interpretation of "sharp" differs somewhat between different applications (Fig. 3-1), even if one restricts the discussion to the repulsive mode. In atomic resolution imaging of crystals, the relevant property is the presence of atomic scale asperities on the tip, while the macroscopic shape of the tip is unimportant [Fig. 3-1 (a)]. When one images samples with large, steep features, the final opening angle of the tip is the most important parameter [Fig. 3-1 (b)]. In the intermediate regime, where the size scale of the sample features are comparable to the radius of curvature of the tip itself, the appropriate figure of merit is this macroscopic tip radius [Fig. 3-1 (c)]. This last regime is the one addressed by the tip making process described in this paper.

The great majority of SFM work today uses microfabricated springs¹, because of their high resonance frequencies and the fact that large numbers of them can be manufactured in parallel. Tips have been fabricated on such springs by many techniques, including

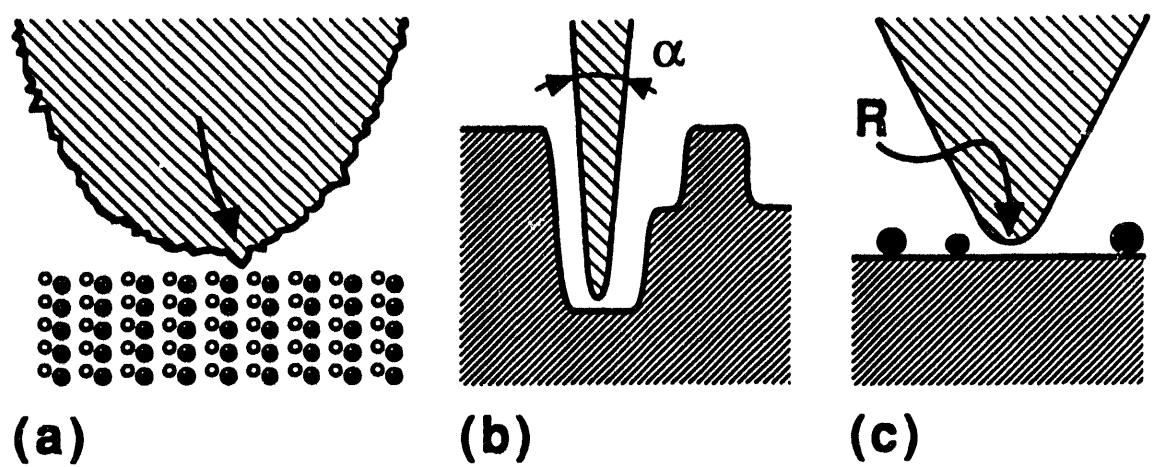


Figure 3-1. Tip sharpness criteria. For atomic resolution imaging of crystals (a), the relevant property is the presence of atomic scale asperities on the tip (arrow), while the macroscopic shape of the tip is unimportant. When one is imaging samples with large, steep features (b), the final opening angle of the tip is the most important parameter. In the intermediate regime (c), where the size scale of the sample features are comparable to the radius of curvature of the tip itself, the appropriate figure of merit is this macroscopic tip radius. Regime (c) is the one addressed by the tip making process described in this paper.

glued-on diamond shards,² photolithographically defined corners,¹ metal cones evaporated through apertures,³ pyramids formed in anisotropically etched pits in silicon³, etched silicon posts⁴, and electron beam driven vacuum deposition.⁵⁻⁷ The diamond shards used in early work² are potentially atomically sharp, since they are defined by the intersection of fracture planes. However, the mass of any shard that can be handled, even with micromanipulators, greatly exceeds that of a microfabricated SFM spring, so the addition of the shard (plus adhesive) seriously lowers the resonance frequency of the spring. For this reason, and because of the amount of labor involved in their assembly, the diamond shard tips have fallen out of favor. The other tip making methods all suffer from some fundamental limitation to the sharpness which can be achieved: The pyramid method is limited by the nonzero etch rate on the Si₃N₄ tip of the etchant used to remove the silicon, and by unavoidable asymmetry in the photolithography used to create the etch pit (since the tip is defined by the intersection of four, not three, planes). The etching methods are limited by the enhanced etch rate on highly curved surfaces near the tip, while the aperture-evaporation scheme by the finite angular width of the source, and the electron beam deposition by the finite radius of the beam. The aim of the technique described here is to combine the potentially atomic sharpness of intersecting fracture planes with the low weight of microfabrication. The approach used is to create microfabricated structures that can themselves be fractured in a reproducible way so as to form a tip.

Fabrication

The material of a fractured tip must be either single-crystalline or amorphous, since a fracture surface will not be plane on an atomic scale if grain boundaries are present. Luckily, the two most common materials used in microfabricated SFM cantilevers, single crystal silicon and amorphous⁸ LPCVD-deposited silicon nitride, both satisfy this condition. We have used the latter.

The fabrication steps before the actual fracturing are equivalent to the process used for the simplest form of SFM cantilevers without tips.³ Standard microfabrication procedures are used, so several thousand springs can be processed in parallel. Figure 3-2 outlines the process. A film of "low stress"⁹ silicon nitride is deposited on both sides of a silicon wafer, by low pressure chemical vapor deposition (LPCVD). After a hole is opened in the film on the reverse side of the silicon wafer with an SF₆/He plasma etch, the wafer is etched all the way through using an ethylene diamine / pyrochatecol / water / pyrazine mixture (EDP).¹⁰ The EDP is a highly anisotropic etchant, so the etch follows the (111) crystal planes, forming a clean and predictable hole. This step leaves the front side film freely suspended as a membrane over the hole. The front side is now patterned, using the visible edges of the hole to simplify alignment, and the springs are formed from the membrane with the plasma etch. (The membrane is only 250 μm wide, so it is possible to spread photoresist on top of it without breakage.) After the resist is removed, aluminum is evaporated onto the back side for reflectivity. (All structures described in this paper are strong

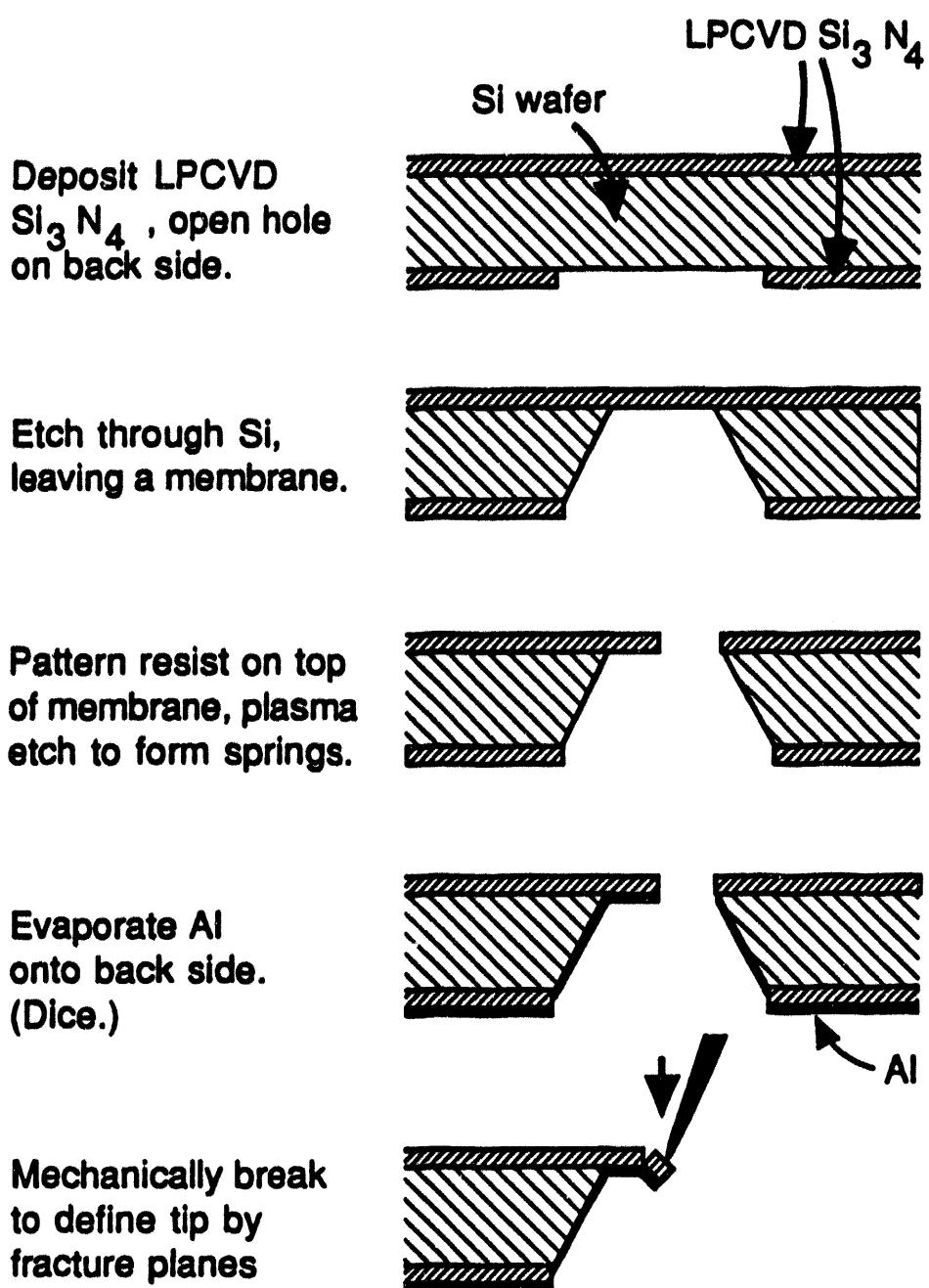


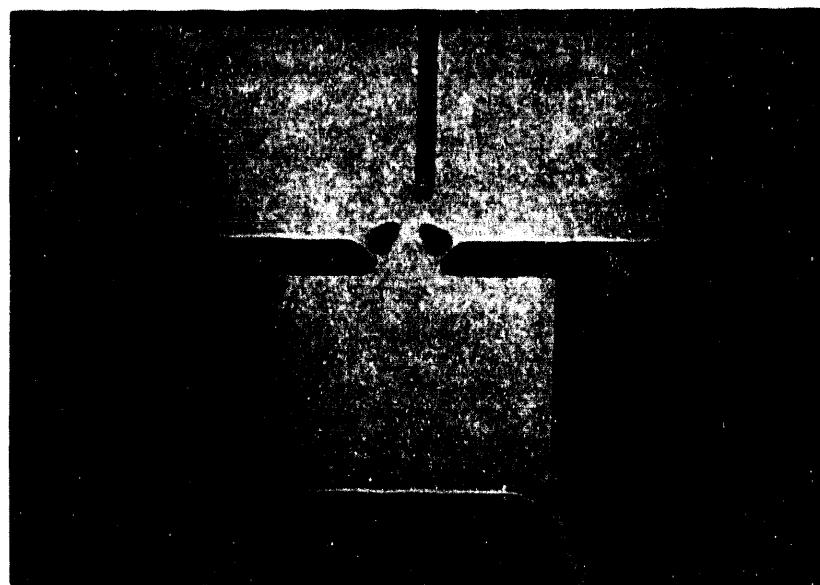
Figure 3-2. The microfabrication process. The steps before the last are equivalent to the simplest process used to make SFM springs without tips. (In fact, both kinds could be processed on the same wafers.) In the last step, special structures are mechanically broken off, leaving a corner-shaped tip defined by the intersection of fracture planes.

enough to survive being dipped into liquids. If very fragile structures are made, one can use an O₂ plasma to remove the resist.) The wafer is cleaved into dies, using cleavage guides formed at the same time as the holes, either before or after the fracturing step where the tips are created.

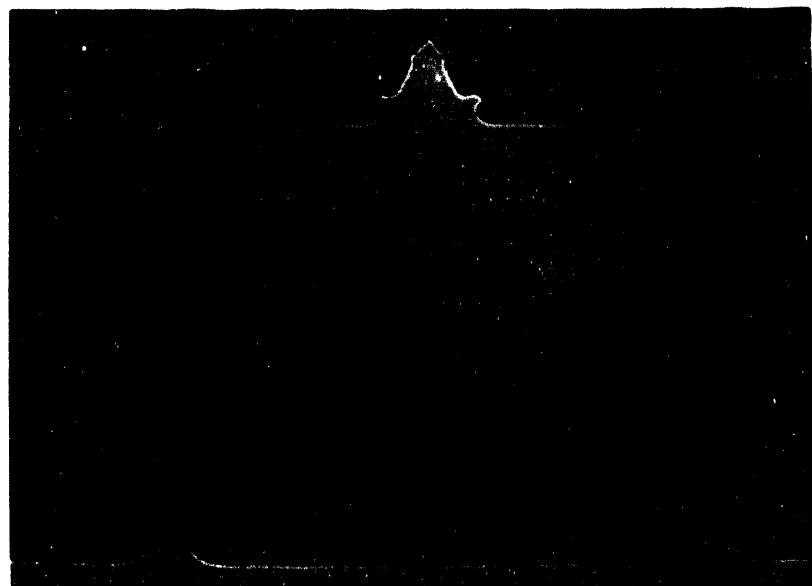
The appearance of the cantilevers at this point is shown in Fig. 3-3(a). It differs from the final shape [Figs. 3-3(b) and 3-5] by the addition of two "break plates". To create the tip, first the left, then the right plate are mechanically broken off using a micromanipulator probe. While this is done, the cantilever must be supported from the back side with another probe to avoid breaking the legs [Fig. 3-4]. The relatively long, narrow guide structures that connect the breakplate to the cantilever near the tip area are designed to remain intact until after the tip region breaks, and help guide the fracture. The fracture plane of the right break plate intersects that left by the left plate at right angles. The cube corner defined by these two intersecting fracture planes and the top surface of the film forms the tip. Since this tip does not protrude from the film, the spring is operated at a 20° angle to the sample surface. The surface of the film, while not a fracture plane, is flat to within 0.6 nm rms as far as can be measured with SFM.

The fracturing process is quite reproducible, with 90-95% of the resulting tips appearing perfect in an optical microscope. Once the manipulators are set up, the process takes only about one minute per tip.

A silicon nitride cantilever spring made this way is shown in Fig. 3-5. It has a total length of 50 μm, a mirror length of 30 μm, and an



(a) 10 μm —



(b) 10 μm —

Figure 3-3. Tip fracturing. The shape (a) of the cantilever spring as patterned differs from the final shape (b) by the addition of two "break plates." These are mechanically broken off, and the fracture plane of the right plate intersects that left by the left one so as to create a cube-corner-shaped tip.

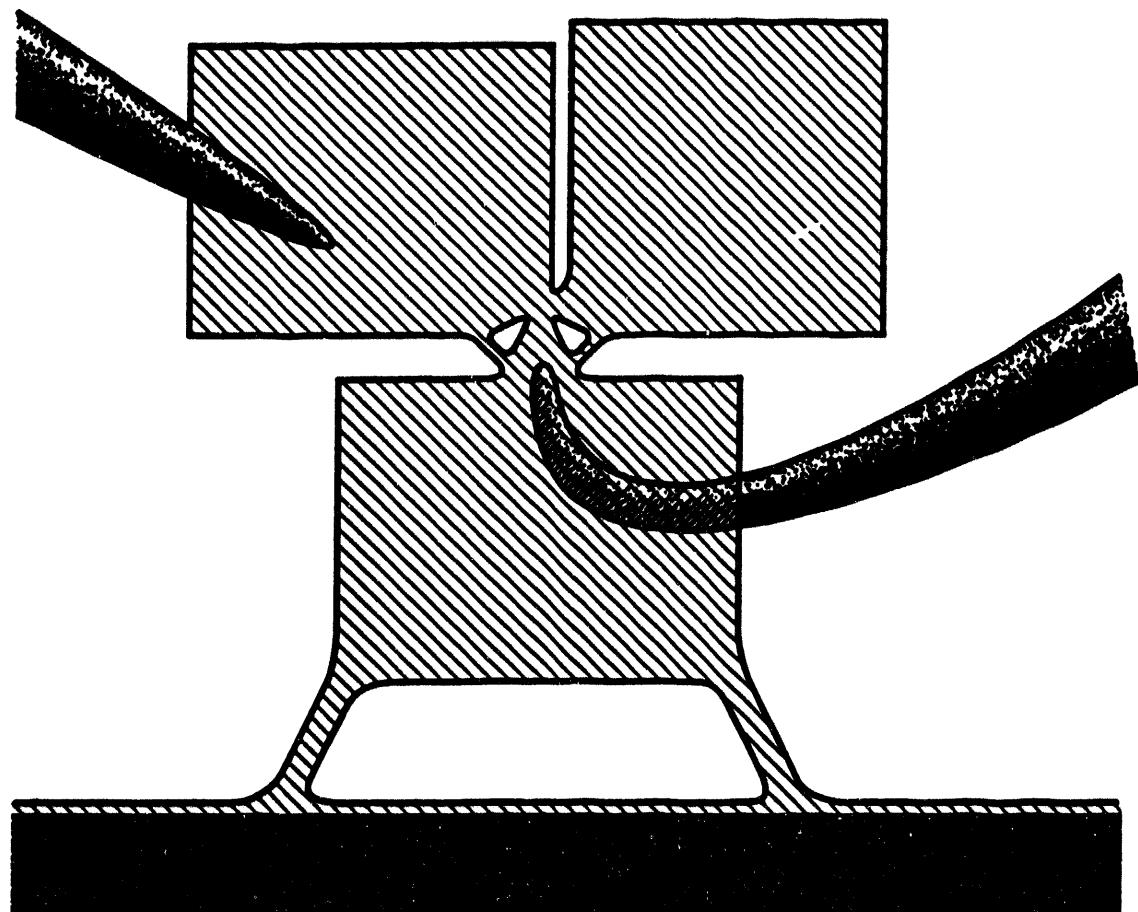
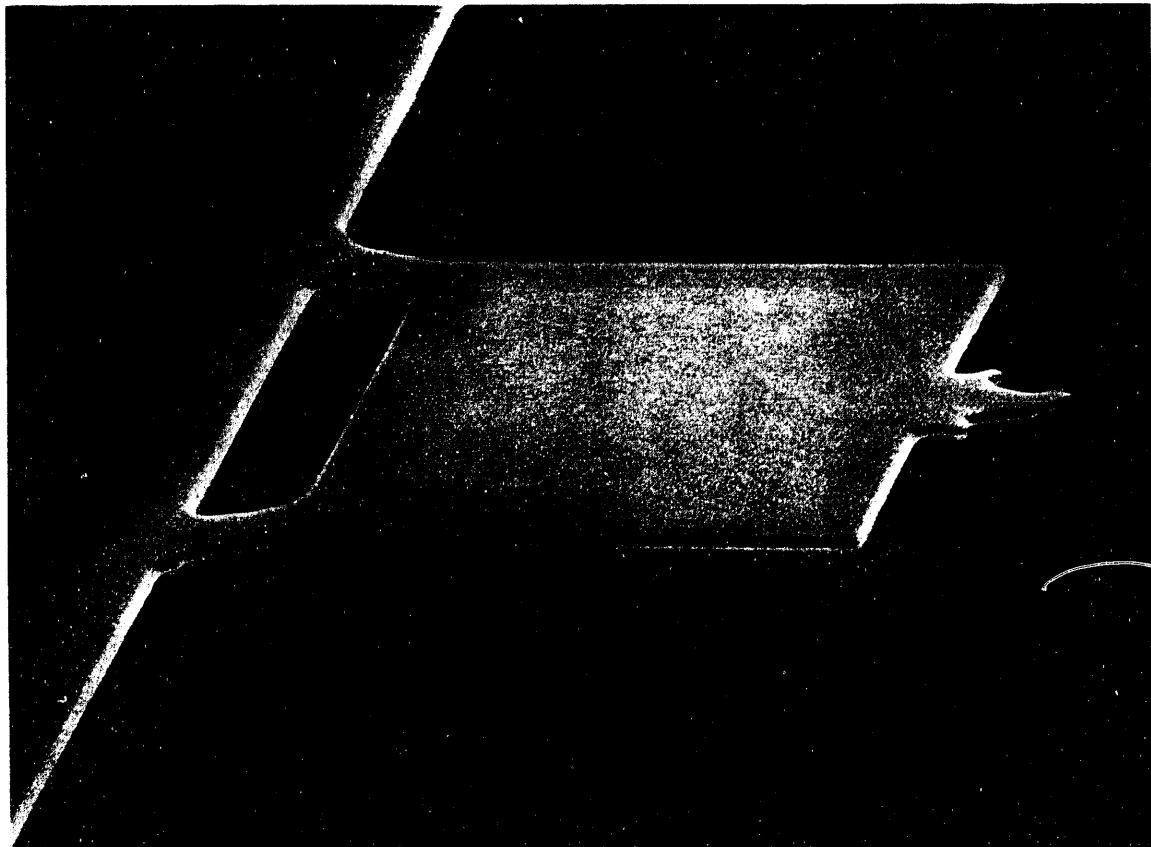


Figure 3-4. The break plates are fractured off using a micropositioner probe tip, while supporting the spring from the back side with another tip to avoid damaging the legs. The procedure takes only about one minute per tip once the probes are set up.



50 μm

Figure 3-5. A finished 50 μm long cantilever spring.

effective length of 44 μm (from the tip to the midpoint of the leg), yielding a D/r of 0.68. The leg width is 3 μm and the thickness is 0.7 μm , resulting in a force constant of 1.1 N/m and a measured resonance frequency of 201 kHz.

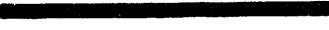
An example of a torsion spring is shown in Fig. 3-6(a). A 50 \times 50 μm mirror plate (with a fractured-tip structure extending from it) is suspended by 15 μm long silicon nitride wires. The cross-sectional dimensions of the suspension wires are 0.7 \times 0.7 μm , as can be seen from the broken wire in Fig. 3-6(b). These wires are surprisingly durable: we have deflected mirrors suspended by 20 μm long wires through 180° without breakage. The spring shown has a resonance frequency of 36 kHz and a force constant of 0.06 N/m. The mirror extends beyond the suspension point, thus allowing the D/r ratio to exceed unity. The spring in Fig. 3-6(a) has an effective length of 32 μm , yielding $D/r = 1.53$. (The D/r ratio is less than 2 only because of the extending fractured-tip structure, and could approach 2 for other tip types.)

The nearest higher detectable resonant mode of the spring shown is at 254 kHz, corresponding to a lateral force constant of 14 N/m — 240 times stiffer than the torsional mode — in agreement with theoretical calculations. In our experience, a certain vibration source (the pulsing of an inertial walker), which induced large vibrations in cantilever springs, was barely noticeable when a torsion spring was used.

Evaluation of Tips

SEM images of the tip are shown in Fig. 3-7. To avoid charging effects, the sample has been coated with a nominally 5 nm thick



(a) 50 μm 



(b) 1 μm 

Figure 3-6. (a) A torsion spring. The suspension is placed symmetrically through the center of gravity of the mirror plus tip structure, thus maximizing immunity to lateral vibrations. (b) A suspension wire which has been broken close to its base, showing the $0.7 \times 0.7 \mu\text{m}$ cross section.



(a)

5 μ m



(b)

100 nm



Figure 3-7. (a) The tip area after fracturing. (b) Closeup of the tip. As is seen, the overall cube corner shape is maintained at least to the 10 nm size scale. To take these SEM images, a nominally 5 nm conductive coating had to be applied to avoid charging effects. This coating, the electron beam spot size and the hydrocarbon accumulation effect together account for the observed tip radius.

conducting Pt-C film.¹¹ As seen, even though the fractures do not always follow perfect planes, the cube-corner shape of the tip is maintained at least down to the 10 nm size scale. Only an upper limit on the tip radius can be learned from such images, since the observed apparent radius of 10-15 nm is entirely explainable by the Pt-C film, the electron beam spot size, and the hydrocarbon accumulation artifacts due to pump oil contamination in the SEM chamber (this is the same effect which was used to make the first e-beam defined tips⁵). The image in Fig 3-7(b) was the first scan of this area at any magnification, and subsequent scans showed successively larger apparent tip radii because of hydrocarbon accumulation. We estimate 10 nm to be a conservative upper limit on the tip radius.

The performance of the fractured tips on biological samples can be judged from Fig. 3-8. The sample is type I collagen, a fibrous protein molecule 280 nm long and 1.5 nm in diameter, adsorbed to mica. (This sample will be described in more detail in chapter VIII.) It was imaged at 143 K under isopentane, using our low temperature SFM. We used a cantilever spring similar to the one in Fig. 3-5 (except 100 μ m long instead of 50) in the repulsive mode at a force of 250 pN. Each of the fibrous features seen in the image is a single collagen molecule. The apparent full width of the molecules in this and similar images is about 4-5 nm, corresponding to a tip radius of 1.3-2.1 nm.

Some caution has to be taken when evaluating the tip radius in this fashion, since the apparent height of the molecules in these images is less than the expected diameter. The above value for the

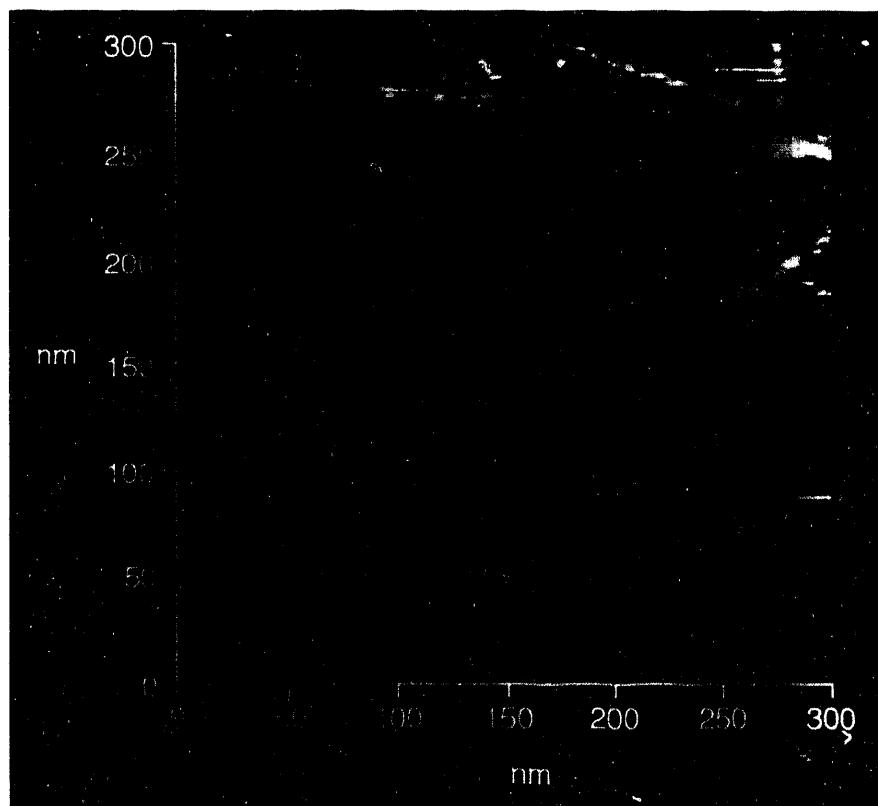


Figure 3-8. An SFM image acquired with a fractured tip. The sample is type I collagen, a 280 nm by 1.5 nm fibrous protein molecule, adsorbed to mica, imaged at 143 K under isopentane. From the apparent width of the collagen molecules in such images, 4-5 nm, we estimate a tip radius of 1.5-2 nm.

tip radius assumes that this is because of distortion by the tip force, and that this distortion is elastic, so that the molecule regains its shape as the tip is removed. Another possible interpretation is that the molecules are partially embedded in contamination, distorting their apparent height and width, in which case it is more difficult to draw any conclusions about the tip shape.

While the wide opening angle of the tips is not a limitation in images like Fig. 3-8, it does become apparent when tall and steep features such as steps are present [Fig. 3-9 (a)], and can cause spherical objects to take on a slightly triangular / pyramidal appearance [Fig. 3-9(b)].

Possible improvements

The fractured-tip process can be applied equally well to other materials. Single crystal silicon is a widely used material for SFM springs, and may be advantageous: it is likely to cleave easily and cleanly along crystal planes (if the wafer orientation is chosen appropriately), yielding sharp and predictable tips. It is, however, subject to oxidation in air, which is likely to round off an ideal tip to about a 2 nm radius. (Silicon nitride also oxidizes in air, though to a somewhat lesser extent.) The ideal material is probably single crystal diamond films, which may become available in the future.

The tips generated by the simple one-mask process described do not protrude from the spring surface. Extruding tips can be made by adapting this process, one simple way being to prepare a sloping surface on which the tip region of the film is deposited [Fig. 3-10(a)]. Such a sloping surface can be achieved using a rapidly etching sacrificial layer [Fig. 3-10(b)]. Another improvement would be to

(a)



(b)

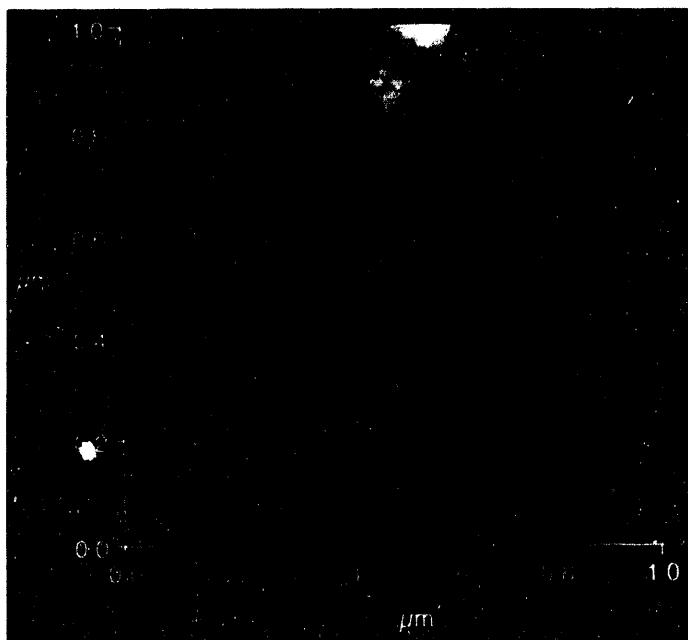


Figure 3-9. The limited opening angle of the tip is noticeable only on tall, steep sample features. (a), a 30 nm step on mica is imaged as a 20° slope, 20° being the inclination of the tip to the sample. Spherical objects may take on a slightly triangular / pyramidal appearance in the images (b). (The sample in both cases is coated ferritin on mica, imaged in hexadecane — see chapter VIII.)

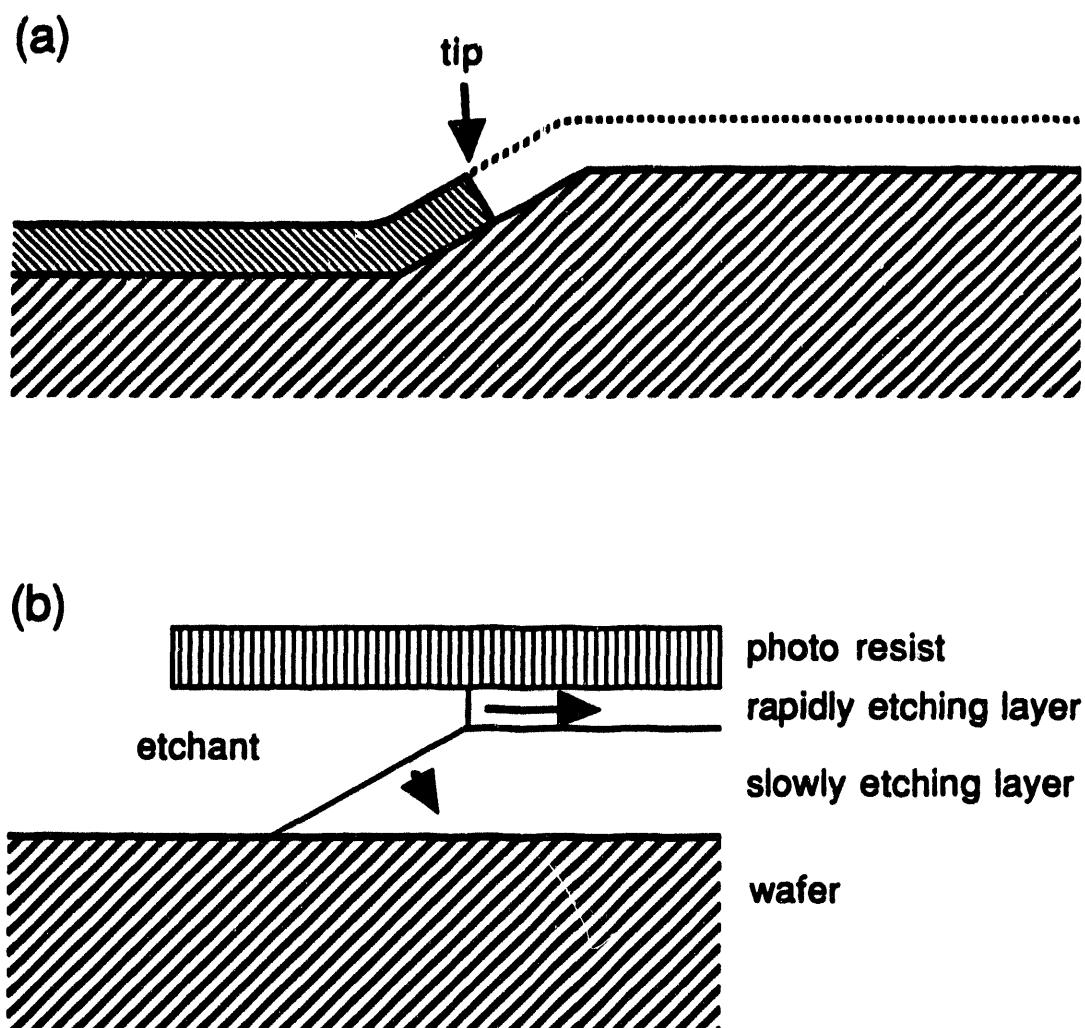
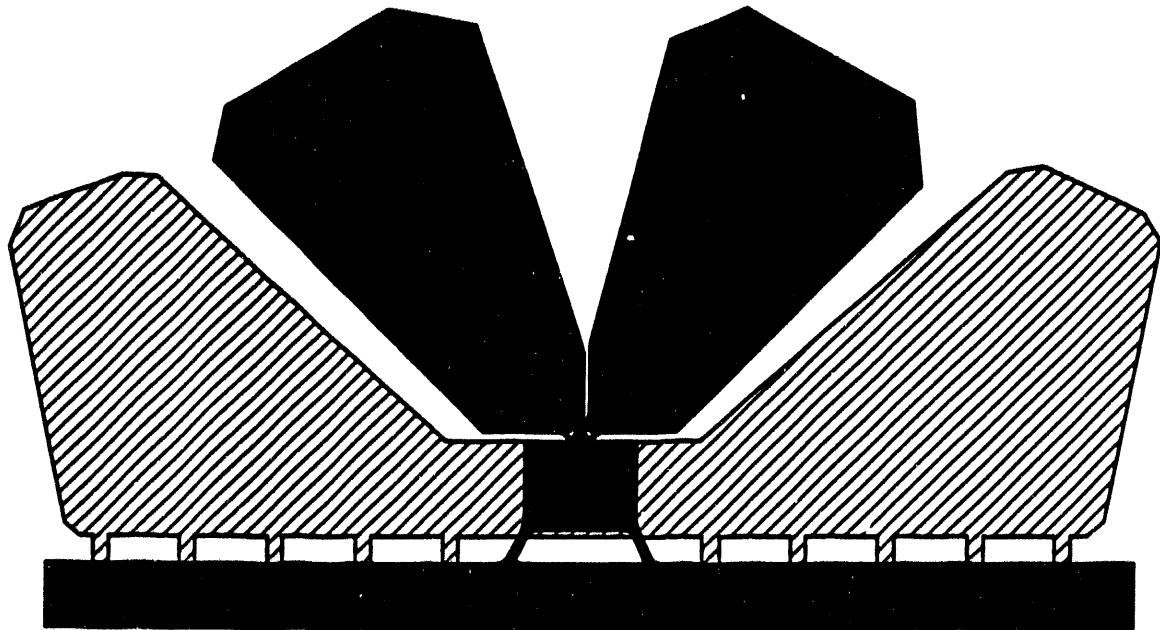


Figure 3-10. The tip-making process can be adapted to create protruding tips. One simple way is to deposit the tip region on an inclined surface (a). Such a sloping surface can be created using a faster-etching layer deposited on top of a thick buffer layer (b).

(a)



(b)

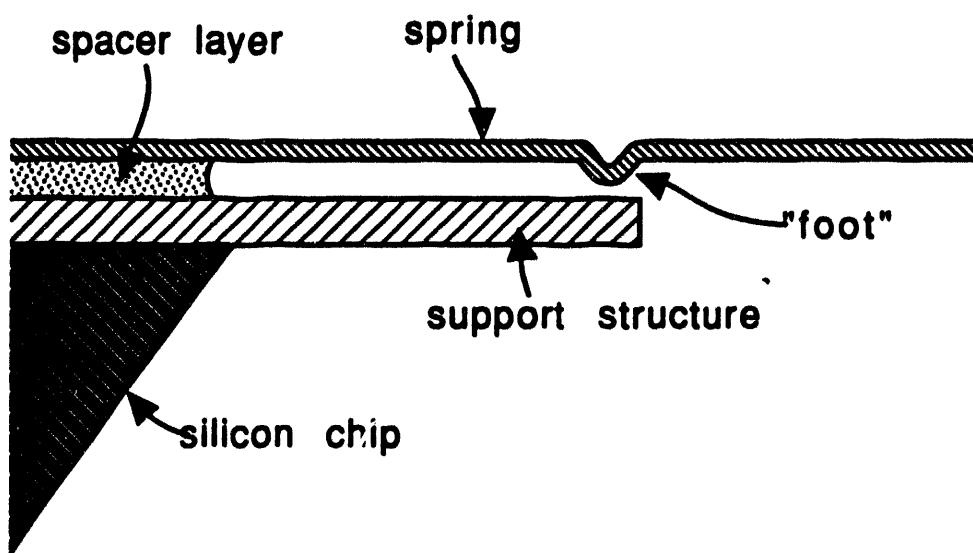


Figure 3-11. The fracturing can be done without micromanipulators if a structure is supplied which can support the spring during fracturing (a). The problem of the spring sticking to the support structure can be avoided by endowing the spring with small "feet" (b).

design the break plates in such a way that the fracturing could be done using simple tools such as tweezers. This would allow each tip to be created by the user immediately before being scanned, minimizing any tip degradation from contamination or diffusive relaxation. The difficulty in implementing this for the current designs is the need to support the spring during fracturing, to avoid damage to the legs or suspension wires. This could be done by adding a second, stronger structure below the spring, which would support it during fracturing, but which could itself be removed afterwards [Fig. 3-11(a)]. To avoid having the two structures stick to each other, as they can be expected to do if allowed to touch¹², the spring could be endowed with small "feet" [Fig. 3-11(b)] which would assure that only point contacts are possible. (Such feet are easily made by etching small depressions in the intermediate layer before the top film is deposited.¹³)

References:

- 1 T. R. Albrecht and C. F. Quate, *J. Appl. Phys.* **62**, 2599 (1987)
- 2 S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, and J. Gurley, *J. Appl. Phys.* **65** 164 (1989).
- 3 T. R. Albrecht, S. Akamine, T. E. Carver, and C. F. Quate, *J. Vac Sci Technol.* **A8**, 3386 (1990).
- 4 R. B. Marcus, T. S. Ravi, T. Gmitter, K. K. Chin *et al*, *Appl. Phys. Lett.* **56**, 236 (1990).

5 K. L. Lee, D. W. Abraham, F. Secord, and L. Landstein, J. Vac. Sci. Techn. B 9, 3562 (1991).

6 D. Keller, D. Deputy, A. Alduino, and K. Luo, Ultramicrosc. B .42-44, 1481.(1992)

7 Hongyu Ximen and P. E. Russel, J. Vac. Sci. Technol. B 9, 2733 (1991).

8 K. E. Petersen, Proc. IEEE 70, 420 (1982)

9 M. Sekimoto, H. Yoshihara, and T. Ohkubo, J. Vac. Sci. Techn. 21, 1017 (1982).

10 A. Reisman, M. Berkenblit, S. A. Chan, F. B. Kaufman, and D. C. Green, J. Electrochem. Soc. 126, 1406 (1979).

11 A. L. Robards and U.B. Sleytr, *Low temperature methods in Biological Electron Microscopy*, Elsevier, Amsterdam, 1985, p. 376.

12 R. L. Alley, G. J. Cuan, R. T. Howe, and K. Komvopoulos, Technical Digest, IEEE Solid-State Sensor and Actuator Workshop, Hilton Head Island, South Carolina, USA, 1992, p. 202.

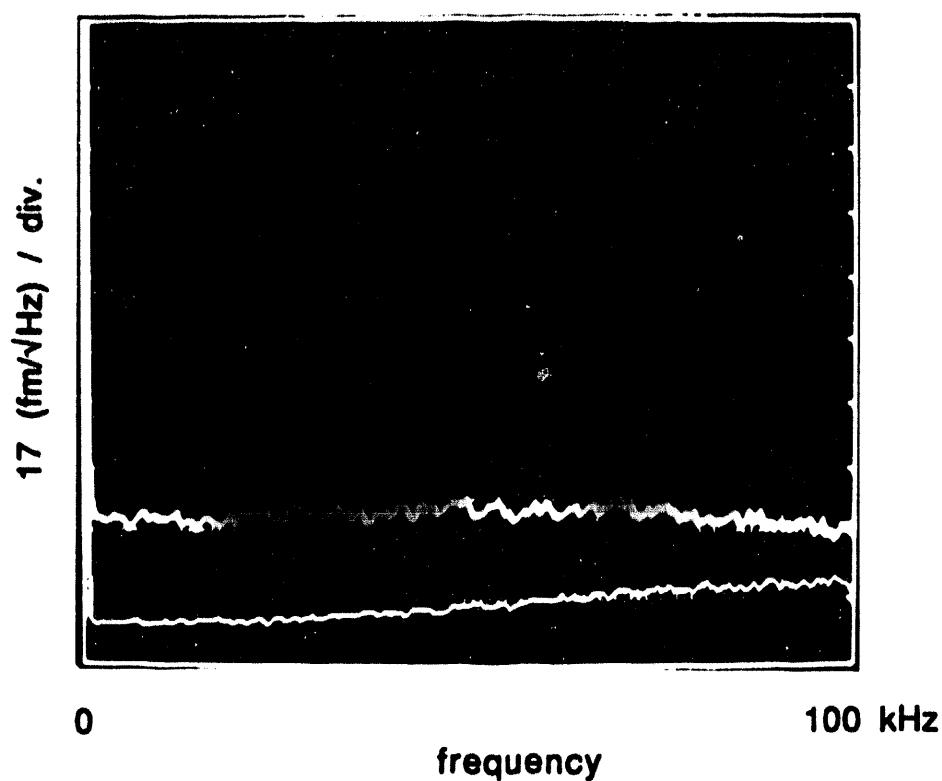
13 W. C. Tang, T.-C. H. Nguyen, and R. T. Howe, Sensors and Actuators 20, 25 (1989).

Chapter IV: Noise Measurement

We characterized the noise of our SFM springs in a test rig, not in the low temperature SFM to be described in the next chapter, to avoid certain complications due to stray reflections at air/quartz interfaces. For this measurement, and in all later data taking, we used a Sharp LTO23MF0 780 nm diode laser, power regulated with the Sharp regulator IR3CO1. Because this is an extreme multi-mode laser, it has considerable noise in wavelength and amplitude due to mode competition. However, this mode competition takes place on a very fast characteristic timescale (sub-ns), and in the low frequency range it manifests itself as a stable and predictable white noise. We found this to be vastly preferable to the drift and sudden dramatic mode jumps characteristic of single mode diode lasers. In a well aligned beam deflection system, neither wavelength nor amplitude noise affects the output signal.

210 μ W of this light was transported through a single mode optical fiber. The beam emerging from the fiber, which is circularly symmetric and diffraction limited, was focused by a miniature graded-index lens¹ to a circularly symmetric waist with an e^2 radius of 25 μ m. At the waist the beam was reflected off a spring of the type shown in Fig. 3-5, and after a distance of 58 mm it was captured on a split photo diode² with a responsivity of 0.45 A/W at this wavelength. Figure 4-1(a) shows the measured rms noise spectrum in the differential diode current. (The strong spring resonance at 201 kHz is outside the frequency range shown. The

(a)



(b)

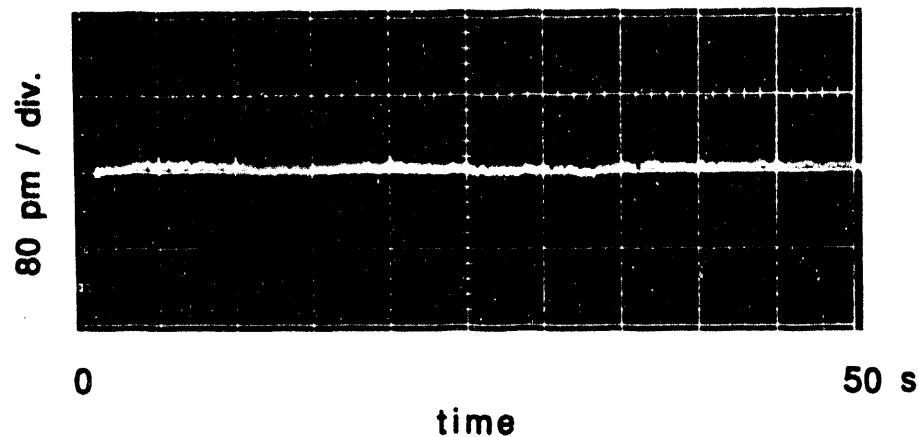


Figure 4-1. (a) RMS noise spectrum from 0 to 100 kHz of the measured deflection of the cantilever spring shown in Fig. 3-5, using 210 μ w of laser power at a wavelength of 780 nm. Both axes are linear. The lower trace was acquired with the laser turned off and shows the contribution of the electronics. (b) Time trace of the same signal over 50 s, in a 0-11 kHz bandwidth.

apparent rolloff above 70 kHz is due to the preamplifier.) The lower trace was measured with the laser off, and represents the noise of the electronics; this could be made negligible by a simple redesign. The measured spectrum is flat from 300 Hz to 170 kHz, and the level of this white noise is $37 \text{ fm}/\sqrt{\text{Hz}}$, or $1370 \text{ (fm)}^2/\text{Hz}$. Of this, we calculate that thermal noise contributes about $250 \text{ (fm)}^2/\text{Hz}$ at low frequencies. (Thermal excitation of an SFM spring manifests itself as a force noise with a white spectral density³ $F_{th} = 4k_B T/\omega_0 Q$. At low frequencies the resulting spring deflection noise is $N_{th} = F_{th}/k = (4k_B T/\omega_0 Q)^{1/2}$. This is non-negligible only because of the comparatively low Q (≈ 43) which is due to air damping. In vacuum, where the Q is likely to be several thousand⁴, this effect is minute.) If we correct for this thermal contribution, and the measured electronic noise of about $100 \text{ (fm)}^2/\text{Hz}$, the remaining white noise level is $1070 \text{ (fm)}^2/\text{Hz}$, or $32 \text{ fm}/\sqrt{\text{Hz}}$.

The low frequency behavior can be judged from the time trace in Fig. 4-1(b), which shows a total peak-to-peak noise of about 20 pm over 50 s in a 0-11 kHz bandwidth. (The calculated contribution from the white noise quoted above is 3.8 pm rms in this bandwidth.) Published interferometric techniques display noise levels of 30-40 pm rms^{5,6} (not p-p) in similar frequency ranges, or drift rates of 0-300 pm/min.⁷ The stability of the beam deflection method is related to the fact that it measures angular quantities, and thus is insensitive to lateral thermal drifts.

For the parameters in this experiment, the shot noise level predicted by (19) is $19 \text{ fm}/\sqrt{\text{Hz}}$. The measured rms white noise is only a factor of 1.7 larger than this value. We attribute the

remaining discrepancy mainly to the wavelength noise of the laser. This does in theory not affect the signal, but could have propagated through if the alignment of the optics was imperfect. Other contributions come from the fact that the beam waist radius was mismatched to the mirror size, and the fact that the mirror is slightly curved because of stress gradients in the film created during deposition, though both of these effects can be estimated and have been found to be small. These cantilevers curve about 0.3 μm over their 50 μm length. Such a curvature, if known, can in principle be compensated for by focusing the beam slightly in front of or behind the mirror, in such a way that the reflected wavefronts are plane as they leave the mirror.

The curvature of the mirror is also affected by the metallization applied to increase the reflectivity (an effect which can be exploited to partly cancel out the intrinsic curvature). However, a metallized cantilever acts like a bimaterial thermometer, making the output signal sensitive to temperature variations, including ones created by variations in the laser power. For applications where this can be a problem, such as in variable temperature SFM's, it is possible to achieve usable reflectivity without any metallization, by adjusting the film thickness so that the front and back surfaces of the transparent silicon nitride spring form a Fabry-Perot cavity tuned for reflection. The maximum theoretical reflectivity of a tuned silicon nitride film of index of refraction 2.0 is 50% in air or vacuum. We have attempted to tune the thickness of springs like these by timed etching in hot phosphoric acid, and have achieved a reflectivity of 42% at optical wavelengths without coatings.

The predicted shot noise for the torsion spring, at the same beam parameters as above, is $8.3 \text{ fm}/\sqrt{\text{Hz}}$. In air, thermal noise dominates over shot noise for these springs because of their low Q value of about 19, so their shot noise properties can only be taken advantage of, and would have to be evaluated, in vacuum.

References:

- 1SEFOC SLW-1.0-25 from NSG America Inc., 28 Worlds Fair Dr., Sommerset, NJ 08873, USA. The lens has a diameter of 1 mm, but the e^2 diameter of the beam as it leaves the lens is only 250 μm , since only a fraction of the large (0.45) numerical aperture is used. This yields the wanted waist radius of 25 μm at a convenient distance of 12.4 mm, and decreases the amount of spherical abberation.
- 2S994-13 from Hamamatsu Photonics, 2444 Moorpark Ave. San Jose, CA 95128, USA. This is actually a quadrant diode, here used with the segments pairwise in parallel.
- 3C. V. Heer, *Statistical Mechanics, Kinetic Theory, and Stochastic Processes*, 431 (Academic, New York, 1972).
- 4T. R. Albrecht, S. Akamine, T. E. Carver, and C. F. Quate, *J. Vac Sci Technol. A8*, 3386 (1990).
- 5D.Rugar, H. J. Mamin, R. Erlandsson, J. E. Stern, and B. D. Terris, *Rev. Sci. Instr.* **59**, 2337 (1988).
- 6C. Schönenberger and S. F. Alvarado, *Rev. Sci. Instr.* **60**, 3131 (1989).

7D. Rugar, H. J. Mamin, and P. Guethner, Rev. Sci. Instr. **55, 2588
(1989).**

Chapter V: Design Concerns for Biological Applications

The first thing one needs to do before starting to design a low temperature SFM, is to estimate just how low a temperature is necessary to achieve the improved sample rigidity. The answer is not obvious. However, there are indications from several techniques, including Mössbauer absorption¹, X-ray crystallography², infrared absorption³, and molecular dynamics⁴, that many proteins undergo a "glass transition" at a characteristic temperature, typically 180-200 K, above which the molecule moves among many different "conformational substates" (different local minima of its conformational potential energy surface), and below which it is frozen into only one such state. It appears reasonable to expect much of any rigidity improvement to have taken place once the temperature is below this range. Therefore, we took it as our design requirement to reach a temperature well below 180 K.

Perhaps the most important concern in designing a low temperature force microscope is to prevent even molecular amounts of contaminants, such as water ice, from condensing on the sample surface or tip. There are two options: one can operate the microscope under ultra-high vacuum conditions (attempts⁵ to operate an SFM at a pressure of 5×10^{-6} Torr (at 140-200K) have encountered severe contamination problems, indicating that higher vacuum than this is needed), or one can operate it *under a liquid*.

Clearly, the ideal environment for studying biological systems is their native aqueous one, but at low temperatures this is not an option. Various techniques have been developed to deal with this

problem in the context of electron microscopy, and many of these, in particular rapid-freezing followed by freeze-fracture, should be well suited to SFM. Indeed, as was argued in the introduction, a rapid-frozen sample can in some circumstances be *more* native-like than one in aqueous buffer, once the latter sample has been adsorbed to a substrate. Since rapid-freezing involves techniques like plunging a thin sample into a suitable liquid or gas jet, it cannot easily be done *in situ*, certainly not in UHV. Thus, to accommodate rapid-frozen samples, the system must allow transferring and loading of samples when both the sample and the microscope are at low temperature. This appears difficult to implement for a UHV system. For this reasons, we chose the option of operating under a liquid.

The simplest way to achieve a stable, low temperature is, of course, to use a bath of boiling cryogenic liquid, such as nitrogen or helium. However, the vibrations associated with the boiling of the cryogen must be isolated from the microscope, while maintaining good thermal contact. One way to solve this non-trivial problem is to work in helium in a "superinsulated"⁶ dewar (which does not need a nitrogen jacket), and pump on the He gas until the liquid goes into the superfluid state, where boiling ceases. However, the problem can be avoided entirely by working at a melting point instead of a boiling point.

The microscope we have built operates immersed in a liquid, typically isopentane, with the temperature stabilized at 143 K by a bath of melting n-pentane. Several factors went into the choice of

isopentane as the immersion liquid. The handling of the cryogen is much simplified if it remains liquid at room temperature, is reasonably non-toxic, and is available in quantity at high purity. The pentanes satisfy these properties, and have some of the lowest melting points of any compounds that do. They are also established as good rapid-freezing media, giving us the option to freeze samples by plunging into the microscope well. A mixture of isopentane with 15% by volume n-propanol may be an even better choice because of its ability to dissolve a small amount of water, which allows it to dissolve fine ice particles settling from the gas, turning somewhat less rigorous the requirement of maintaining a water free atmosphere in the chamber. Pure n-propanol melts at 144 K but is not a good choice, since it turns prohibitively viscous near its melting point.

Being non-polar, these liquids are far from the natural aqueous environment, so the possibility of denaturation must be taken into account. This is likely to be less of a problem for rapid-frozen samples, and it has recently been suggested that imaging biomolecules in a medium in which they are insoluble may actually be advantageous, in that it improves their adhesion to their substrate.⁷

References:

¹F. Parak, E. N. Frolov, R. L. Mössbauer, and V. I. Goldanski, *J. Mol. Bio.* **145**, 825 (1981)

2H Hartmann, F. Parak, W. Steigemann, G.A. Petsko, D. R. Ponzi, and H. Frauenfelder, Proc. Nat. Acad. Sci. USA **79**, 4967 (1982).

3I. E. T. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach, A. H. Xie, and R. D. Young, Phys. Rev. Lett. **62**, 1916 (1989). *Note however the comments in:*

A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter J, and W. A. Eaton, Science **256**, 1796 (1992).

4R. J. Loncharich and B. R. Brooks, J. Mol. Bio. **215**, 439 (1990).

5C. B. Prater, M. R. Wilson, J. Garnæs, J. Massie, V. B. Elings, and P. Hansma, J. Vac. Sci. Techn. B **9**, 989 (1991).

6R. C. Richardson and E. N. Smith, "Experimental Techniques in Condensed Matter at Low Temperatures," Addison Wesley, New York, 1988.

7H. G. Hansma, J. Vesenka, C. Siegerist, G. Kelderman, H. Morrett, R. L Sinsheimer, V. Elings, C. Bustamante, and P. K. Hansma, Science **256**, 1180, (1992).

Chapter VI: The Microscope System

SFM Platform

As described in earlier chapters, our system is based on the now common optical beam deflection method,^{1,2} where a laser beam is reflected off the back side of the SFM cantilever spring, so that any deflection of the spring causes an angular deflection of the reflected beam, which is detected as a differential signal from a two-segment photo diode. The entire SFM, including all the optics except the laser, operates immersed in liquid. However, most of the beam path is confined to the interior of a solid prism, thereby drastically reducing beam disturbances by index-of-refraction gradients or particulates in the beam path.

We have used three kinds of springs/tips: (i) the microfabricated Si_3N_4 cantilever springs of our own design³, with tips created from the film itself by controlled fracture, described in chapter III, (ii) the same springs modified with GaN grown by an electron beam process,⁴ and (iii) commercial silicon cantilevers with etched silicon tips⁵ ("Ultralevers" from Park scientific Instruments).

The SFM, shown in Figs. 6-1 and 6-2, is built on a circular invar base plate, 81 mm in diameter and 17 mm thick. 780 nm light from a laser diode⁶ enters through a single-mode polarization preserving optical fiber, passes through the graded-index focusing lens⁷ and the quartz prism, is reflected off the cantilever back into the prism, is reflected once more by an aluminized face of the prism, and ends up on the two-segment photo diode⁸ which is held in near contact with the prism surface. The photo diode (in its commercial TO-5 package)

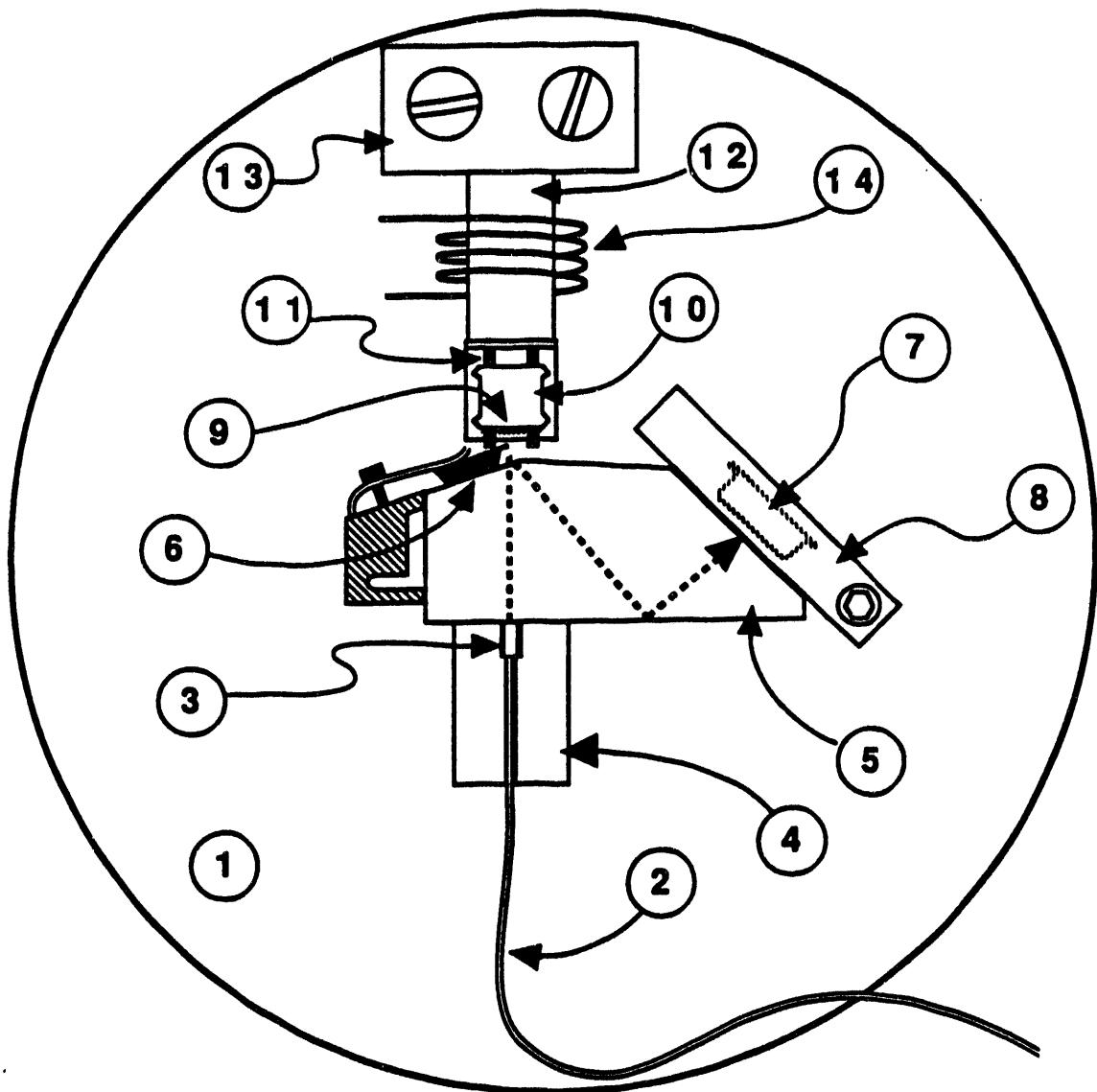


Figure 6-1. The SFM platform viewed from above. The dashed line represents the path of the laser beam. (1) invar base plate (diam. 81 mm), (2) optical fiber, (3) graded-index focusing lens, (4) quartz support sleeve, (5) quartz prism, (6) SFM cantilever, (7) two-segment photo diode, (8) photo diode translation stage, (9) sample, (10) sample carrier, (11) quartz rails, (12) PZT scanner, (13) Macor scanner mount, (14) kicker coil for sample approach.

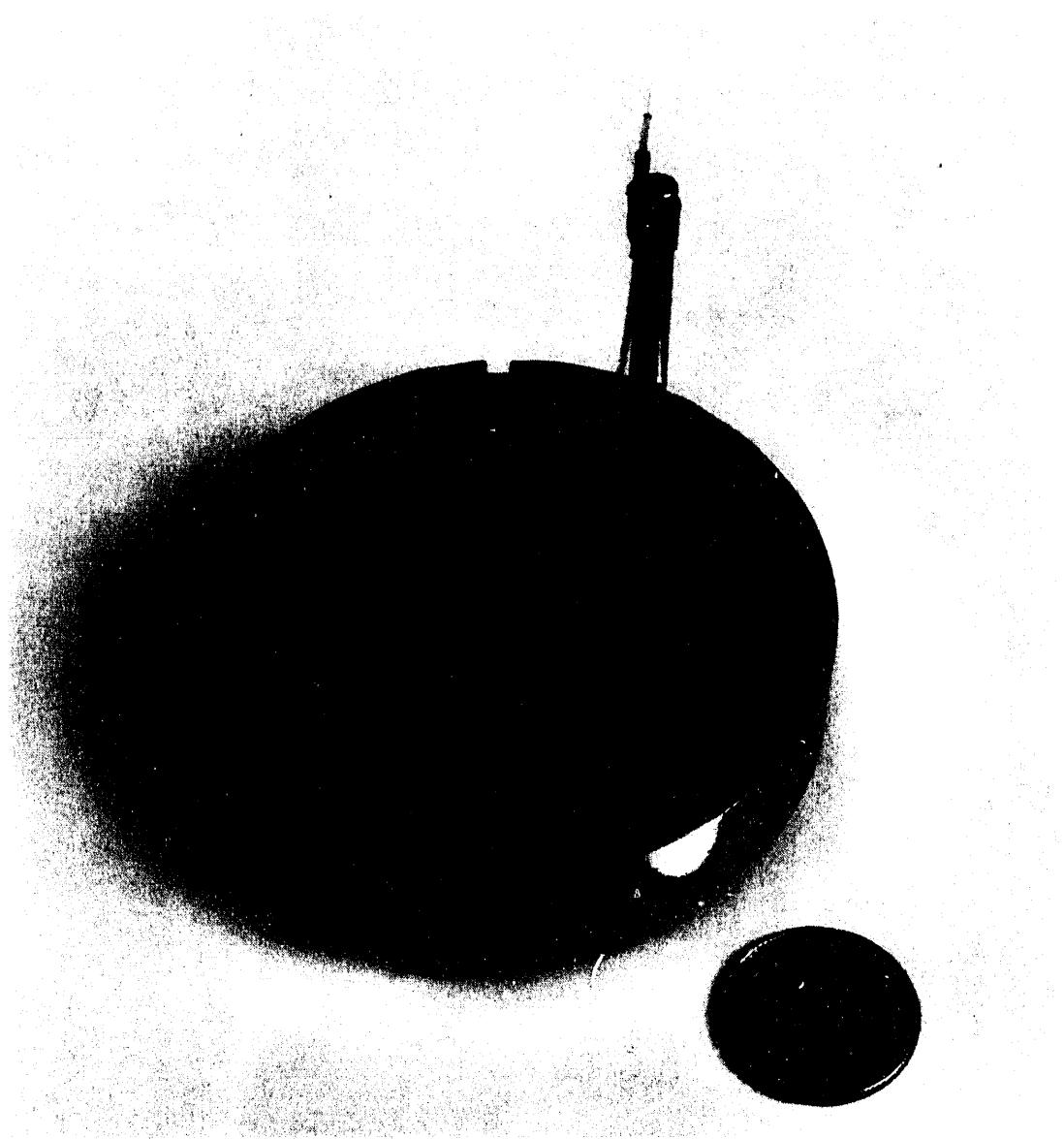


Figure 6-2. The SFM platform. (This photograph was taken before the kicker coil was added.)

resides in a miniature one-dimensional translation stage used to center the diode on the beam, using a guided hex-key tool from above. One dimensional motion is sufficient, since it is of no consequence where in the vertical direction the beam hits the diodes, as long as it doesn't miss entirely. The vertical position of the entire translation stage was adjusted once and for all when the microscope was assembled. This vertical adjustment, as well as coarse horizontal adjustment, is made possible by mounting the stage to the vertical wall of a recess in the base plate, and using adjustment screws to define its position before its mounting screw is tightened down. (The photodiode is specified for operation only in the temperature range -20 - +80°C, but has performed flawlessly at 143 K (-130°C) even after many temperature cycles.) The cantilever chip is held against the prism surface by a BeCu spring clip, whose adjustment screw threads into a small invar block attached to the prism with epoxy^{9,10}. The prism is clamped to the base plate, the top surface of which has been ground flat for stable contact.

The light output from the single mode fiber is circularly symmetric and diffraction limited, making it easy to use optimal focusing (see chapter II). Standard "single mode" fibers actually allow two polarization modes which are degenerate in wavelength. If the fiber is bent or twisted, these modes easily mix due to stress-induced birefringence, creating variations in the polarization state of the fiber output. Such variations can in turn create a apparent signal through any polarization asymmetry, such as of stray reflections at interfaces. This is the reason we used a polarization

preserving fiber — it admits only one mode, hence is totally insensitive to such disturbances.

Wherever a laser beam exits or enters a solid surface at a non-perpendicular angle, it will be refracted by an amount which will depend on the index of refraction of the surrounding medium, which in turn depends on the temperature. This could create an unnecessary temperature dependent offset of the spring deflection signal, unless such contributions are made to cancel. In the present design, the beam exits and reenters the prism at equal angles near the cantilever (when the cantilever is at its rest position), giving zero total refraction angle independent of the index of refraction (Fig. 6-3). (While an absolute *translation* of the beam does occur, this effect is small due to the short lever arms involved, though it has to be taken into account when aligning the cantilever to the beam.) At the photo diode the beam exits the prism nearly perpendicularly (though not exactly so, to avoid backreflection).

The same effect prevents the use of an immersed (conventional) focusing lens, since its effective focal length would become strongly dependent on temperature (and on the choice of immersion liquid). This is why we use a graded index¹¹ (GRIN) focusing lens, which does not rely on interfaces for focusing. In fact, the fiber end, the lens and the prism (together with a quartz support sleeve) are potted together using optical epoxy, maintaining alignment permanently. (The exposed optical epoxy is then covered with Stycast⁹ epoxy, which is more nearly inert to the immersion liquids.)

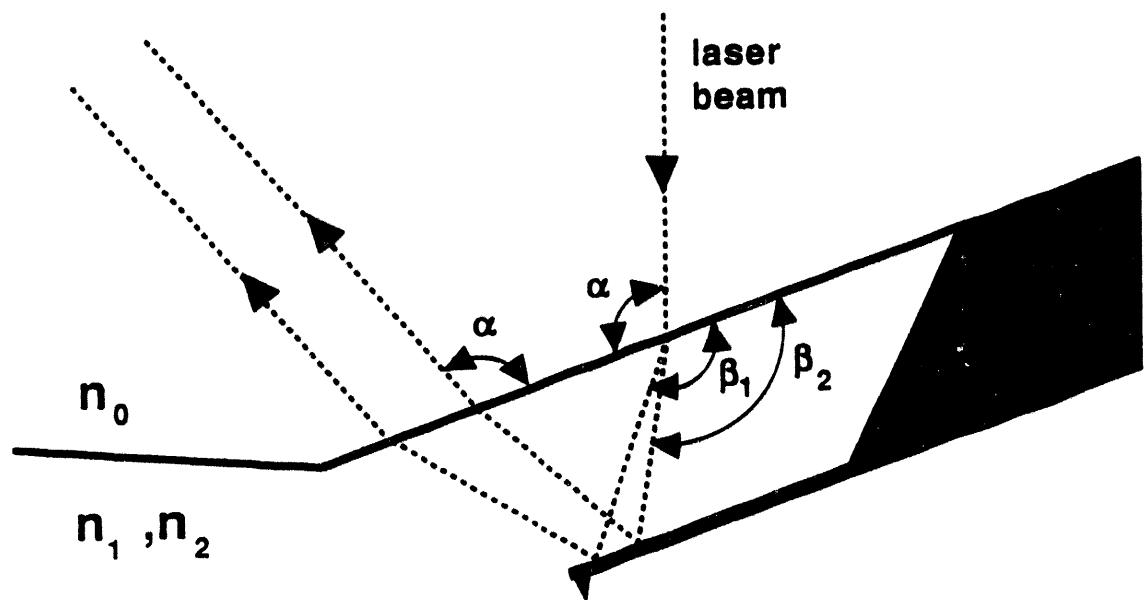
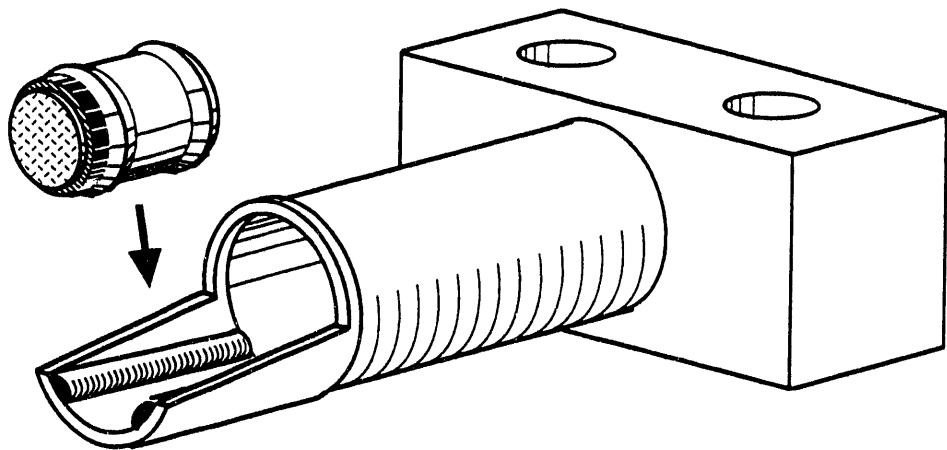


Figure 6-3. Closeup of the path of the laser beam near the cantilever. When exiting the quartz prism, the beam is refracted to some angle β which depends on the index of refraction of the surrounding medium, but the final beam angle α after it has reentered the prism is unaffected because of symmetry. The translation of the beam is not canceled by symmetry, but is small because of the short lever arms involved (the thickness of the cantilever chip is $500\mu\text{m}$). However, the translation of the beam focus relative to the cantilever has to be taken into account when the cantilever is aligned in air for operation in a liquid.

The sample is mounted on a magnetic stainless steel disc, held magnetically onto a thread-spool-shaped carrier containing a rare earth cobalt permanent magnet¹². This carrier rests on two quartz rails on an Al cradle at the end of the piezoelectric tube scanner (Fig. 6-4). The thread-spool shape of the carrier guarantees stable four-point contact with the rails. The carrier is electrically grounded through a thin Cr film applied to the rails. The sample approach is done either through an "inertial walking" motion¹³ of this same piezoelectric, or by applying magnetic field pulses with a kicker coil^{14,15} mounted concentrically with the scanner. The latter method has proven more reliable in our system. The piezoelectric material of the scanner is PZT 5A, and it is 0.5 inch long, 0.25 inch in outer diameter and has a wall thickness of 0.03 inches, yielding an XY range of 3.6 μm at room temperature and 2.0 μm at 143 K. It is epoxied⁹ to a ceramic¹⁶ mount, which is screwed to the base plate. Invar was chosen as the base plate material to minimize the differential thermal expansion relative to the PZT, though stainless steel or Macor may be better choices since they are nonmagnetic and thus would not interact with the magnetic sample carriers, simplifying sample exchange.

There is no provision for alignment of the beam onto the cantilever *in situ*. The platform is sufficiently stable that the cantilever can be aligned once and for all, by hand, in air, before cooling down (with appropriate offset, typically around 90 μm , to compensate for the shift in index of refraction when liquid is added, see Fig. 6-3).

(a)



(b)

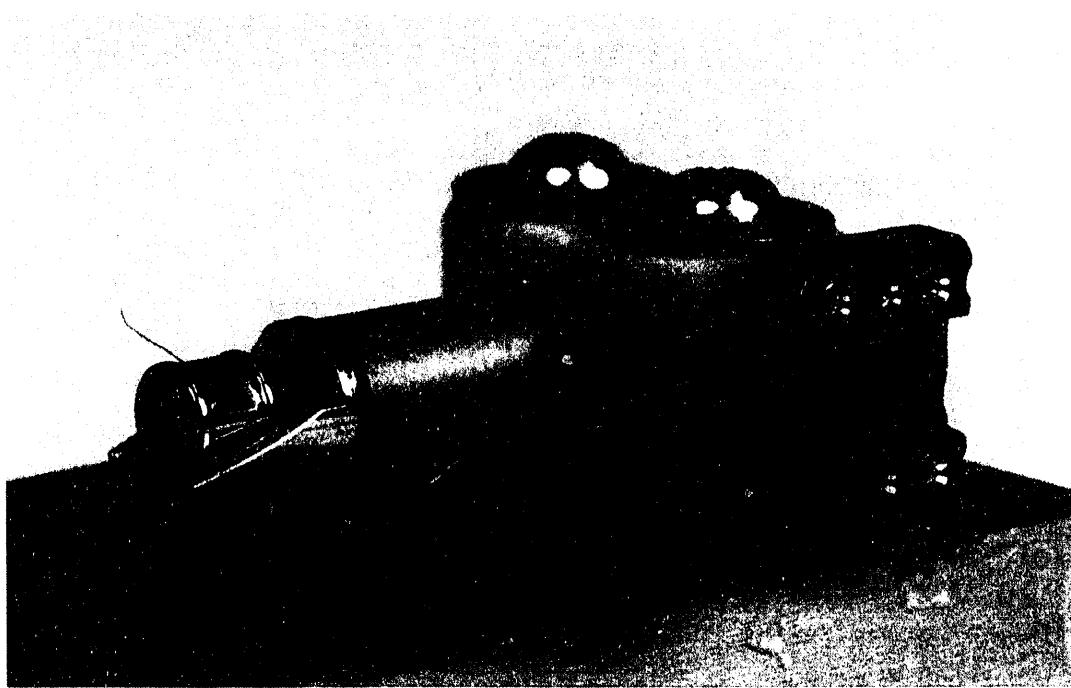


Figure 6-4. The piezoelectric scanner assembly. (The thin grounding wire seen touching the sample carrier in (b) is an obsolete feature; the grounding is now done through a thin metal film evaporated onto the rails.)

This simple design has several features: samples can easily be exchanged by dropping a new carrier onto the rails from above; the only mechanical connection to the outside world is the adjustment screw for the photo diode translator, and this adjustment is done only once per cooldown, after which the tool is removed; the mechanical/thermal loop from tip to sample is very short since both the scanner mount and the prism are clamped to the baseplate, and the cantilever chip is clamped to the prism. A drawback is the low resonance frequency of the scanner, about 4 kHz, caused by the comparatively large weight of the sample carriers.

Cooler and Dry Chamber

Figure 6-5 shows the whole system. The SFM platform rests on the bottom of a well filled with a liquid, typically isopentane (2-methyl butane) or (in our early work) n-pentane. The well is surrounded by a larger, closed vessel filled with n-pentane, some of which can be frozen by flowing liquid nitrogen through a cooling coil, and which in turn rests in a glass dewar. After the flow of nitrogen is stopped, the n-pentane slowly melts, and the whole cooler equilibrates at the melting point of n-pentane, 143 K. The iso-pentane stays liquid since its melting point is lower (113 K). While the rest of the n-pentane vessel is made from stainless steel, the microscope well wall is copper, and forms part of a copper heat distributor, with fins extending throughout the n-pentane space. To decrease the heat load, which is dominated by radiation from above, we cover the top of the glass dewar with a teflon lid, the bottom of which is covered with aluminized mylar. In practice, the rate of temperature drift at the microscope stage stays below 10 mK/min

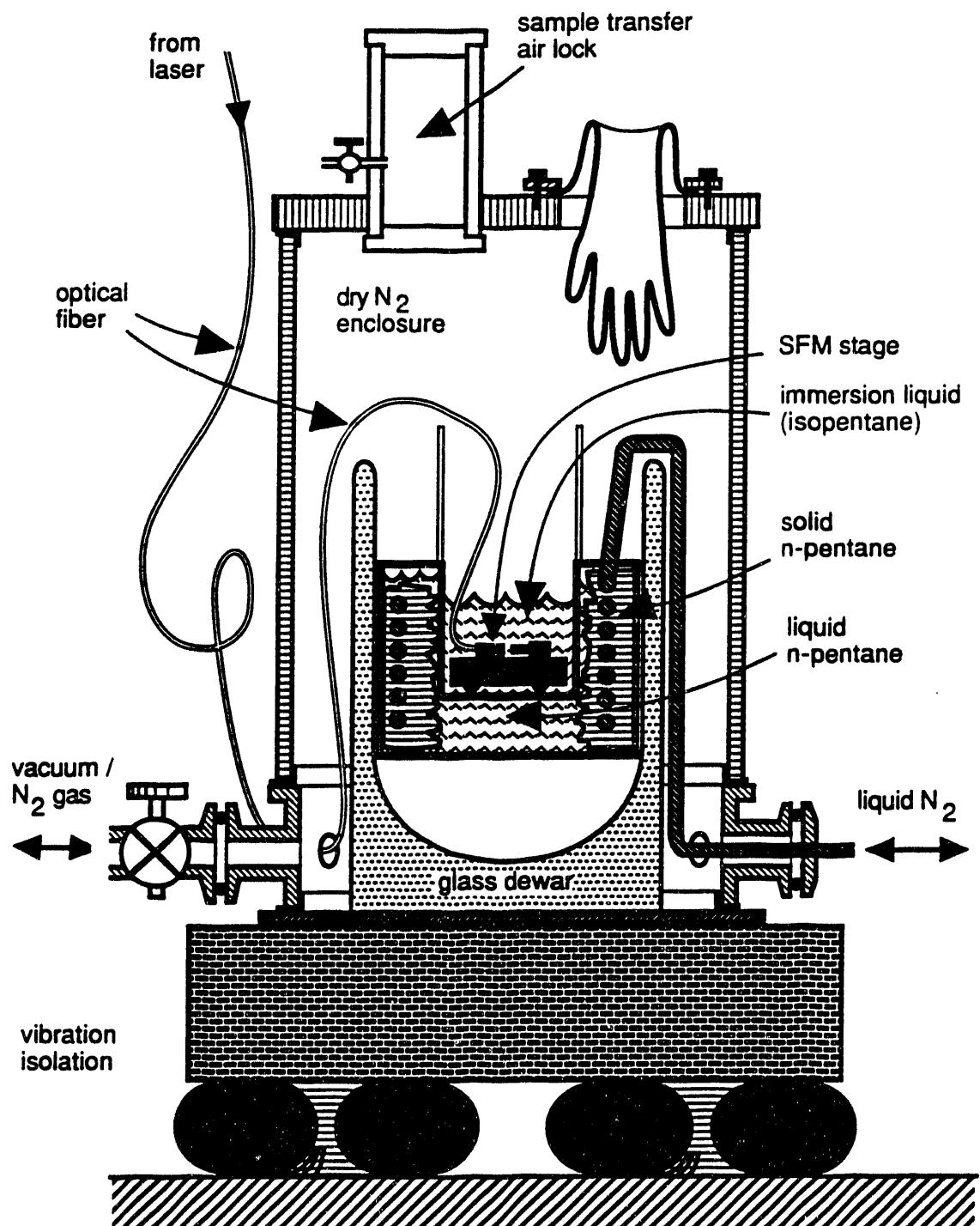


Figure 6-5. Cooling system, dry chamber and vibration isolation.

for about 4 hours. By re-freezing the n-pentane at this time, the system can be kept cold indefinitely, allowing many samples to be examined per cooldown.

In spite of the prism, the laser beam still travels through the liquid for a short distance (twice the chip thickness =1 mm) near the cantilever, and is therefore not entirely immune to particulates in the liquid. A potential source of such particulates is water vapor (or other condensable contaminants) in the gas above the isopentane, which could condense into "snow" and settle into the liquid. Such contaminants therefore have to be avoided. For this reason, the SFM dewar has been enclosed in a vacuum-tight chamber, which can be evacuated to a few mTorr, and backfilled with dried nitrogen gas, before the isopentane is introduced. The chamber consists of an Al bottom plate, a stainless collar with kwik-flange feedthroughs (for electronics, vacuum, N₂ gas, liquid nitrogen and the optical fiber), a Pyrex cylinder, and an acrylic top plate, all joined with Buna-n L-gaskets. The top plate contains ports for a pair of butyl rubber dry box gloves¹⁷, which are used for sample loading and diode adjustment, and a sample transfer airlock, which can be separately evacuated and backfilled. A cover for the glove ports, with connection to the pumping line, allows the outside of the gloves to be evacuated at the same time as the chamber, so that the gloves can be outgassed while remaining relaxed. The L-gasket joints are clamped so as to seal even at the slight positive pressures in the chamber which may occur when the gloves are being used. To remove any water or other condensable contaminants in the isopentane, it is

stored with 4A molecular sieves¹⁸, and is re-distilled into an evacuated metal bottle before use. Vibration isolation is provided by epoxying the bottom plate of the chamber to a concrete block which rests on three rubber inner tubes, and by physically disconnecting the vacuum pump and liquid nitrogen lines during scanning. This gives a main resonance frequency of about 3 Hz, with a Q around 2. The low Q is essential since any sustained low frequency oscillations, such as pendulum motions of a suspended system, would cause movement of the liquid. So as not to compromise the acoustic isolation provided by the chamber wall, no rigid connections have been made from the feedthroughs to the microscope cooler except through the concrete block; the liquid nitrogen lines are decoupled as far as possible using stainless hose.

Electronics

STM ELECTRONICS

Figure 6-6 shows an overview of the electronics, computing and data acquisition components of the SFM system.

Many of the elements used (specifically, the feedback, high voltage, and signal conditioning amplifiers, the XYZ matrix mixer, and the magnetic walker controller) are electrically identical to the electronics developed for STM purposes by earlier members of the Clarke group.^{19,20} In the following paragraphs, I will briefly describe the components which are different: the preamplifier, the ramp generator, the computer system and data acquisition hardware; the interrupt circuit, and the laser regulator.

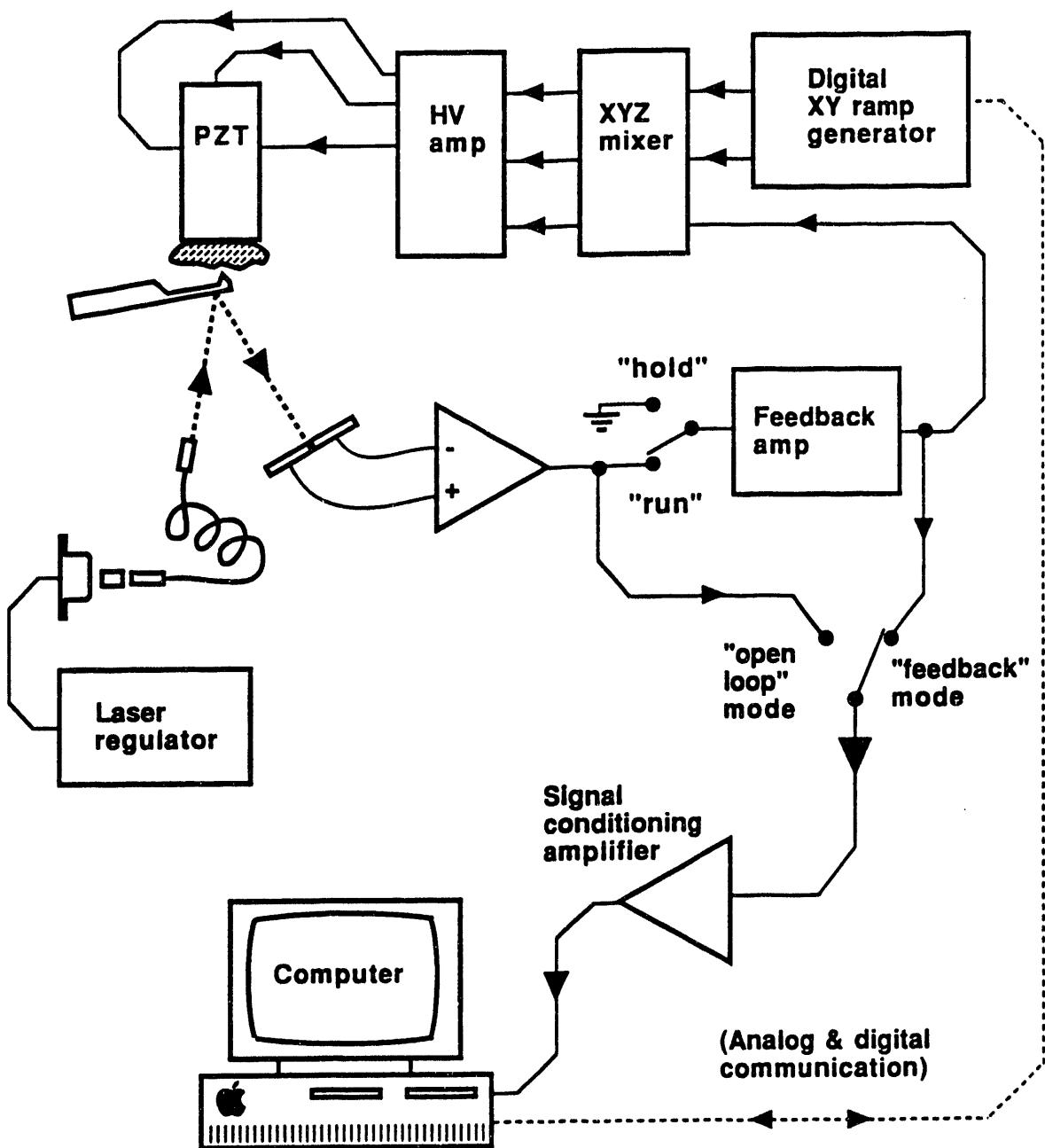


Figure 6-6. Overview of the electronics and data acquisition system.

LASER REGULATOR

For reasons that were discussed in the noise measurement section, we use a Sharp LTO23MF0 780 nm laser diode, which is an extreme multi-mode laser. The TO-5 diode laser package contains an internal photo diode which reads the light coming out the back surface of the laser chip. This is used to feed back to the laser drive current, so as to maintain a constant output power from the laser, using a dedicated Sharp integrated circuit²¹. (The laser and regulator together cost less than \$30.)

PHOTO DIODES AND PREAMPLIFIER

The photo diode we used, S994-13 from Hamamatsu, is actually a quadrant diode with the elements used pairwise in parallel. (If one were to utilize all four quadrants, lateral friction forces could be monitored simultaneously.) A schematic of the preamplifier is shown in Fig. 6-7. Each half is connected to a low noise FET operational amplifier²² in the "photoamperic" configuration, i. e. no bias voltage is applied across the diodes, and the op-amp converts the zero-bias photo current to a voltage. At zero bias voltage there is zero dark current, and therefore less noise. The difference between the two voltages is calculated by an instrumentation amplifier,²³ and the sum by a simple summing operational amplifier. If required, the ratio $A-B/(A+B)$ of these output signals could be calculated (using an analog divider,²⁴ or numerically if both $A-B$ and $A+B$ are acquired by the computer) to suppress any variations in the laser light intensity. We have not found this necessary. The noise of this preamplifier is dominated by the thermal (Johnson) noise in the

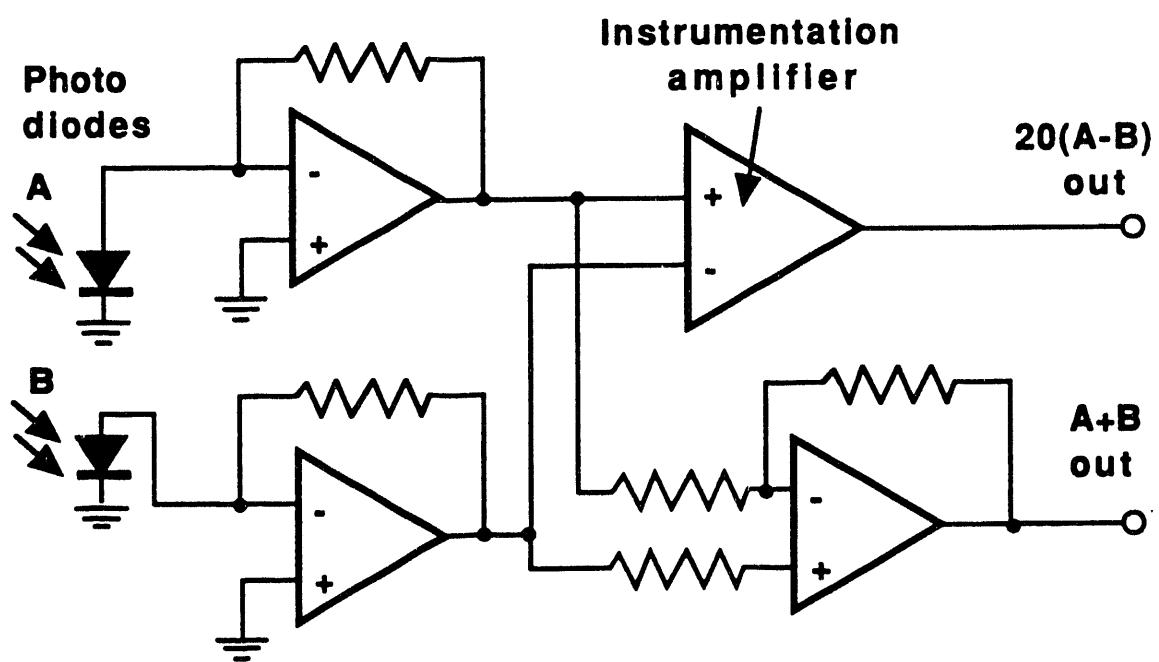


Figure 6-7. Schematic of the preamplifier.

op-amp feedback resistors, and is small compared to the photon shot noise.

DIGITAL RAMP GENERATOR

To simplify operation of the SFM, and to achieve better control and more consistency of the ramps, I designed a digital replacement for the analog ramp generator described in reference (20). The concept is given by Fig. 6-8: pulses from a variable rate clock are fed to an X counter. When the counter reaches a set endpoint, its direction is reversed, and an appropriate number of pulses are allowed to reach a Y counter. The direction of the Y counter is similarly reversed when it reaches a certain count. The outputs of these counters are digital-to-analog converted to make two analog ramps: a fast X ramp and a stepwise Y ramp of appropriately slower rate. These ramps are of fixed amplitude, and are scaled to output ramps of the desired size by two (buffered) potentiometers, ganged in tandem on the same front panel knob. Pulses to trigger the data acquisition board in the computer are generated simply by selecting an appropriate output pin of the X counter, and using this as the trigger. Thus there are only three major front panel controls: a speed, a ramp size, and a scan number selector, which controls how far the Y counter is advanced per X ramp, and how many trigger pulses per ramp are sent to the acquisition board.

As an additional feature, the analog ramps are sent, before they are scaled, through an analog XY rotator,²⁵ with which the scan orientation can be rotated in the XY plane by an arbitrary angle. This allows us to resolve situations where an observed feature is suspected to be an artifact affected by the scan direction.

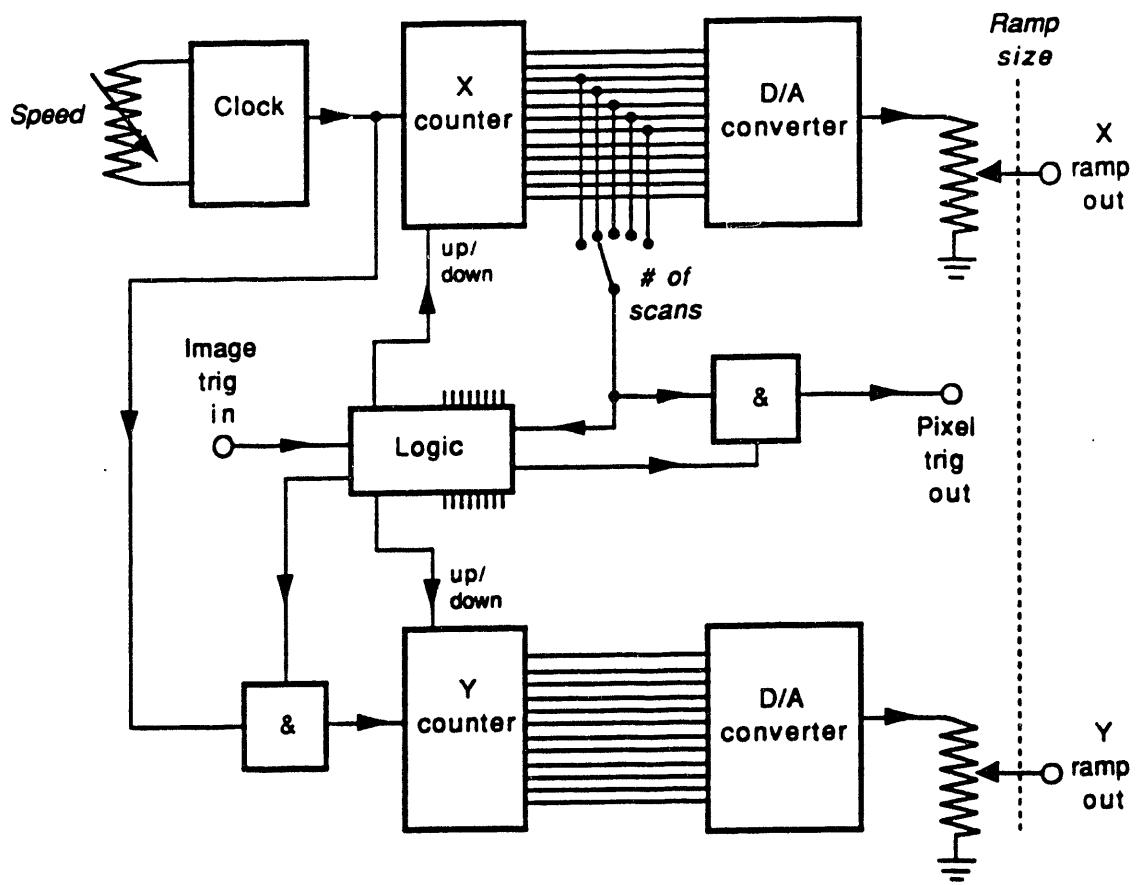


Figure 6-8. Schematic of the digital ramp generator.

The X ramp is active at all times (though its magnitude may be turned down to zero), while the Y ramp is triggered by the computer. (The purpose of this arrangement is to avoid unnecessary distortions at the beginning of the scan, caused by the hysteretic behavior of the scanner — see chapter VII.)

I want to stress the great importance of careful grounding, shielding and layout when digital components are combined with noise sensitive analog circuits.

INTERRUPT CIRCUIT

One item of electronics is specific to this system: If a large suspended particle strays into the beam, it can create a stray signal which makes the feedback circuit push the sample hard into the tip, risking tip damage. To avoid this risk, a comparator circuit looks for the accompanying drop in the sum signal from the two photo diode segments (the SFM deflection signal being their difference) due to scattering off the particle, and when triggered interrupts the feedback loop and pulls the tip back a settable distance to safety. If the system is well prepared, very few such particles are present, and this circuit is only a safety precaution.

MAC DATA ACQUISITION HARDWARE

Data are acquired into a Apple Macintosh II computer through a data acquisition board,²⁶ which contains eight 12 bit differential input A/D converter channels, two D/A converters, five timers and eight digital I/O channels. The A/D converters are used to sense the main image signal (the output of the signal conditioning amplifier), and also the ramp size and rotation angle signals from the ramp

generator, the raw preamplifier signal (which is used for to detect tip-sample contact during walkup), and, when applicable, a second simultaneous channel of image data. Several digital channels are used to communicate with the ramp generator: detecting the status of the ramps (for timing purposes), triggering image aquisition, and resetting the ramp generator if the user aborts an image. Another digital channel triggers pulses from the magnetic walker controller. One D/A converter is used to generate test ramps during walkup, and to perform cross-calibration and elasticity measurements (see chapters VII and XI, respectively). The period of the X ramp is measured with the help of a timer.

To allow imaging at high rates, a second plug-in board transfers the acquired data directly into RAM memory by DMA, relieving the processor of this task. The maximum data acquisition rate is limited by the settling time of the A/D converter, and is about 95 000 samples per second, or one 256×256 pixel image in 1.7 seconds.

SOFTWARE

To control the SFM, and acquire, analyze and display the images, I have developed²⁷ a software application for the Macintosh, which now exceeds 20 000 lines of working C code. (For scale, this thesis contains about 3000 lines of text.)

The application contains routines to control the magnetic walkup (generating magnetic field pulses alternating with test ramps to check if the tip is within reach), to communicate with the ramp generator and acquire image data, to apply various filters to the images, to make quantitative measurements, and to display the images in topview or 3D formats. Many data files can be open

simultaneously, each associated with one or more windows containing different views, line scans, etc.

The filtering routines include facilities to:

- Flatten the image by fitting (in the least squares sense) a plane to the data, and subtracting this plane. The fit can be done considering either the entire image or only a specified "region" (a union of rectangles).
- Remove slow drift by adding appropriate constants to each horizontal line, attempting to flatten a specified region.
- Apply average, median or high pass convolution filters.
- Remove horizontal glitches by replacing those pixels with an interpolation between adjacent lines. This is useful for 3D display, where a tall glitch can visually dominate an image, distracting from the points of interest.

The data analysis routines include facilities to:

- Draw height profiles along arbitrary lines in an image.
- Measure the height of features (in three different senses)
- Measure the height difference between different regions of the image (by fitting a pair of parallel planes).
- Measure the contour length of curved objects.
- Calculate the 2-dimensional Fourier transform of the image.
- Characterize the nonlinear behavior of the scanner (using reference samples, see chapter VII) in order to allow accurate measurements in the above routines.

Examples of the user interface will be shown in chapter VII (in the context of calibration) and chapter IX (in the context of length measurements on collagen molecules).

Experimental Procedure

With the laser power turned down to $4\mu\text{W}$, we remove the scanner assembly, which makes it possible to observe the beam through an optical microscope. The laser wavelength, 780 nm, is nominally in the near infrared, but is seen by most people, and at this power the beam can safely be watched directly. The cantilever is positioned by hand relative to the beam focus (with appropriate offset, as mentioned above), using tweezers under the microscope, and its spring clip is tightened down. Positioning to 5 μm precision is easily achieved by hand, and is quite sufficient.

The metal pentane bottle is evacuated, then cooled with liquid nitrogen. A valve is opened to the vapor above the isopentane supply, allowing isopentane to distill over into the bottle. The bottle is warmed to room temperature, pressurized with nitrogen gas, and placed in the chamber, together with any dry samples to be studied. The chamber is evacuated to a few millitorr (typically overnight), then backfilled with nitrogen gas which has been passed through a liquid nitrogen cold trap to remove water vapor and other condensable contaminants. Isopentane from the bottle is transferred into the microscope well, and liquid nitrogen is passed through the cooling coil until most of the n-pentane is frozen. After the temperature stabilizes, the photo diode is adjusted, and samples can be plunged into the isopentane, placed on the scanner rails, walked up, and scanned. Typical scan forces are 0.1-0.5 nN. A constant-force

scan takes typically 5-10 minutes, while images in the variable deflection mode are acquired in about 20 seconds.

After all samples have been imaged, the isopentane is sucked out, and the gas in the chamber is evacuated through a cold trap, trapping the isopentane vapor. Thus no isopentane is ever released into the room (nor is any n-pentane, since the cooler vessel is hermetically closed).

References:

- 1**G. Meyer and N. M. Amer, *Appl. Phys. Lett.* **53**, 1045 (and 2400, erratum) (1988).
- 2**S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, and J. Gurley, *J. Appl. Phys.* **65** 164 (1989).
- 3**M. G. L. Gustafsson and J. Clarke, to be submitted to *Rev. Sci. Instr.*
- 4**These tips were grown by P. E. Russel, whose help is gratefully acknowledged.
- 5**Ultralevers from Park Scientific Instruments, 1171 Borregas Ave., Sunnyvale, CA 94089, USA.
- 6**LTO23MF from Sharp Electronics Corporation, Sharp Plaza, Mahwah, NJ 07430-2135, USA.
- 7**W1.0-.25-NC-.83 from NSG America, Inc., 28 Worlds Fair Drive, Somerset, NJ 08873, USA.
- 8**S994-13 from Hamamatsu Corporation, 420 South Ave., CN 420, Middlesex, N.J. 08846-258M, USA. (This is actually a quadrant diode with the elements used pairwise in parallel. If one were to utilize

all four quadrants, lateral friction forces could be monitored simultaneously.)

⁹Stycast 2850FT from Emerson & Cuming Inc., Woburn, MA 01888

¹⁰The invar block has three "legs" which are epoxied to the quartz prism. The slight flexibility that this introduces prevents the differential thermal contraction from fracturing the quartz.

¹¹C. Gomez-Reino, Internat. J. Optoelectr. **7**, 607 (1992).

¹²90C1798B (1/8" diameter by 1/8" long cylindrical) from Dexter Permag, 1159 Sonora Court, Sunnyvale, CA 94086, USA.

¹³J. W. Lyding, S. Skala, J. S. Hubacek, R. Brockenbrough, and G. Gammie, Rev. Sci. Instr. **59**, 1897 (1988).

¹⁴D. P. E. Smith, S. A. Elrod, Rev Sci Inst **56**, 1970-1971 (1985)

¹⁵E. D. Ganz, Ph. D. thesis, Physics Department, University of California at Berkeley, 1988.

¹⁶Macor machinable ceramic from Leed Plastics Inc., 793 East Pico Blvd., Los Angeles CA 90021.

¹⁷5B3032 or 5BSF3032 from North Hand protection, 4090 Azalea Drive, Charleston, SC 29405.

¹⁸J. T. Baker, Phillipsburg, NJ 08865, USA. For solvent drying procedures, see for example J. A. Riddick, W. A. Bunger, and T. K. Sakano, *Organic Solvents - Physical Properties and Methods of Purification* (John Wiley & sons, New York, 1986), p. 795 ff.

¹⁹E. D. Ganz, Ph. D. thesis, Physics Department, University of California at Berkeley, 1988.

20R. E. Thomson, Ph. D. thesis, Physics Department, University of California at Berkeley, 1991.

21IR3CO1 from Sharp Electronics Corporation, Sharp Plaza, Mahwah, NJ 07430-2135, USA.

22OPA111B from Burr Brown Corporation, P. O. Box 11400, Tucson, AZ 85734, USA.

23INA101 from Burr Brown Corporation, P. O. Box 11400, Tucson, AZ 85734, USA.

24E. g. DIV100 from Burr Brown Corporation, P. O. Box 11400, Tucson, AZ 85734, USA.

256125A analog image rotator from Optical Electronics, Inc., P. O. Box 11140, Tucson, AZ 85734, USA.

26MIO-16 from National Instruments Corporation, 12109 Technology Boulevard, Austin, TX 78727-6204, USA.

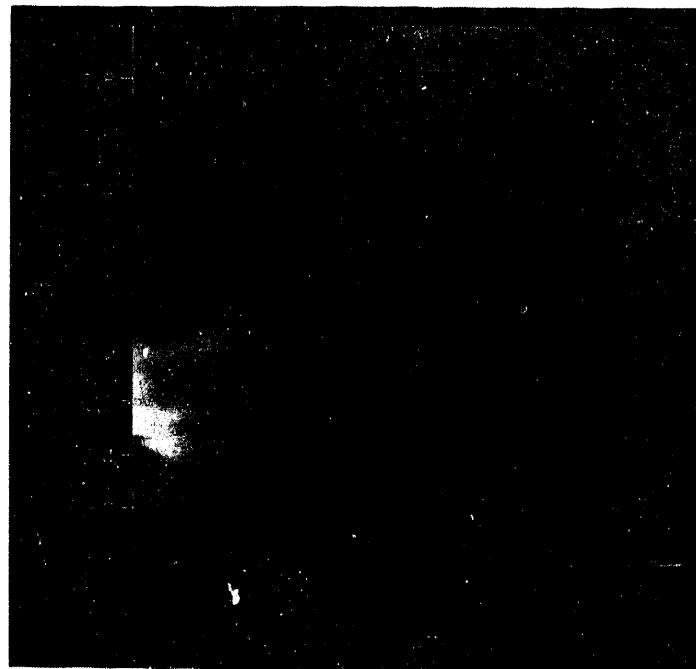
27About half of the programming was done by Marc Hertlein, whose help is gratefully acknowledged.

Chapter VII: Calibration

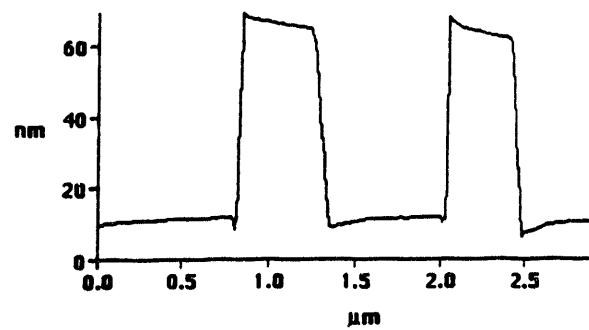
Hysteresis of PZT

Modern scanned probe microscopes almost all use a lead zirconate titanate (PZT) ceramic as the piezoelectric material of the scanner. These materials have large piezoelectric coefficients, but are plagued with nonlinearities, hysteresis and drift. If a voltage ramp is applied to a PZT scanner, its response is nonlinear, depends on the amplitude and rate of the ramp, and continues in a logarithmic fashion for many minutes after the signal is removed. Figure 7-1 illustrates the problem: In Fig. 7-1(a), an image of a square-ruled grating sample, what should be parallel lines and right angles are distorted, and the apparent size of the squares varies over the image. Figure 7-1(b) is a line scan of a step grating (both of these gratings are described later in the context of calibration) acquired in the feedback mode, and demonstrates how the same problem manifests itself in the Z direction. The tops and bottoms of the grating lines, which are flat on the sample, are distorted in the image: After the tip has traversed an upwards step, the tendency of the PZT is to keep creeping upwards. The feedback circuit has to compensate for this creep, causing the feedback voltage, and thus the apparent Z height, to turn back down. For comparison, Fig. 7-1(c) shows the same line scan taken in the open loop mode, where this effect does not occur since the voltage on the Z piezo is held fixed. (One might think here that the open loop mode is the answer to the problem. This is not the case; the open loop mode has its own limitations. In the open loop mode, the tip force increases

(a)



(b) Feedback mode



(c) Hold mode

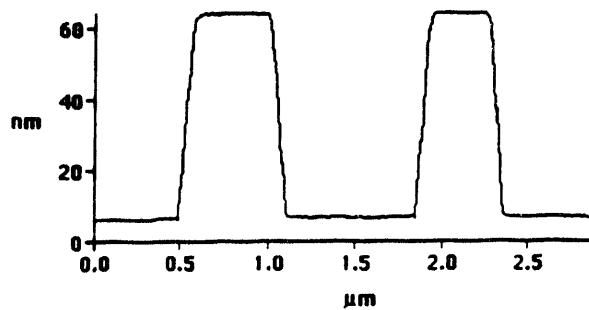


Figure 7-1. Illustration of the effects of the hysteretic behavior of the piezoelectric scanner. (a) A regular, cross-ruled grating is distorted so that lines appear curved and the squares appear to be of unequal size. (b) The same problem in the z direction distorts the profile of a step grating imaged in the feedback mode, while (c) the same profile taken in the open loop mode is undistorted.

dramatically when the tip scans over tall features, often to levels that can be destructive to the sample or tip.)

This property of PZT makes it very difficult to use an SPM as a metrology tool; any attempt to make a precise measurement must be accompanied by extensive characterization of the scanner. In fact, no amount of characterization can completely eliminate the problem unless the entire motion history, including all position changes, is recorded when the SFM is used, and the scanner response is simulated using a computer model. This is feasible in principle, but impractical.

This section describes our attempts to characterize the scanner of our low temperature SFM. Since we need the calibration at 143 K as well as at room temperature, we cannot use benchtop measurement tools such as interferometers or inductive motion gauges. Instead, we have used grating samples of known dimensions, and analyzed the actual SFM images.

XY Calibration

For calibration in the XY plane, we used commercially available cross-ruled grating replicas¹ intended for TEM calibration. These are made by depositing a thin film of SiO₂ on a cross-ruled grating, then freeing the film and supporting it on a copper TEM grid. The thin film is too flimsy to be scanned directly in the SFM. To transfer the replica to a more suitable substrate, we adapted an established procedure for preparing shadowed samples for TEM. We first floated the grid on the surface tension of a drop of nitric acid, which

dissolved the copper grid and left the replica floating. We then used a wire loop to catch the replica on a thin water film (as in a child's soap bubble tool) and transfer it - all the time floating on surface tension - to a series of drops of deionized water, where the remains of acid are rinsed off. Finally we catch the replica on a block of mica.

When an SFM image is acquired with a given scanner, a feature at a real space point (X_r, Y_r) will be reproduced at some image coordinates (x, y) , which are related to the real space coordinates through a transformation $(x, y) \rightarrow (X_r, Y_r)$. If the scanner were ideal, the transformation would be a simple proportionality

$$\begin{cases} X_r = Ax \\ Y_r = Ay \end{cases} \quad (1)$$

for some constant A. For a real, imperfect scanner, we will have to consider a more general transformation. If we limit ourselves to second order terms, the most general transformation has the form

$$\begin{cases} X_r = A_x x + A_{xx} x^2 + A_y y + A_{yy} y^2 + A_{xy} x y \\ Y_r = B_x x + B_{xx} x^2 + B_y y + B_{yy} y^2 + B_{yx} x y \end{cases} \quad (2)$$

Figure 7-2 illustrates the role of the various terms of equation (2): A_x and B_y are the "usual," linear calibration coefficients in the X and Y directions, A_y and B_x describe "skew", A_{xx} and B_{yy} represent "nonlinearity", while A_{yy} and B_{xx} describe "curvature," and A_{xy} and B_{yx} are "trapezoidal" terms. Not all of these terms are needed: the

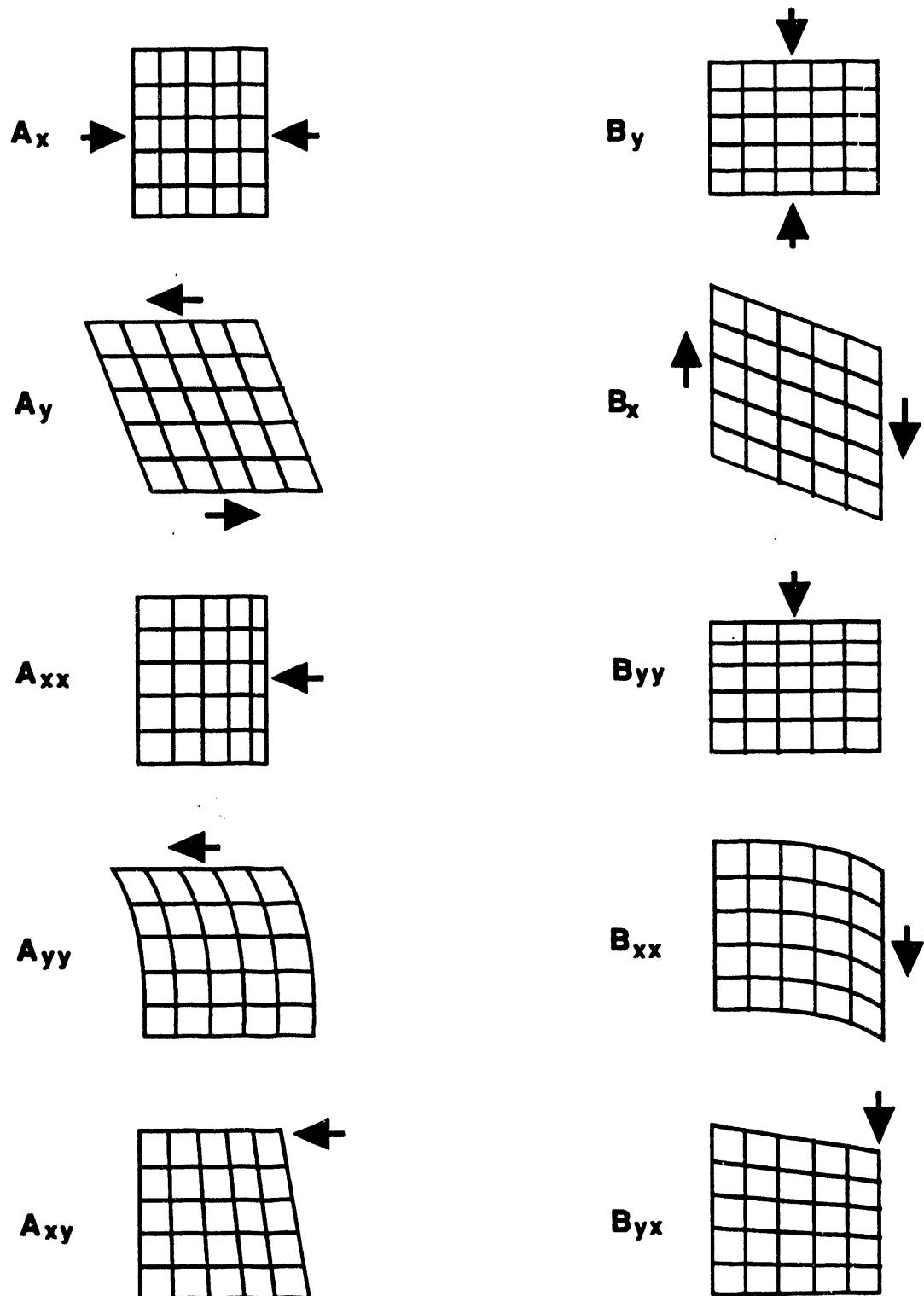


Figure 7-2. The geometric significance of the various terms of equation (2). Each figure shows the effect of an increase in a particular term on an SFM image of a square grid.

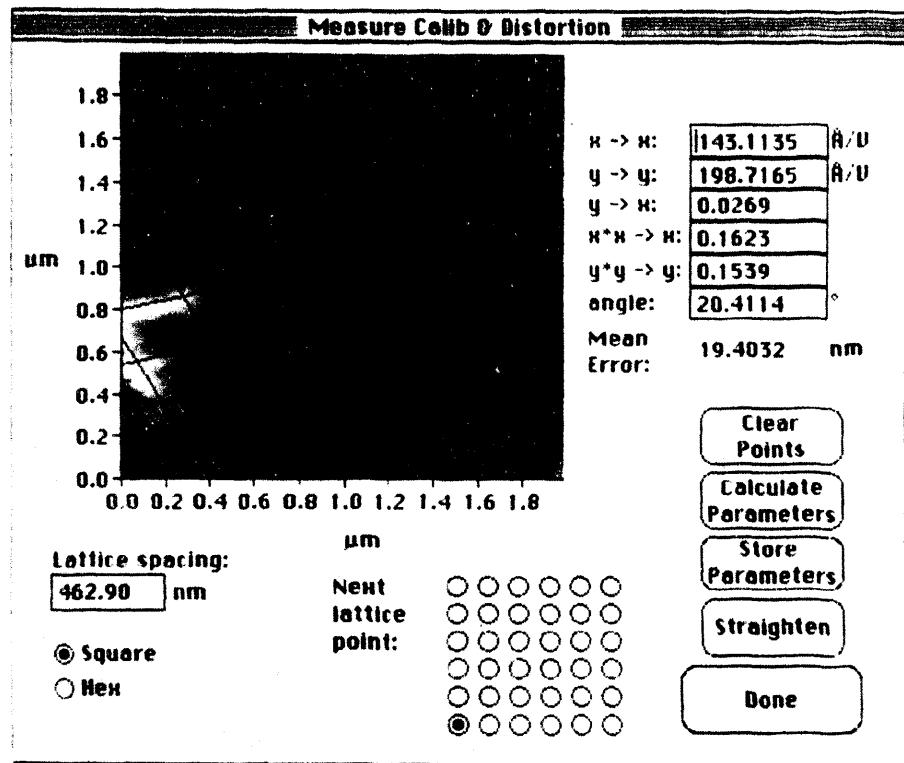
two skew terms taken together describe one single actual distortion plus a rotation. The rotation is irrelevant, since the orientation of the sample is unknown anyway. Furthermore, the curvature and trapezoidal terms are in one sense higher order, since they are second order coupling terms between X and Y motion, and already the linear coupling terms (the skew terms) are zero in the ideal case. We have not observed any visible curvature or trapezoidal effects in grid images, so these terms are small. If we drop these terms, and one skew term, we gain an important benefit: the resulting transformation

$$\begin{cases} X_r = A_{xx}x + A_{xxx^2} + A_{yy} \\ Y_r = B_{yy} + B_{yyy^2} \end{cases} \quad (3)$$

is easily invertible, while the full transform is not. This simplification is extremely valuable since we will need to calculate the inverse transform 65 000 times in order to straighten a single image.

To measure these coefficients for our particular scanner, we analyze SFM images of cross-ruled grids, such as the one in Fig. 7-1 (a), using the dialog box in Fig. 7-3 (a). The user points to the vertices in the grid image with the mouse, and the program finds the set of coefficients which gives the best fit (in the least squares sense) of these points to a square grid with the known lattice spacing. (Actually, in addition to the five A and B coefficients mentioned above, one needs to fit two translational coordinates and a rotation angle, since the position and orientation of the sample are

(a)



b As acquired

c Corrected

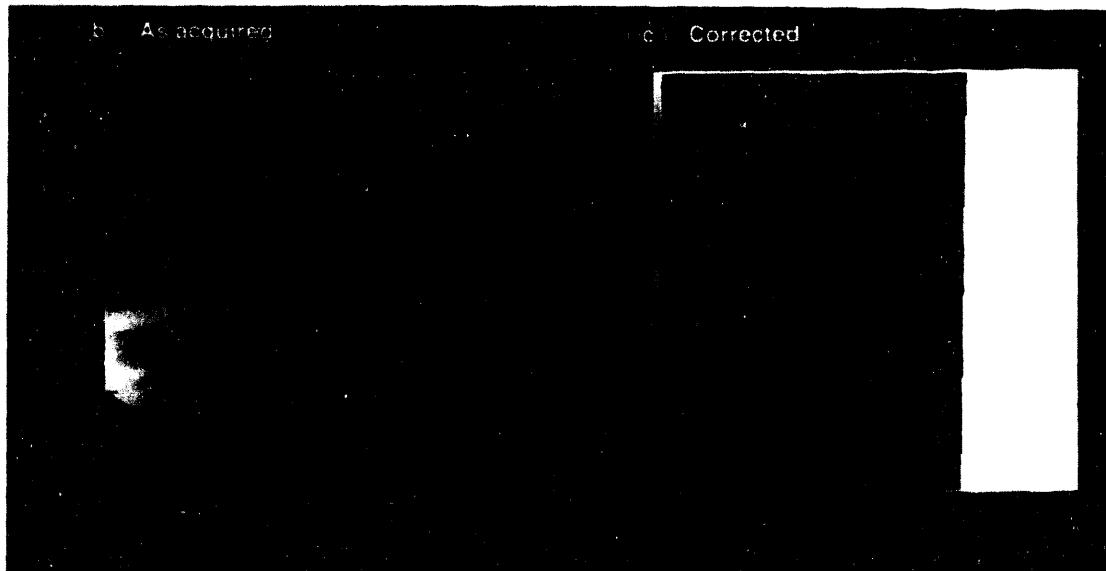


Figure 7-3. (a) The Macintosh dialog box used to find the calibration coefficients. Using the mouse, the user specifies the position of lattice vertices in the image, and indicates in the array of buttons (at bottom center of the dialog) which lattice points they correspond to. The calibration coefficients are then found through a least squares fit. These coefficients can then be used to correct other images acquired under similar conditions (b), (c), or to correct coordinates when making distance measurements.

unknown. The angle cannot be least squares fit directly, since it enters into the transform in a nonlinear way. Instead, the fit of the other coefficients is iterated over for different values of the angle, and the best angle is found with the secant method.) As seen in Fig. 7-3 (a), the grid image corresponding to the resulting fit is then superimposed on the image so the user can verify that a reasonable fit was found, and catch any entry errors.

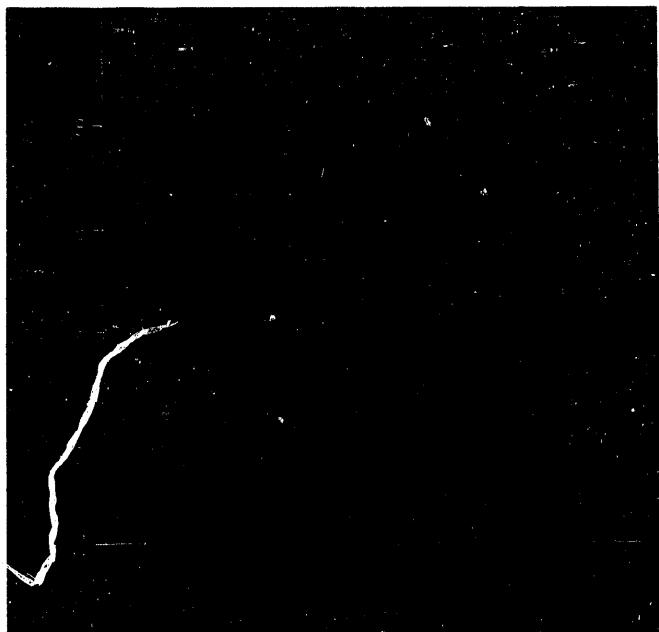
Most images in this thesis are shown uncorrected, but the full correction transform has been used for all distance or curve length measurements.

Once these distortion coefficients have been measured, they can be used to "straighten" other images (taken under the same conditions of temperature, image size and scan speed) to remove the distortions, as shown in Fig. 7-3 (b)-(c) [The reason Fig. 7-3 (c) is turned so narrow by correction is that the original grid image was from one of our earliest runs. After that run, we compensated for the ratio of the X and Y calibrations by adjusting the relative amplitudes of the ramps (using the XYZ matrix mixer, see the electronics section of chapter III). This eliminated the need for such drastic corrections in later data.]

Relative Calibration

The line spacing of the grating replicas, 462.96 nm (21 000 lines/cm) is too wide to allow direct calibration of images smaller than about 1.5 μm . Atomic lattices such as that of mica (Fig. 7-4) can be used for calibration at very small scales, but the limit, set by pixelation, is about 40 nm [Fig. 7-4 (b)]. To address the intermediate size scales, I have written a routine which allows us to use one

(a)



(b)



Figure 7-4. (a) Atomic lattices, like these images of muscovite mica, can be used for calibration at very small size scales, but because of the finite number of pixels in our images, typically 256×256 , they are not usable at image sizes much above 50 nm (b).

image, with known calibration coefficients, as a reference for another, partly intersecting image (typically a subset), and thus deduce the calibration coefficients of the second image. In this way we can calibrate the scanner behavior at intermediate scan sizes without resorting to extrapolation.

The procedure is as follows: the user marks corresponding points in the two images with the mouse (Fig. 7-5). The points in the reference image are transformed to real space coordinates, using the known coefficients for this image. The coefficients of the second image can then be found through a least squares fit of the corresponding points in the second image to these real space positions, in the same manner as in the previous section. These fits can often be done with considerably higher precision than the original grid fits, since very sharp image features (such as intersecting collagen molecules) can be used. Therefore, the problem of error accumulation is not too severe even when the process is repeated.

Using this scheme, we have determined a set of "best fit" coefficients as a function of image size and speed (and temperature). However, the scatter in the fit values between different images is about 5% (1 σ), probably caused mainly by the long term drifts associated with any change of imaging area or retractions of the tip. This implies that an uncertainty of this magnitude remains in any distance measurement.

Z Calibration

To make accurate Z measurements, one would have to go through a calibration process similar to the one described above, and

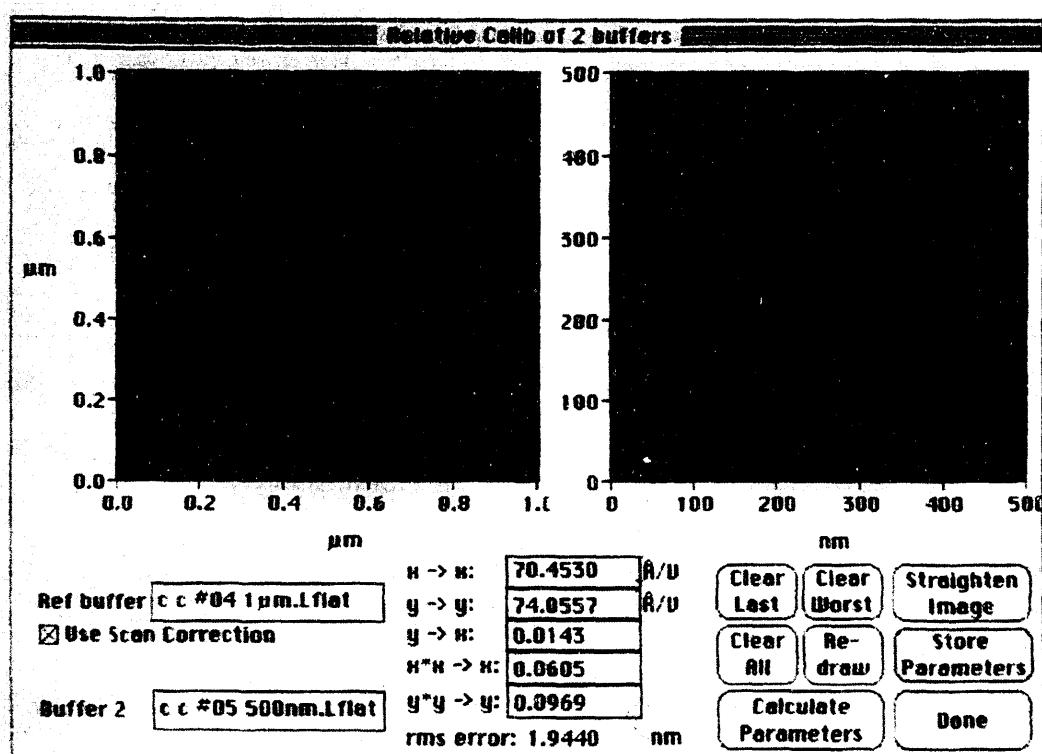


Fig. 7-5. The dialog box used to calibrate one image using as the reference another, partly intersecting image, for which the calibration coefficients are known. The user indicates corresponding points in the two images, and the calibration coefficients are found by a least squares fit.

determine a set of calibration coefficients. Though we have not done this, we have determined an overall, average Z calibration, both at room temperature and at 143 K, using a silicon dioxide grating with a known step height. The grating was fabricated by photolithography, and the step height was determined with profilometry, interferometry and ellipsometry, those instruments in turn being calibrated with commercial reference standards.² Figure 7-6 shows an SFM image of this grating, taken in the open loop mode. By measuring the apparent step height in such images, we could calculate the Z calibration coefficient.

There are two modes of SFM operation which yield quantitative Z information — the feedback mode and the open loop mode — and these derive this information through different means. Thus there are really two Z calibration quantities to determine: the Z response of the piezoelectric scanner, which is a fixed property of the scanner (though it is a function of temperature, and varies slowly with time and thermal cycling), and the responsitivity of the spring deflection measurement, which varies from run to run with the particular spring used and its exact alignment to the beam. However, it is straightforward to determine the ratio of these by applying known voltages to the Z piezo while in hard contact with the sample, and measuring the resulting signal from the detection system. Through this "cross-calibration", one can determine either of the Z calibration coefficients once the other is known.

To some extent, the optical beam deflection method of SFM detection provides internal self-calibration. The conversion from tip Z motion to angular deflection of the reflected beam depends in

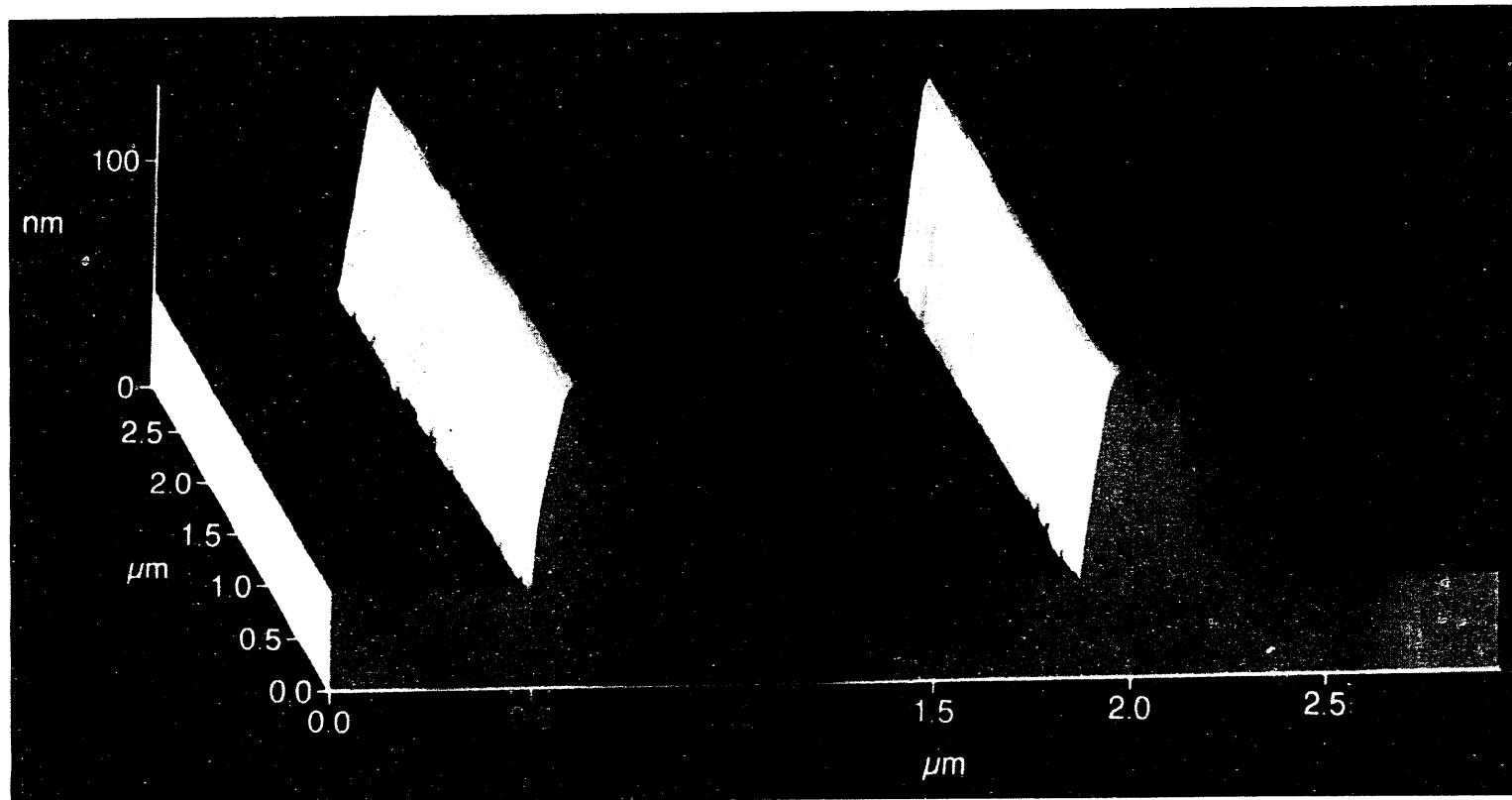


Figure 7-6. An SFM image of the SiO_2 grating used to calibrate the scanner in the z direction. The grating was defined photolithographically, and has a step height of 57 nm and a line pitch of 1.2 μm . This image was acquired in the open loop mode.

principle only on the effective length of the SFM spring (see chapter II), which is a known geometric quantity. The only remaining information we need in order to be able to calculate the relation between the measured signal (from the photo diodes) and the tip motion, is the angular width of the beam. This can be characterized by moving the photo diode in known steps and recording the signal. Thus we have an independent way of determining the Z calibration of the scanner. In practice, there are stray effects, e. g. having to do with stray reflections interfering with the main beam, which can make the Z calibration measured this way differ slightly from the actual value, but this method is valuable as a consistency check. We have made crude measurements using this scheme, and the resulting Z calibration coefficients agree with those determined using the grating images to within 8%.

References:

- 1** Pelco grating replica #604 from Ted Pella, Inc., P.O. Box 2318, Redding, CA 96099, USA.
- 2** VLSI Standards, Inc., 2660 Marine Way, Mountain View, CA 94043, USA.

Chapter VIII: Samples / Images

We have used our low temperature SFM to study several types of biological samples under different conditions: at room temperature in air, at room temperature under liquid, and at 143 K under liquid. In this chapter I will present representative images from these runs, with the emphasis on what can be observed about the samples, and what the abilities and limitations of the SFM are. I will postpone to chapter IX the discussion of the differences between the two temperatures.

To make it possible to compare the low temperature and room temperature data fairly, we needed to make all other scan conditions as comparable as possible. In particular, we needed to acquire room temperature data under liquid (to avoid the meniscus effect, see the introduction), and to have this liquid be as similar as possible to the pentanes used in the low temperature work. For this reason (and because our SFM was not designed to operate in conducting liquids), we chose to take the room temperature in-liquid data under either isopentane or hexadecane, which is a similar aliphatic hydrocarbon but does not evaporate as readily as the pentanes at room temperature.

One might think that the Hansma argument for working under a liquid — that the air/water interface, which is the cause of the meniscus force, is eliminated — would only apply to water or liquids in which water is soluble, since any interface between immiscible liquids carries an interfacial tension. However, our experience is

that hydrocarbons work quite well. In fact, we have not detected any sign of a meniscus effect while operating our SFM in hexadecane, n-pentane or isopentane. Our scan forces in these liquids are usually in the 0.1-1 nN range, while in air forces of 50-100 nN are typical.

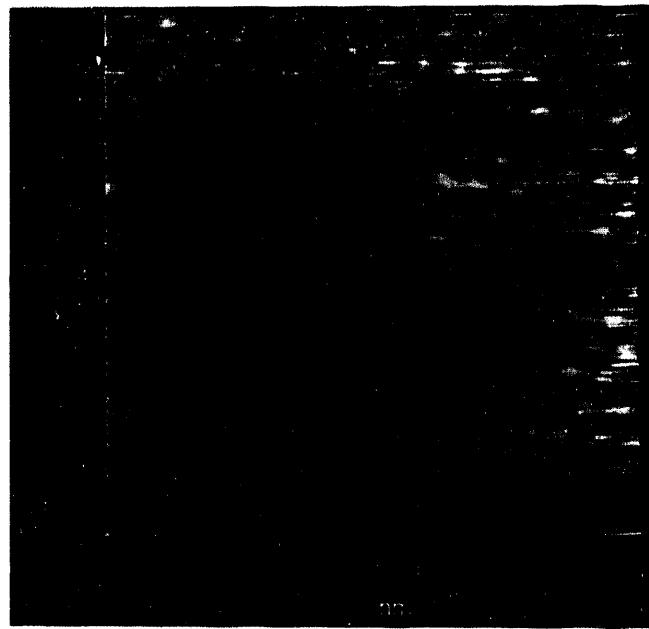
Mica

Figure 8-1 demonstrates the atomic resolution capability of our SFM, acquired in air at room temperature on the canonical choice of material for such demonstrations: muscovite mica. This highly layered material cleaves easily between layers, leaving clean cleavage surfaces that are atomically flat over large areas (many microns across). The crystal plane exposed by cleavage consists of SiO_4 tetrahedra, pairwise sharing corner oxygen atoms, bound together as a honeycomb lattice [Fig. 8-1(b)]. The spacing between the bright dots in Fig. 8-1(a) is 5.2 nm, which is the hole-to-hole spacing of the mica lattice. While the periodicity is clearly resolved, the full honeycomb lattice is not; instead the image is a convolution of this lattice with the atomic structure of the SFM tip, in a similar way as has been discussed in for STM images of graphite.¹

Ferritin

The first biological sample I will discuss is ferritin², an iron storage protein present in similar forms in most organisms from bacteria to vertebrates. The molecule is shaped as a spherical shell, with a hollow interior in which an iron compound is accumulated or released according to the need of the organism. In our specific

(a)



(b)

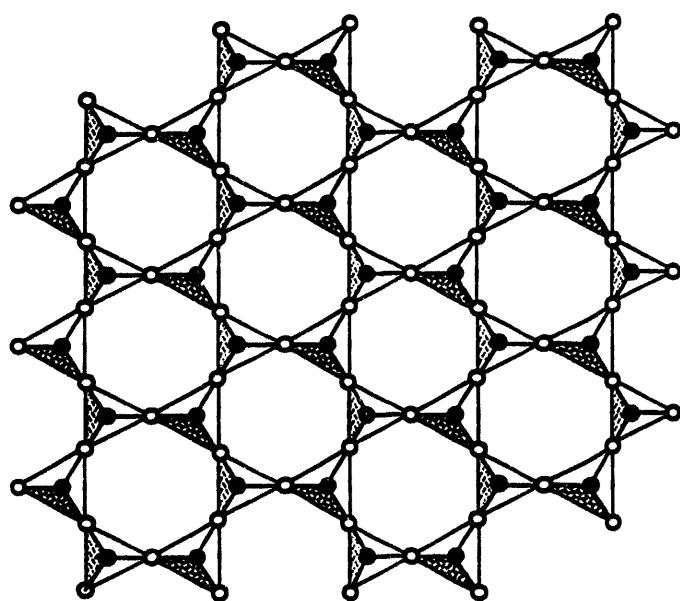


Figure 8-1. (a) An SFM image of freshly cleaved muscovite mica, acquired in air. (b) The cleavage plane of muscovite mica consists of SiO_4 tetrahedra in a honeycomb arrangement. The highest oxygen atoms are marked •, those marked ○ are located slightly lower.

sample, horse spleen ferritin (Fig. 8-2), this shell consists of 24 subunits, and has a diameter of 13nm.

To see a biological sample with the SFM, it needs to be supported by a substrate which is flat on the size scale of the sample in question. If the sample consists of single molecules, this is a stringent criterion. All our sample substrates are based on mica, which is atomically flat over wide areas.

Ferritin, like most proteins, is negatively charged, and therefore does not adhere to pure mica, which has a negative surface charge. However, the mica surface can be coated with polyvalent cations, which willingly adhere to the mica and supply an excess positive charge to which the ferritin in turn can bind. This is an established practice for TEM sample preparation, and the peptide poly-L-lysine is a common choice for the role of the polyvalent cation. We coated freshly cleaved mica sheets with polylysine, applied a 0.5-5 μ g/ml solution of horse spleen ferritin for 15 s, rinsed with double-distilled water, and dried the sample with a burst of nitrogen gas. Note that the nitrogen burst is by no means equivalent to air drying: it serves the crucial purpose of mechanically removing most of the water, thus minimizing the amount that actually evaporates and deposits its content of solutes as contamination on the sample.

Biological samples such as ferritin are significantly more susceptible to damage by the tip than are crystalline samples such as mica (this fact, of course, was one of the main motivations for developing this low temperature SFM). In air, at the high tip forces caused by the meniscus effect, sample damage can be severe enough to make imaging impossible. What typically happens when one

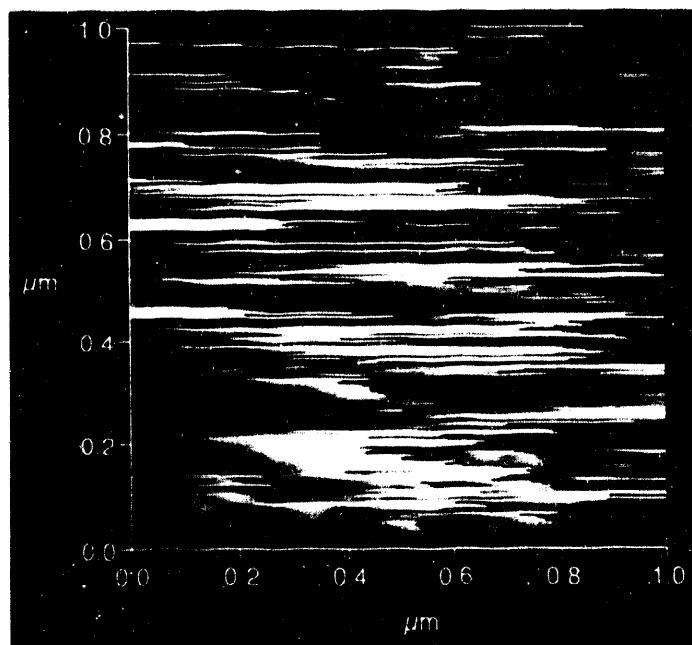


Figure 8-2. A model of horse spleen ferritin. The molecule is built up from 24 subunits which form a spherical shell.

attempts to image a ferritin sample in air at room temperature is illustrated in Fig. 8-3 (a) — the tip does not track, and imaging is impossible. Figure 8-3 (b), a larger scan acquired immediately after (a), shows the reason: the ferritin molecules in the area of the first scan were not only damaged but actually removed by the scan tip. In the surrounding areas, individual ferritin molecules can be seen. The concentration of glitches near the hole is caused by the debris dislodged by the first scan. Figure 8-3 (b) was a rare occasion when a ferritin sample could be imaged in air at all. All other scans of the same sample failed in the manner of Fig. 8-3 (a). That this entire scan appears to have tracked well while others completely fail may point to a "collaborative" aspect of the sample damage: once a molecule is dislodged it may adhere to the tip and make it more "sticky" in some sense, thus increasing the likelihood of further damage in a snowballing fashion.

One direct way to overcome the sample damage problem, while still allowing imaging in air, is to coat the sample with a thin film of a harder material. Figure 8-4 was acquired in air on a sample similar to the one described above, onto which we had sputtered a thin layer of Cr, nominally 2 nm thick. The film anchors the molecules down so that they are no longer dislodged by the tip, and the images are now quite reproducible. However, the presence of the film precludes any hope of resolving details smaller than the film thickness. The images are also characterized by a certain "smeared-out" quality {in the 3D image [Fig. 8-4 (b)] this corresponds to the "tails" on the peaks, which join the substrate smoothly instead of in

(a)



(b)

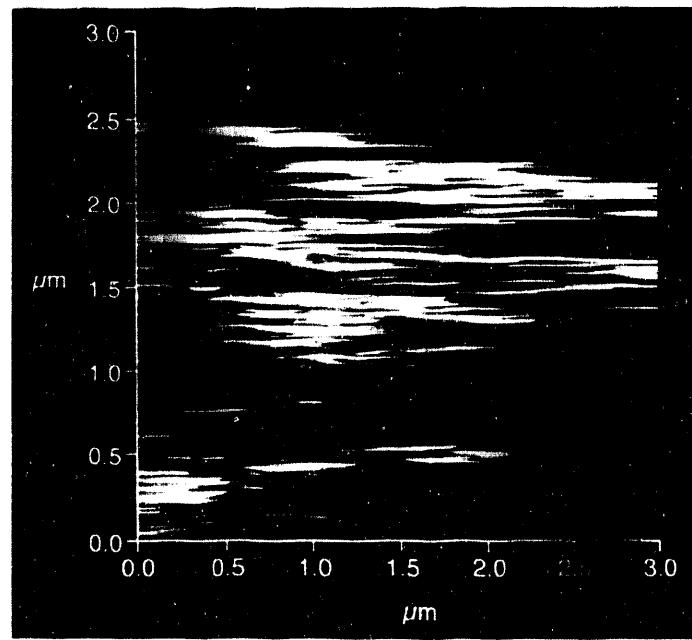
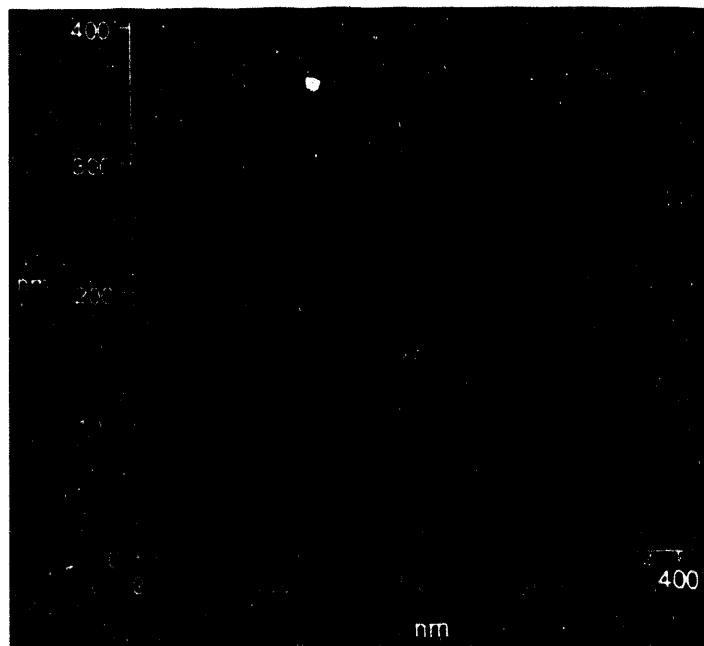


Figure 8-3. An attempt to image ferritin in air. (a) A failed scan. (b) A subsequent scan over a larger area shows how the previous scan removed the ferritin molecules from the scan area. All further scans of this sample failed as in (a).

(a)



(b)

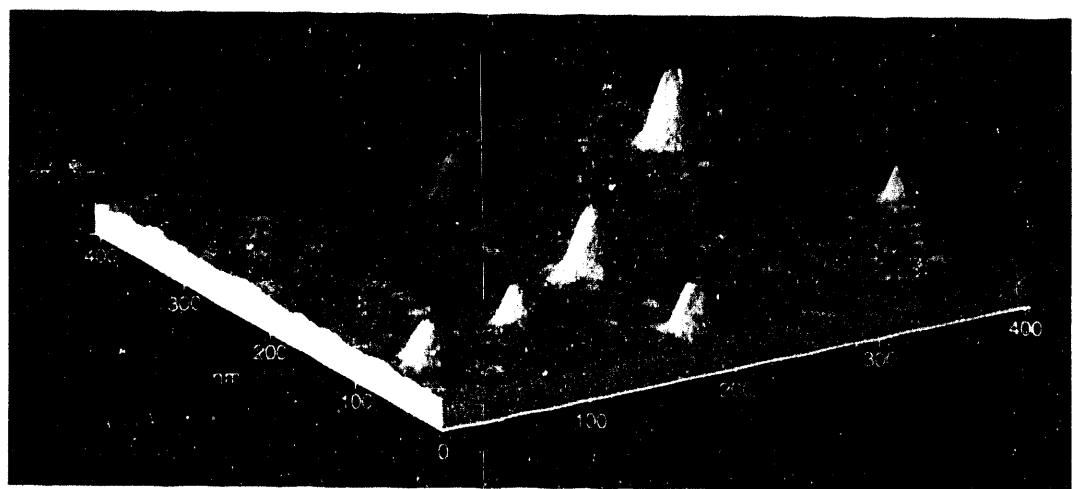


Figure 8-4. A ferritin sample which has been coated with 2 nm of chromium, imaged in air. (a) Top view, (b) 3D view of the same data.

a crisp corner} which is often seen in high force scans, possibly associated with deformation of the substrate itself.

When we imaged the same, coated, sample as in the previous figure under hexadecane, the smeared-out quality disappeared (Fig. 8-5). More importantly, the change from air to hexadecane allowed us to image *uncoated* ferritin samples reproducibly without dislodging the molecules (Fig. 8-6).

Figure 8-7 shows a ferritin sample scanned at 143 K (in n-pentane). This sample had a higher surface coverage — several layers of molecules are visible. Single molecules could be imaged clearly and reproducibly [Fig. 8-8 (a)]. However, the substructure of the ferritin molecule was not resolved [Fig. 8-8 (b)]. [The slight structure seen in the molecules in Fig. 8-8 (b) — the division into two lobes by a diagonal dark band — occurs on all the molecules in the same orientation, a sure sign that it is a property of the tip and not of the sample.]

Purple Membrane

The "Purple Membrane" (PM) is the photosynthetic organ of the salt-loving archebacterium *Halobacterium Halobium* (found, among other places, in the salt flats of the southern San Francisco Bay). It is a patch of the archebacterial cell membrane, and contains a hexagonally packed, near-crystalline array of bacteriorhodopsin (BR) molecules, with a lattice spacing of 6.2 nm. When the BR molecules are exposed to light, they pump protons from the inside to the outside of the membrane, and the resulting concentration difference is used as an energy source by the archebacterium. PM and BR have

(a)



(b)

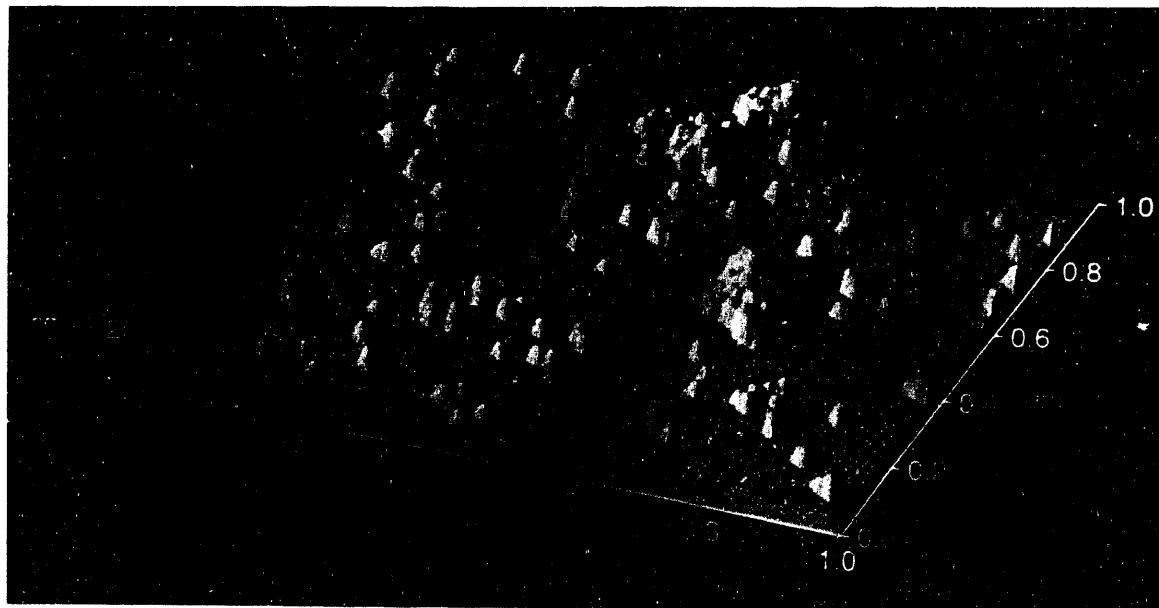


Figure 8-5. The same sample as in Fig. 8-4, imaged in hexadecane (at room temperature). (a) Top view, (b) 3D view.

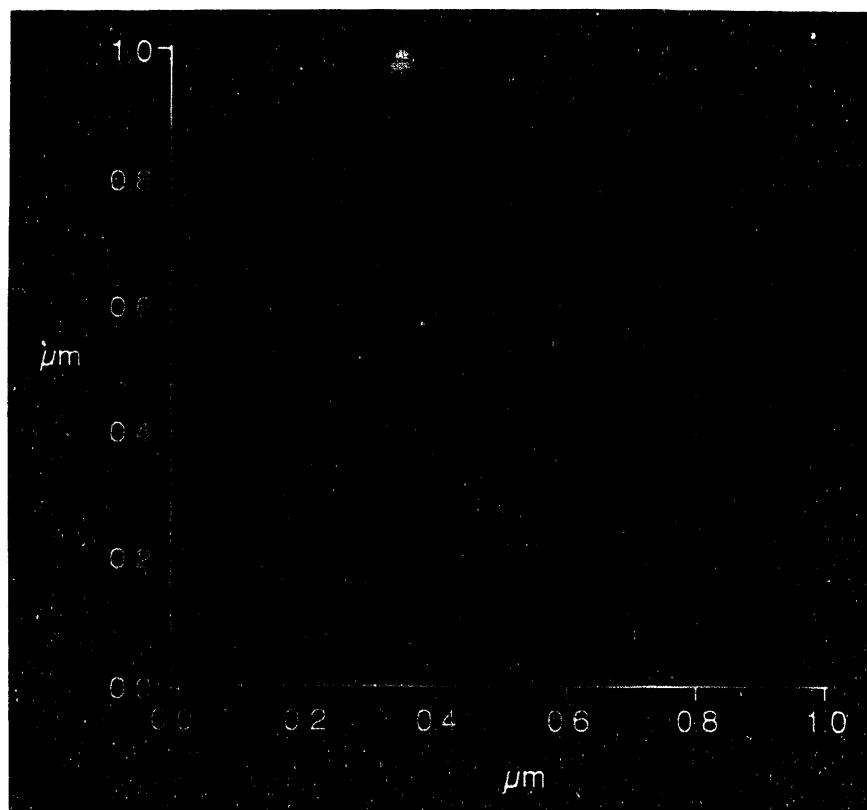


Figure 8-6. An uncoated ferritin sample imaged in hexadecane at room temperature.

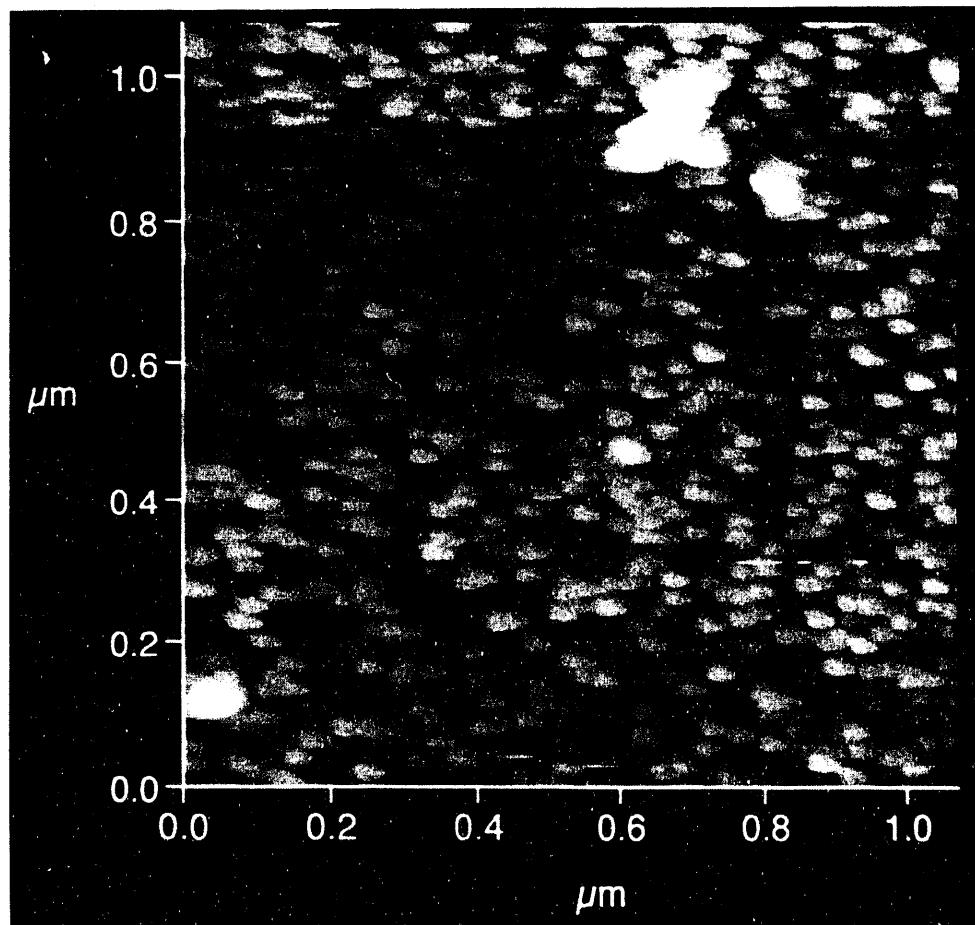
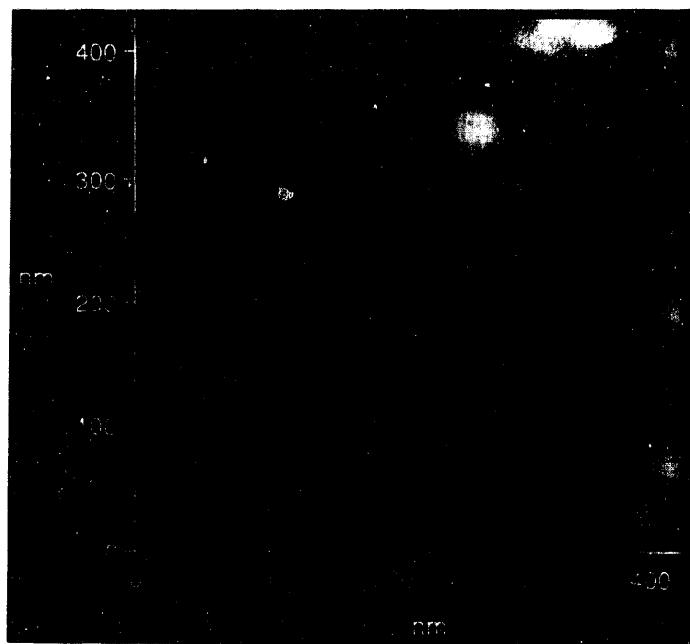


Figure 8-7. An uncoated ferritin sample imaged at 143 K in n-pentane. This sample was heavily covered, and several layers of ferritin can be seen.

(a)



(b)

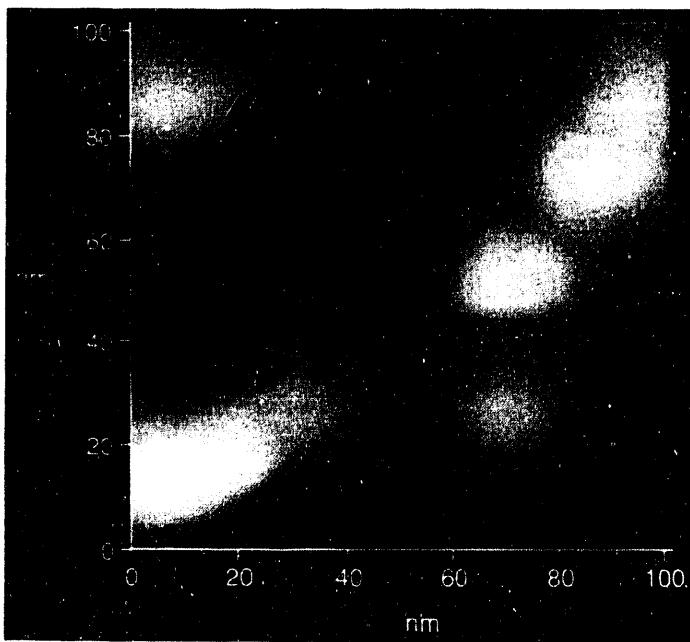


Figure 8-8. Higher magnification scans of the sample in Fig. 8-7. The molecules can be imaged quite reproducibly, but the substructure is not resolved. The two-lobed appearance of the molecules in (b) is a tip shape artifact, not molecular structure.

been subjects of considerable research interest³ since first described in 1971,⁴ and bacteriorhodopsin is now one of the best understood membrane proteins.

From our perspective, there are two reasons to study PM with SFM: (a) the regularity of the BR lattice, which could be expected to be easily visible in SFM, makes it an attractive test sample, and (b) some details of the conformation of the parts of the BR molecules that protrude from the membrane are unknown, so there is the possibility that the SFM could add new information.

Procedures for isolating PM and preparing monolayers of are well established.⁵ Following these, and similarly to the ferritin preparation, we treated freshly cleaved mica with polylysine, applied a solution of PM, and rinsed with double-distilled water after 30 seconds. The samples were then ultrasound treated in distilled water, to remove stacked and overlapping membranes. Finally they were dried with a burst of dry nitrogen gas.

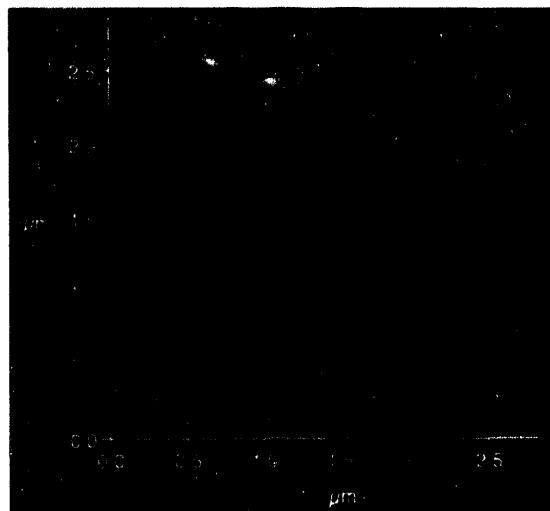
Figure 8-9 is a TEM image of such a sample, prepared by low angle shadowing. Many PM patches, oval shapes about 0.5 μm across, are visible. In some places, membranes that had been overlapping another have been partly removed in the ultrasound step.

Unlike ferritin, PM samples often do withstand SFM imaging in air without being totally disrupted. In Fig. 8-10, which was acquired in air, we recognize the PM patches from the TEM image. [Here, some overlapping patches and fragments — the light areas in Fig. 8-10(a) — remained in spite of the ultrasound step.] However, no structure is resolved on the membrane surfaces, and the resolution is poor, as

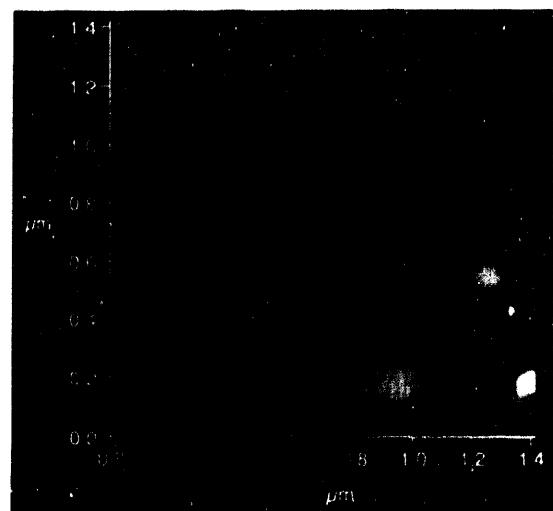


Figure 8-9. A TEM micrograph of purple membrane (PM) on polylysine mica. For TEM, this sample was low-angle shadowed with platinum-carbon.

(a)



(c)



(b)

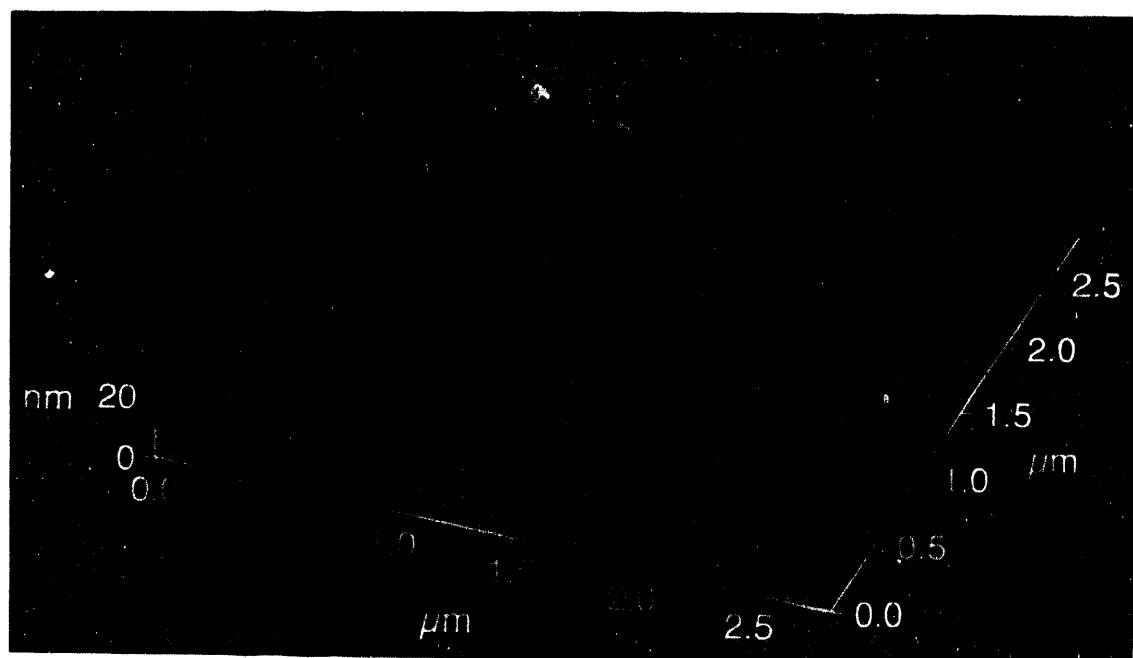


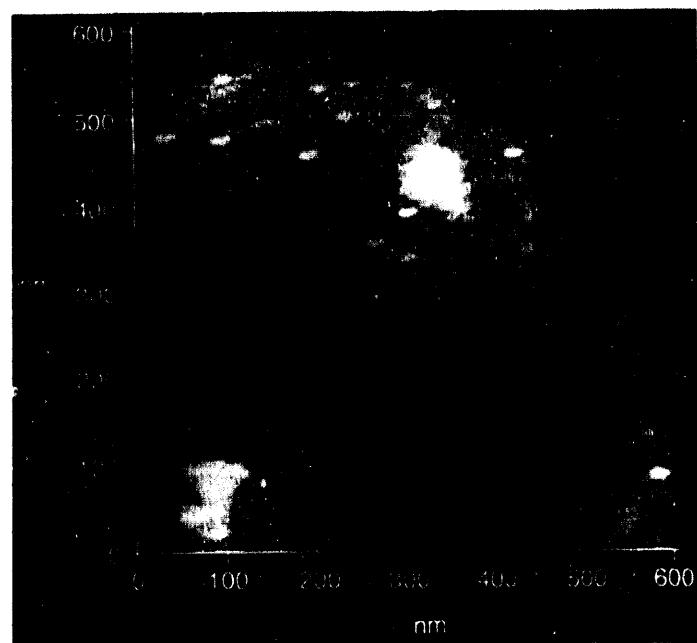
Figure 8-10. SFM images of PM on polylysine, acquired in air. (b) 3D view of the data in (a). Little or no structure is resolved on top of the membrane patches (c).

expected at the necessarily high scan forces used in air [Fig. 8-10(b)].

In hexadecane, using tip forces around 0.5 nN — two orders of magnitude lower than is possible in air — we do resolve details of the structure on the top surface of the membrane (Fig. 8-11). However, the observed structure is irregular and the typical size scale of the features is 15 nm, larger than the BR lattice period of 6 nm. Several effects may contribute to this structure: (a) the roughness (both in a topographic and in a charge distribution sense) of the underlying polylysine layer may distort the arrangement of the BR molecules, (b) surface-tension-related forces during dehydration may create similar artifacts, (c) it may be somehow induced by the scanning tip, or (d) it may be induced by contamination. We find it unlikely that the structure entirely constitutes contamination (on top of an otherwise unaltered BR lattice), because many of the troughs appear to descend deep into the membrane.

In an attempt to avoid possible dehydration artifacts, we devised a scheme to transfer PM samples into hexadecane from the original aqueous solution without ever drying them. The wet sample was first rinsed in acetone (which is miscible both with water and with hexadecane), and then transferred, with a drop of essentially pure acetone, into hexadecane. While some such samples [Fig. 8-12 (a)] showed similar structure as the dehydrated ones, one sample [Fig. 8-12 (b), 8-13] displayed a much smoother surface. (The holes seen in this sample are known to sometimes occur in the ultrasound step, and whether they do or not depends sensitively on the conditions

(a)



(b)

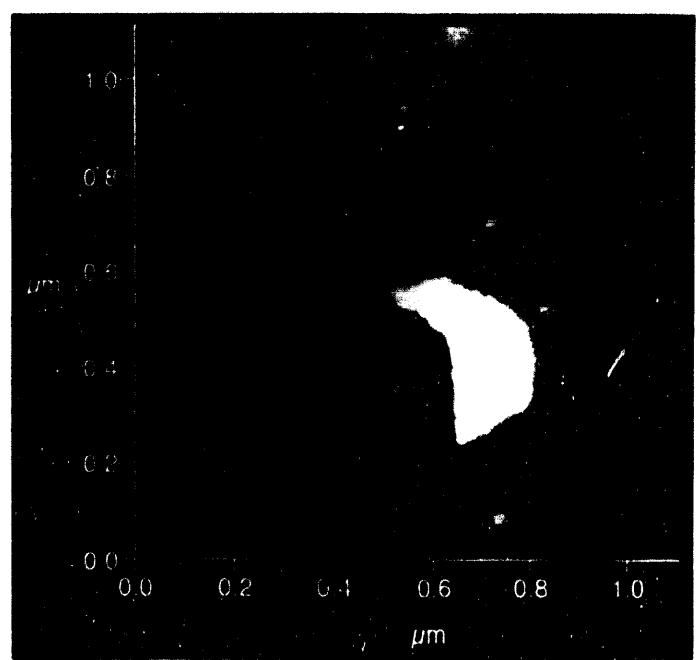
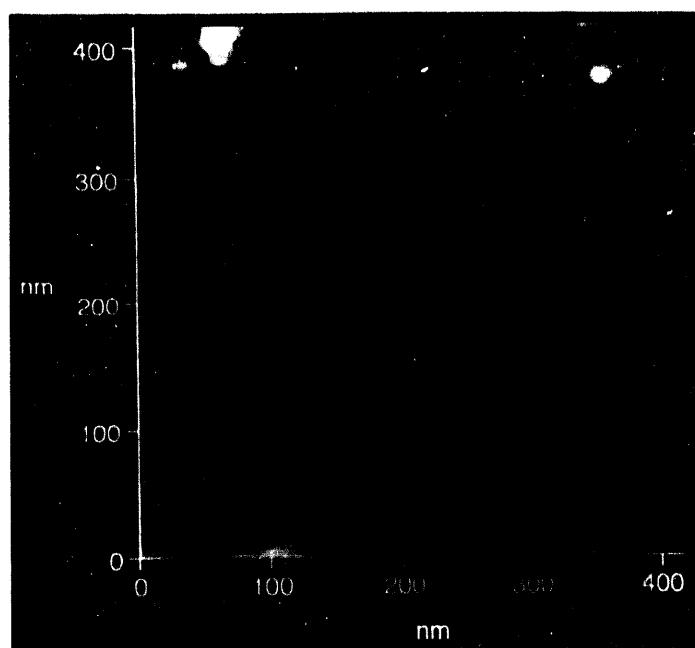


Figure 8-11. Two views of a PM sample imaged in hexadecane at room temperature. Much more detail is resolved than when such samples are imaged in air, but the lattice is not found.

(a)



(b)

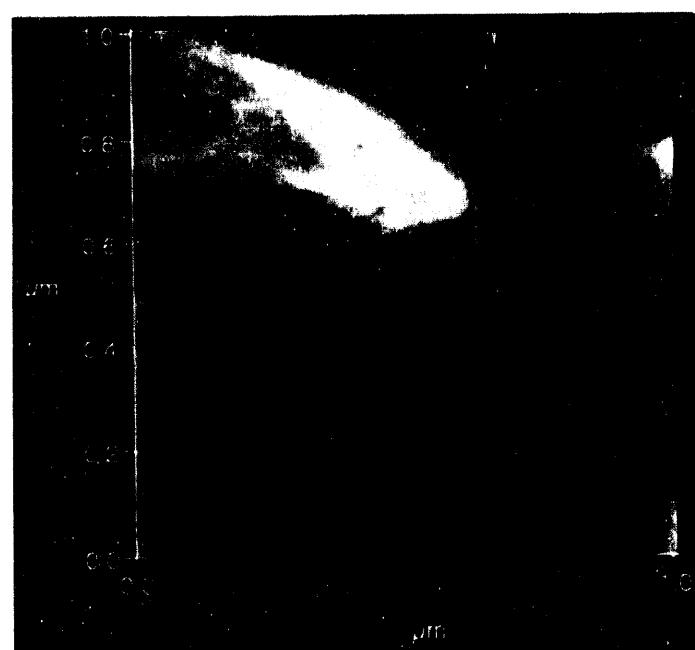


Figure 8-12. PM samples imaged in hexadecane after being transferred from water to acetone to hexadecane without ever being air dried. Some such samples (a) exhibit similar, coarse structure as the air dried samples, but one sample (b) had a much smoother surface.

(a)



(b)



Figure 8-13. (a) A closer look at the membrane surface in the sample in Fig. 8-12 (b) reveals some fine structure of the right size scale to pertain to the BR lattice. (b) The same data, displayed with the gray scale chosen to emphasize the structure. The pattern is oriented at different angles in different parts of the images (arrows), which makes it unlikely that it could be an artifact caused by vibrations or electronic pickup.

during the preparation.) We did not observe a clear BR lattice even on this sample, though some intriguing patterns are seen (Fig. 8-13), most easily if the color scale is adjusted to enhance the surface texture [Fig. 8-13(b)]. The typical repeat period of this pattern is 6.0-6.5 nm, to be compared to the BR lattice period of 6.2 nm. The fact that the observed pattern is oriented at different angles in different parts of the image (arrows) is significant, since any electronic pickup, vibrational noise or other spurious signal source could be expected to give rise to a uniform pattern over the image.

We have also successfully imaged PM at 143 K in n- or isopentane (using the dehydrated type of samples); typical images are shown in Figs. 8-14 and 8-15. These images are not dramatically different from those acquired on similar samples at room temperature in hexadecane: Irregular, mostly coarse structure is seen on the surface of the membranes. Some regions of finer structure [Fig. 8-14 (c)] can be found, as can areas where the structure seems to have some directional order [Fig. 8-15 (b)], but no clear lattice is visible. Fig. 8-15 (b) was acquired in the open loop mode, in such a way that the tip made contact with the sample only on top of the membrane, to lower the force. (This can only be done at the very small "snap-out" that is possible in these liquids, see the introduction.)

[After our PM work was over, the BR lattice was resolved by Hans Butt, working with the Hansma group⁶ at UCSB. They managed to make purple membranes adhere to uncoated mica by adding divalent metal ions (Ca^{2+}) to the buffer solution, which partly compensated the mica surface charge. They imaged these PM samples *in situ* in the aqueous buffer at room temperature. Thus neither any — possibly

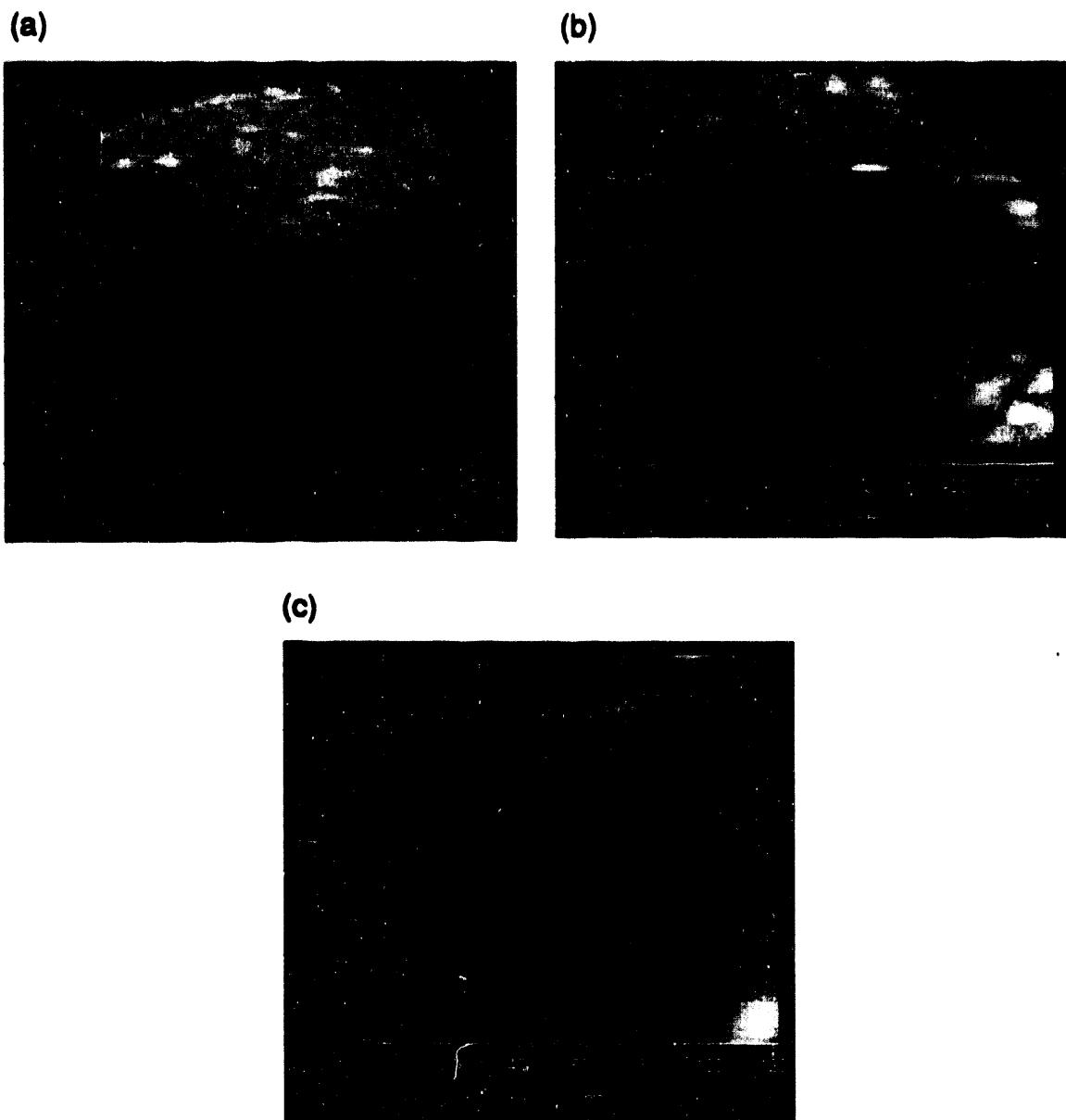
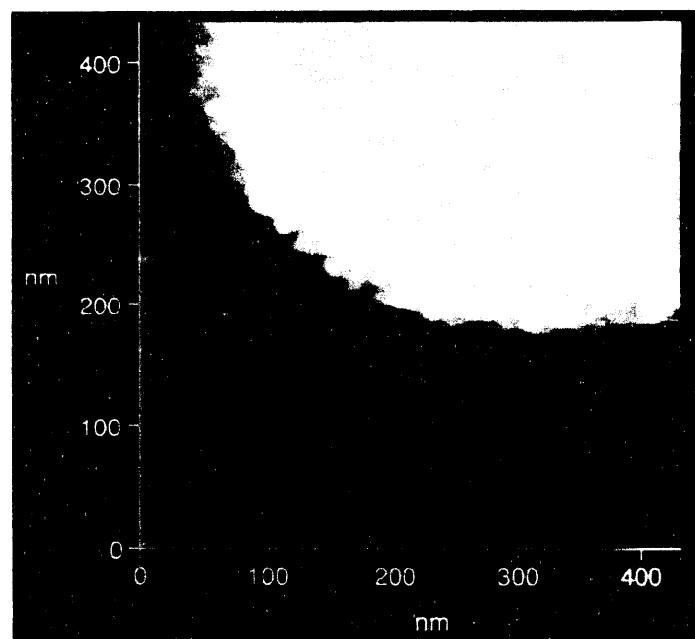


Figure 8-14. (a), (b) Purple membranes at 143 K under n-pentane. (c) Fine structure observed on top of a membrane. [Image (b) was acquired in the open loop mode, in such a way that the tip made contact with the sample only on top of the membrane.]

(a)



(b)

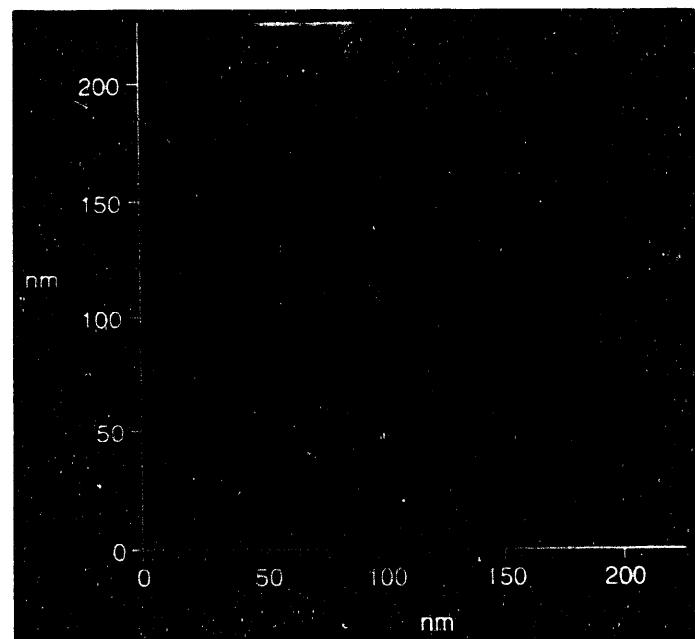


Figure 8-15. PM at 143 K in isopentane. Image (b) is a subset of the top of the membrane shown in (a). The structure on top of some membranes appears to be arranged into rows, but does not form a lattice.

rough — polylysine coating nor any artifact-prone dehydration were necessary. It would be valuable for comparison purposes to image similar samples at low temperature.]

Collagen

The sample on which we have spent most time is collagen, the fibrous protein that is the main structural component of bone, tendon, skin and connective tissues. Collagen occurs in at least 6 different forms that serve different purposes in the body. We studied type I collagen, which is found in skin and tendons and is the single most abundant protein in all higher animals. The type I collagen molecule consists of 3 subunits, two of which are identical, which are arranged in a unique triple helical structure (Fig. 8-16). The resulting molecule is 1.5 nm in diameter with a spiral pitch of 2.9 nm, as measured by x-ray diffraction.⁷ on fibrils. (Collagen tends to aggregate and has therefore not been successfully crystallized.) X-ray diffraction, electric birefringence and light scattering studies⁷ indicate that the length of the molecule is 280-290 nm, while the length calculated from the known sequence and the above diameter and pitch is 296 nm. In the body, these molecules assemble into larger fibrous structures: fibrils, fibers and, ultimately, tendons. The first stages of fibril formation are not well understood, and SFM may be a direct way to study this important process.

Adjacent molecules in such structures normally become joined by covalent cross-links which increase the mechanical strength of the fiber. The collagen we used, which was prepared and generously supplied to us by Arthur Veis⁸, had been derived from rats suffering

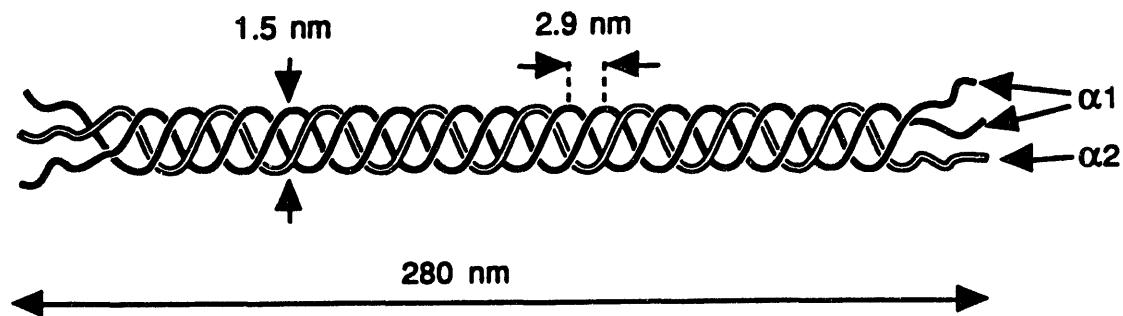


Figure 8-16. The type I collagen molecule.

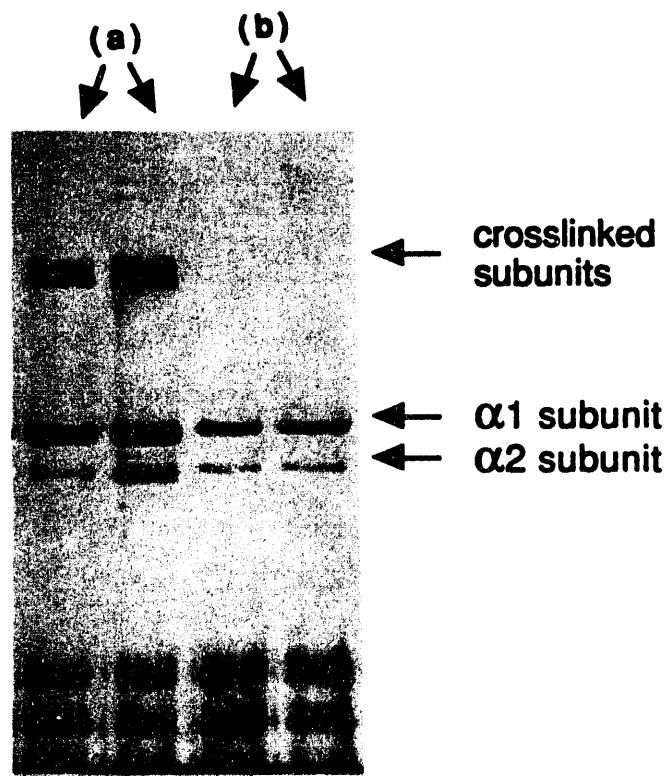


Figure 8-17. An SDS-PAGE gel of (a) collagen retrieved from mica preparations identical to those scanned in the SFM, (b) reference collagen from our stock solution.

from lathyrism, a connective tissue disorder in which these cross-links fail to develop. Such "lathyritic" collagen is therefore easier to separate into single molecules.

Collagen binds well to freshly cleaved mica without the need for polylysine or any other pretreatment. This is probably due to a favorable electrostatic interaction: unlike most proteins, collagen has a positive net charge, while the surface charge of mica in water is negative.

We diluted the collagen solution to appropriate concentration, 0.1 to 1 ng/ μ l depending on the surface coverage wanted, with 0.05N acetic acid, then applied drops (0.5 μ l/mm²) of this solution to freshly cleaved mica sheets (already attached to our sample disks). After exactly 5 min, the surface was washed with double-distilled water for 10 s and dried with a burst of nitrogen gas.

To verify that what we saw in our SFM images was indeed collagen molecules, we repeated exactly the same sample preparation procedure on larger pieces of mica, then redissolved the resulting coating using sodium docecy1 sulfate (SDS) buffer, and analyzed it using SDS polyacrylamide gel electrophoresis (SDS-PAGE, a standard biochemistry tool). The gel electrophoresis separates all proteinaceous material in the samples by molecular weight. The resulting size bands (Fig. 8-17) were indistinguishable from those of collagen controls, indicating that the molecules were not damaged by our deposition procedure. (There are two bands of different density because the SDS treatment separates the collagen molecule into its 3 subunits, two of type α -1 and one of type α -2.)

Furthermore, the observed quantity was consistent with the surface coverage observed in our SFM images.

Though we have been able to image collagen at room temperature, under isopentane (Fig. 8-18), most of our collagen data was acquired at 143 K (including all images described in the rest of this section). Figure 8-19 shows the effect of different concentrations of the applied collagen solution: the samples were prepared from solutions of (a) 10, (b) 2.0, (c) 1.0 and (d) 0.4 ng/ μ l, respectively. Collagen molecules are resolved in all three figures, most clearly at the lowest concentration, where they are well separated. Other examples of single collagen molecules are shown in Fig. 8-20.

From such images it is relatively straightforward to measure the length of individual molecules, something that is not possible with the conventional methods which by their nature average the molecular population. The distribution of lengths, if different from a deltafunction, carries information about to the extent to which the molecules relax from their idealized conformation. To make such measurements, we use the facility in Fig. 8-21 (a). The user draws on the image a polygon which approximates the shape of the molecule, and the length of this polygon is calculated [taking the full correction transform (see chapter VII) into account]. The result of 135 such length measurements is summarized in Fig. 8-21 (b). Clearly, the existence of a few very short objects is an indication that some fragments are present in our collagen. Excluding lengths below 210 nm, we find a mean of 244 nm and a standard deviation of 14.0 nm. The mean is slightly shorter than expected from

(a)



(b)

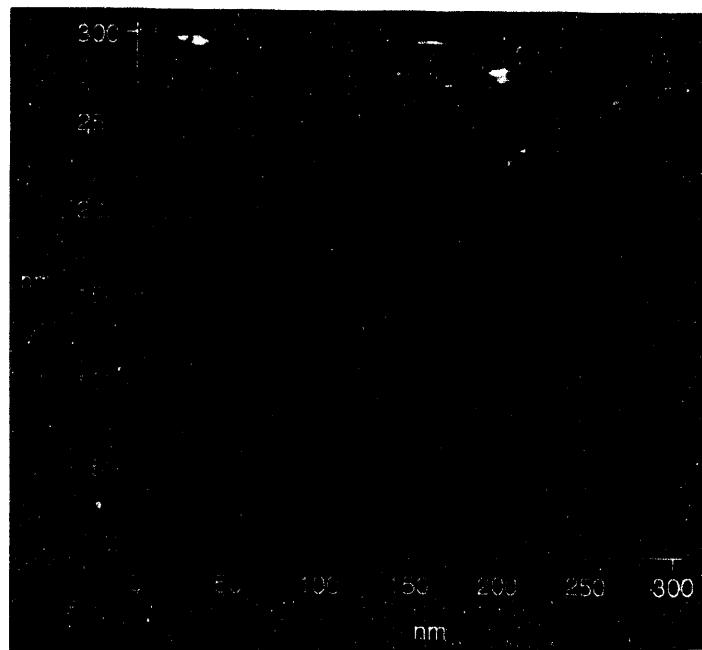


Figure 8-18. Two images of collagen at room temperature under isopentane.

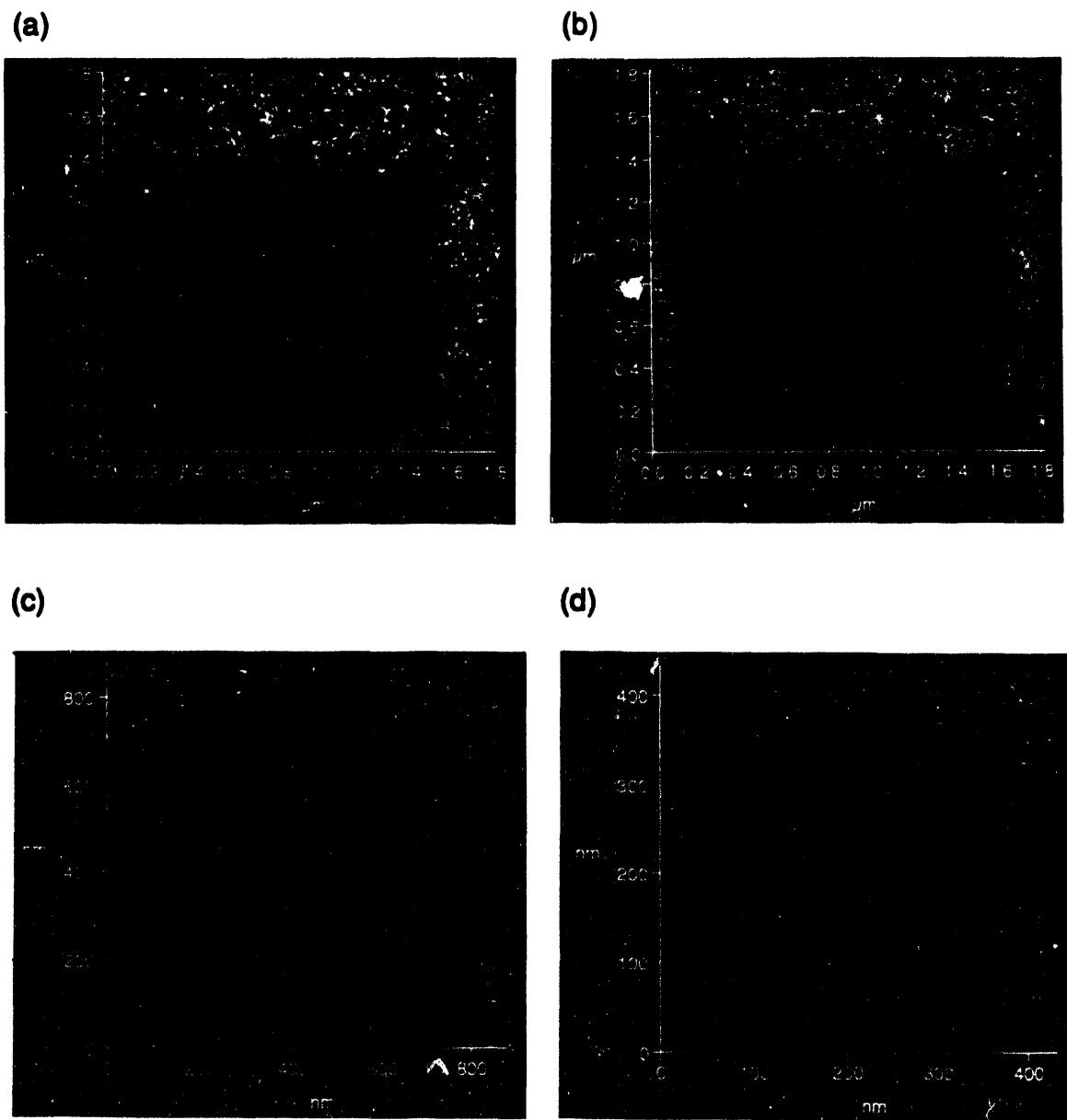
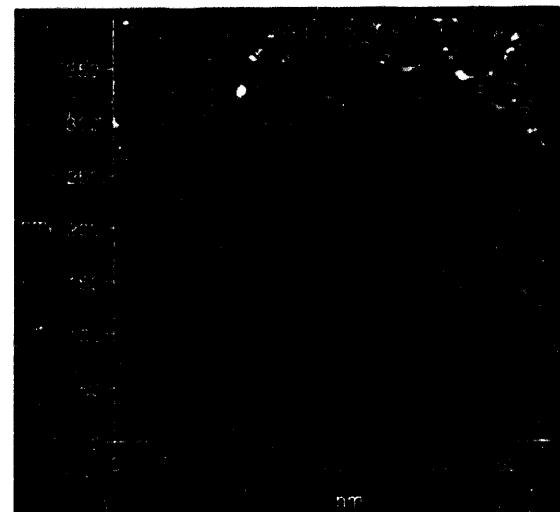


Figure 8-19. Collagen at 143 K under isopentane. The samples were prepared using collagen concentrations of (a) 10, (b) 2.0, (c) 1.0, and (d) 0.4 ng/ μ l. Individual collagen molecules are clearly resolved.

(a)



(b)



(c)

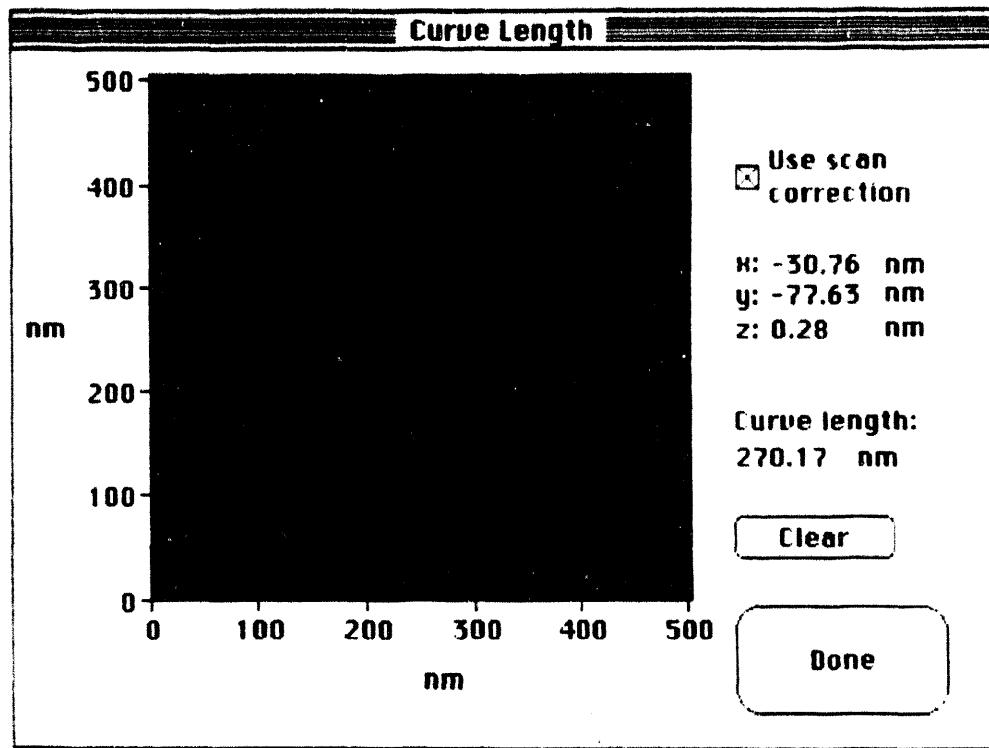


(d)



Figure 8-20. Four sample images of collagen molecules at 143 K under isopentane.

(a)



(b)

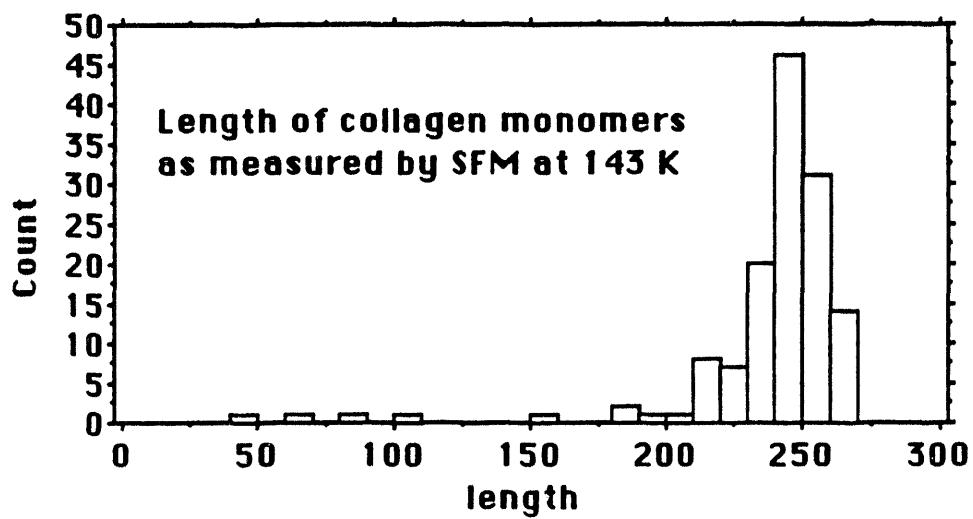


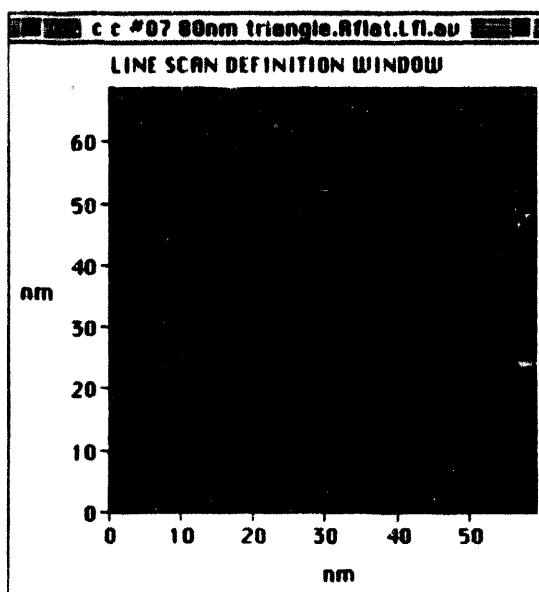
Figure 8-21. (a) The dialog used for measuring curve lengths. (b) A summary of measurements of the lengths of 135 different collagen molecules, using images such as those in Fig. 8-20.

conventional techniques mentioned in reference (2) (280-290 nm), but is consistent with similar measurements from TEM images, in our lab (247 ± 28 nm) and elsewhere⁹ (251 ± 47 nm, calculated including some outliers on the long side which may have been dimers). The smaller standard deviation of the SFM measurements seems to imply that the variability found in the TEM data was mainly due to measurement error, and indicates the higher precision possible with SFM.

The apparent width of the collagen molecules — our simplest measure of the lateral resolution of the SFM — varies from run to run (and with time within a single run) depending on the condition of the tip. At both temperatures, the narrowest widths we measure are about 4 nm (Fig. 8-22), to be compared to the accepted width of the collagen helix as measured by conventional techniques, which is 1.5 nm.

The triple helix of the collagen molecule is not resolved [Fig. 8-23 (a)]. However, the SFM images do show details that are difficult to resolve in TEM images, and which may contain information about the collagen molecule. Examples are pointed out in Fig. 8-23 (b): (i) Some molecules have points of sharply increased curvature, i. e. "kinks" (heavy arrows). If more data of this type were gathered, the distribution of the position at which such kinks occur could be studied. If certain positions of such kinks are found to be overrepresented, these positions would presumably indicate areas where the helix is less stable, and could be correlated to published predictions¹⁰ of the existence of such regions. (ii) In several places, the end of one molecule appears to attach to another molecule

(a)



(b)

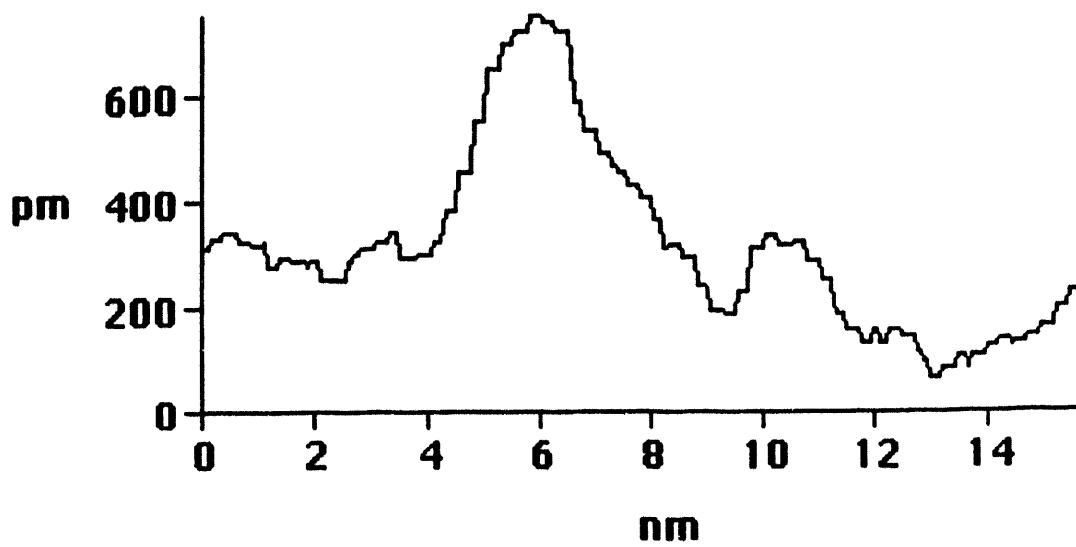
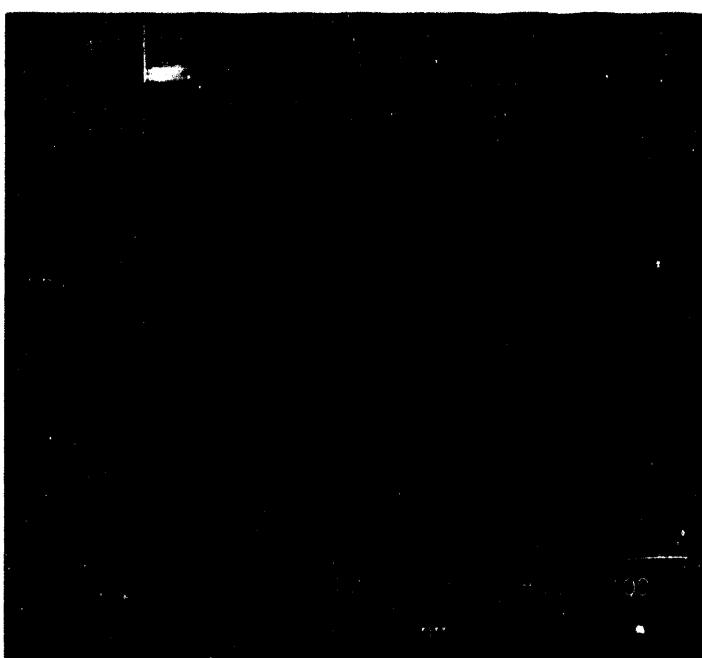


Figure 8-22. A line scan across a collagen molecule, showing the apparent molecular width. The height profile along the line defined in (a) is shown in (b).

(a)



(b)

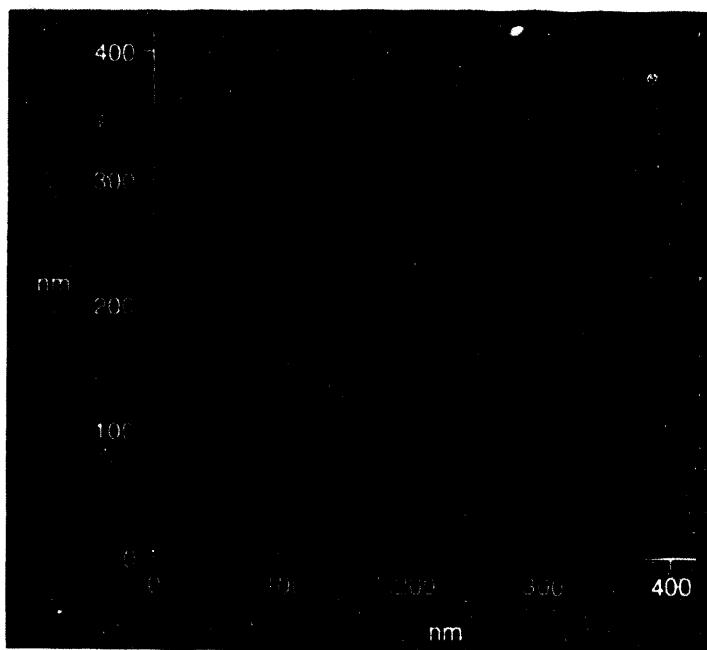


Figure 8-23. (a) High magnification images do not resolve the collagen triple helix. (b) However, they do show details difficult to detect by other methods, such as sharp kinks (heavy arrows), associations (medium arrows) and "bumps" (light arrows). The distribution of the position (along the length of the molecule) of such features may contain information about domains of increased reactivity or lessened conformational stability.

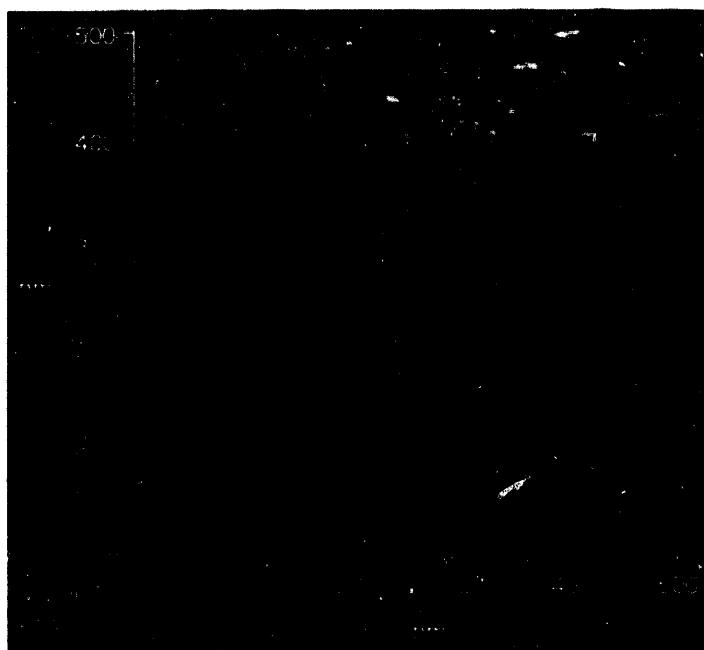
(medium arrows). Statistical studies of these attachments may yield insights into intermolecular interactions. (iii) Conspicuous "bumps" are found on some molecules (light arrows). While these quite likely represent the decoration of the collagen by contamination, one could do similar statistics of the position at which they occur, and any nonuniformity in this distribution would be indicative of the relative reactivity of different sites along the molecule.

One collagen sample we studied (Fig. 8-24) had adopted a different structure, closer to a network then to an assembly of individual molecules. Note that this is not simply a matter of high surface coverage: here very few free molecular ends are visible, most being attached to other molecules. [Compare with Fig. 8-19 (b) and (c).] A recurring element in this sample is a "twine", where two molecules join each other for a distance, possibly in a spiral fashion [Fig. 8-24 (b), between arrows]. Note the increased height of the molecular pair in the "twined" part, and the height variations along the molecules in this region (suggestive of their being intertwined). Such details would be difficult to detect in TEM data.

RNA

The last sample I will describe is double-stranded RNA (dsRNA). The dsRNA molecule forms a double helix similar to that of DNA, with a diameter of 2.6 nm. Our dsRNA samples were prepared (and generously supplied to us) by Yuri Lyubchenko of the University of Arizona.¹¹ Briefly, his procedure was as follows: mica substrates

(a)



(b)

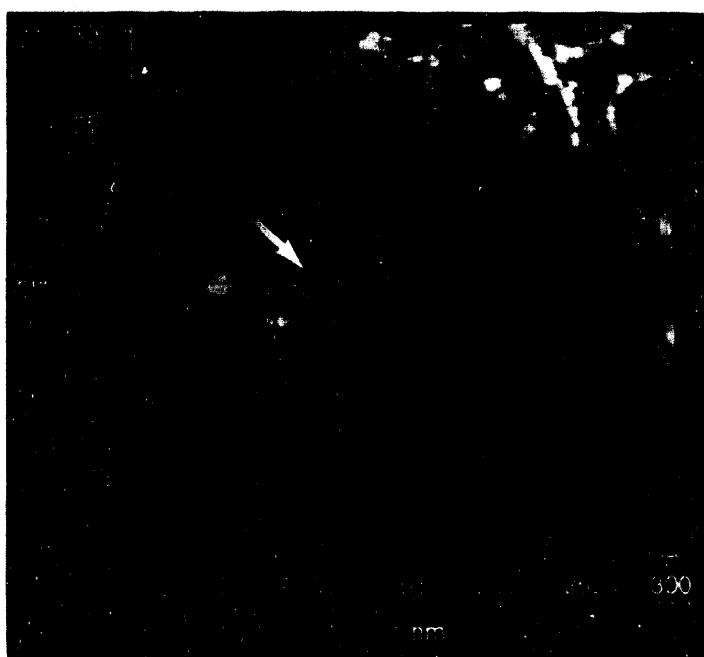


Figure 8-24. A collagen sample that exhibits the network structure. Note how molecules interact in pairs [(b), between arrows], and that very few free ends are visible.

were first exposed to the vapor of 3-diaminopropyl triethoxysilane (APTES), allowing APTES molecules to bind covalently to the mica through their silane groups. The APTES-coated mica sheets were then immersed in a 0.1 μ g/ml solution of reovirus dsRNA in buffer for 2 hours, rinsed in deionized water, blotted at the edge and vacuum dried.

We have imaged these samples in air at room temperature (Fig. 8-25) as well as in isopentane at 143 K (Fig. 8-26). (The fact that these samples can be imaged in air without removing the molecules indicates that the dsRNA is strongly bound to the APTES.) The images may be reminiscent of our collagen data, but these RNA molecules are much easier to see than collagen molecules because of their larger diameter, and unlike collagen they are of greatly varying length.

Two types of features occur repeatedly in the images. In several instances a molecule appears to be folded in half and twisted up on itself like a twist-tie [Fig. 8-26 (a)]. Other molecules appear to be wound many times around a loop, like a coil of rope [Fig. 8-26 (b)]. Figure 8-27 shows what appears to be a case where such a coil has become partly unraveled, demonstrating that it indeed consists of many strands.

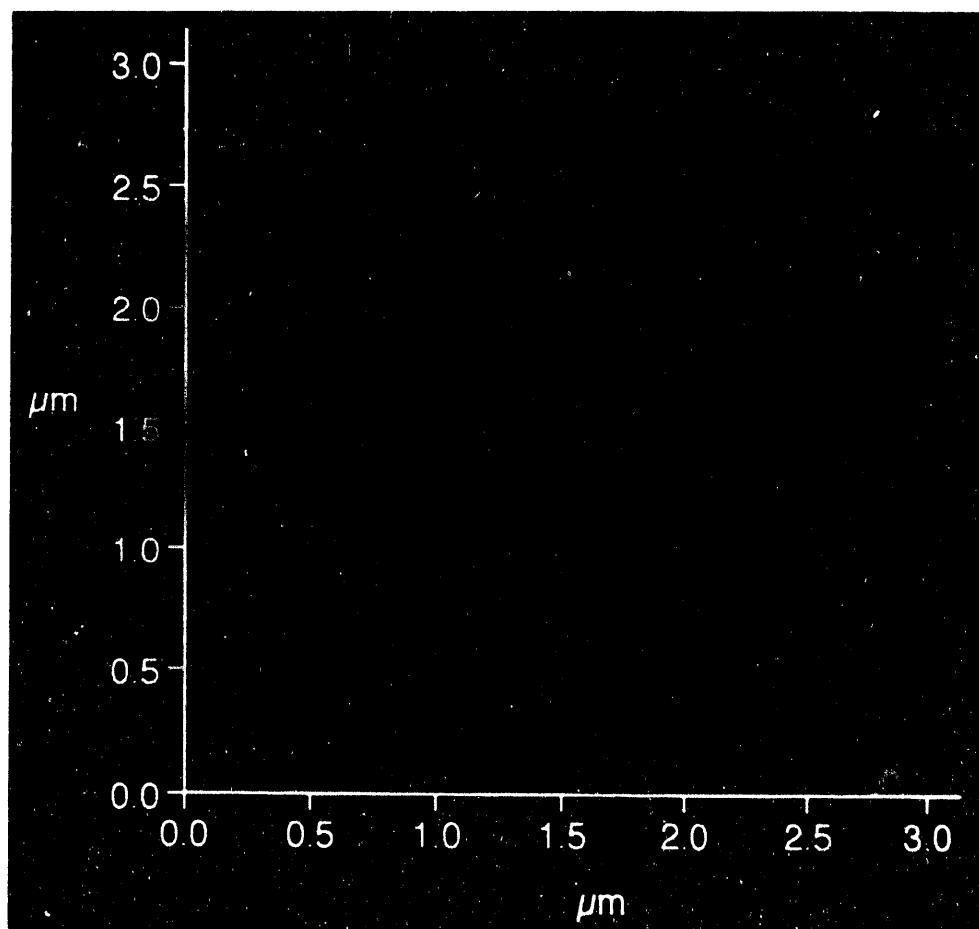
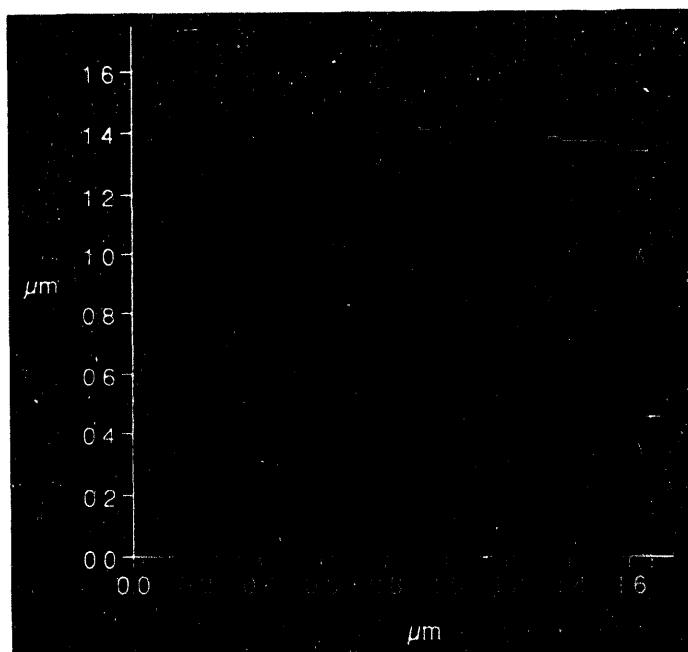


Figure 8-25. Double-stranded RNA on silanated mica, imaged in air.

(a)



(b)

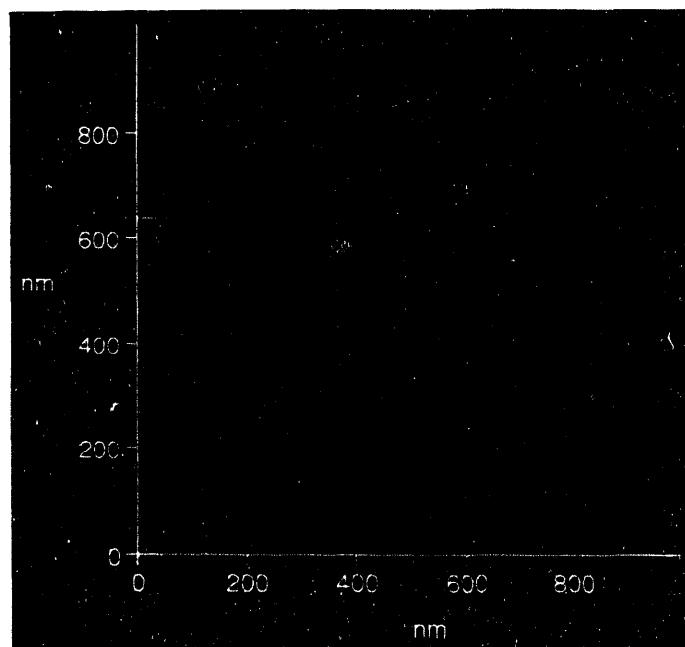
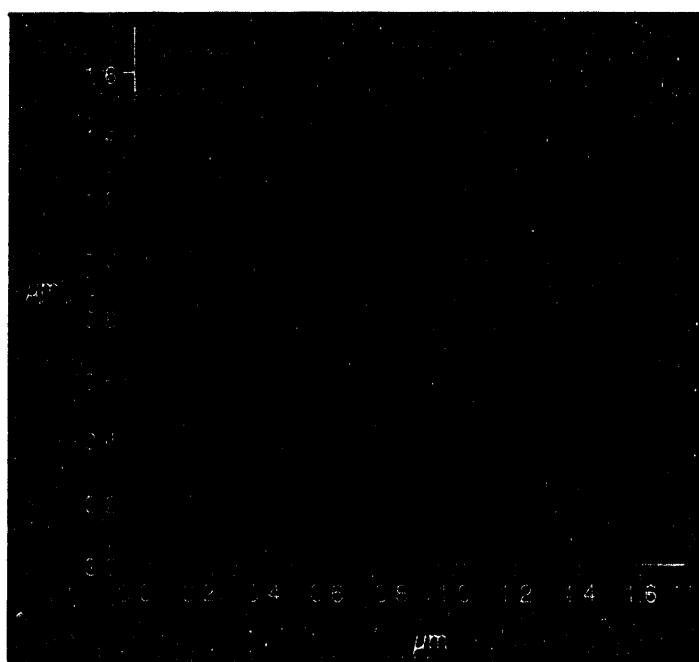


Figure 8-26. Double-stranded RNA on silanated mica, imaged at 143 K in isopentane. Some molecules loop back on themselves to form "coils of rope" (a) and "twist-tie" configurations (b).

(a)



(b)

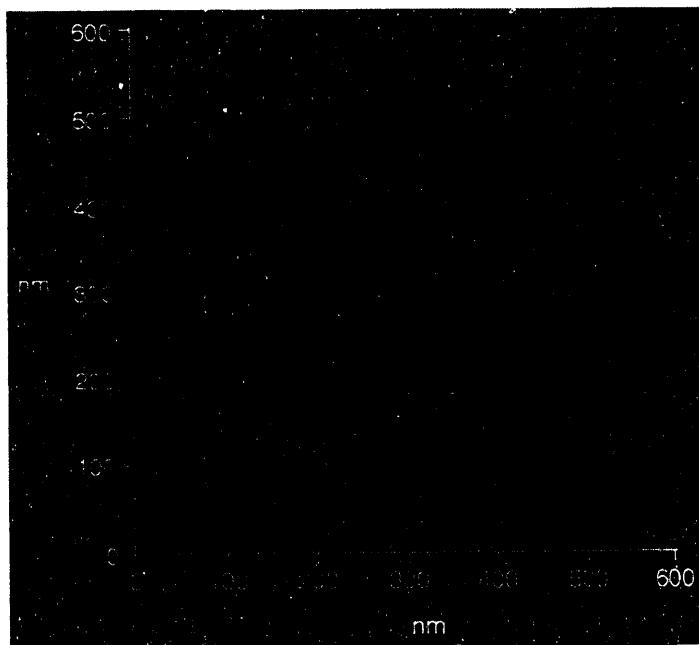


Figure 8-27. Double-stranded RNA at 143 K in isopentane. In the area shown in (a) several "coils" of RNA have partly unraveled — particularly clearly seen in the smaller scan in (b) — demonstrating how they are indeed composed of many turns

References:

- 1 E. D. Ganz, Ph. D. thesis, Physics Department, University of California at Berkeley, 1988.
- 2 E. C. Theil, *Adv. Enzymol. Relat. Areas Mol. Bio.* **63**, 421 (1990).
- 3 h. G. Khorana, *J. Biol. Chem.* **263**, 7439 (1988).
- 4 D. Oesterhelt, and W. Stoeckenius, *Nature. New Biology* **233**, 149 (1971).
- 5 S. B. Hayward, D. A. Grano, R. M. Glaeser, and K. A. Fisher, *Proc.Nat. Acad. Sci. USA* **75**, 4320 (1978).
- 6 H. J. Butt, K. H. Downing, and P. K. Hansma, *Biophys. J.* **58**, 1473 (1990).
- 7 K Kuhn, *in Structure and Function of Collagen Types*, R. Mayne and R. E. Burgeson, eds, Academic, New York, 1987, p 4-7; P. H. von Hippel, *in Treatise on Collagen*, G. N. Ramachandran, ed., Academic, New York, 1967, p. 260-263.
- 8 A. George and A. Veis, *Biochemistry* **30**, 2372 (1991).
- 9 A. Veis, personal communication.
- 10 H. Hofmann, T. Voss, and J. Engel, *J. Mol. Bio.* **172**, 325 (1984).
- 11 Y. L. Lyubchenko, B. L. Jacobs, and S. M. Lindsay, *Nucleic Acids Research*, **20**, 3983 (1992)

Chapter IX: Cold vs. Warm

Our motivation in building a low temperature SFM was the hope that the samples would turn more rigid, thus less deformable and better able to withstand the tip force. This would improve imaging in two respects: less sample deformation would improve the resolution, and less irreversible sample damage would improve the reproducibility. In the previous chapter, I have described the images acquired under different conditions with little commentary; now I will turn to the question of evaluating what advantages, if any, low temperature operation of an SFM can confer.

While it is difficult to be quantitative due to the variability from run to run, the general impression left by the data presented in the previous chapter is that the resolution achieved at 143 K is similar to the one achieved at room temperature. The run-to-run variation in image quality seems to imply that the resolution determining factor is the shape (and contamination status) of the tip. Figure 9-1 demonstrates that the streaky quality seen in some images is related to tip contamination: at the point of the scan indicated by arrows, the imaging quality suddenly *improved* in this respect, presumably due to contamination moving off the tip onto the sample. The image in Fig. 9-2 (of ferritin molecules in hexadecane at room temperature) is a clear example of tip shape artifacts, the repetition of the same pattern in the same orientation being an obvious indication that the pattern derives from the tip and not from the sample. (One can look at this as a reversal of the roles of tip and sample: we are actually imaging the scan tip with a ferritin

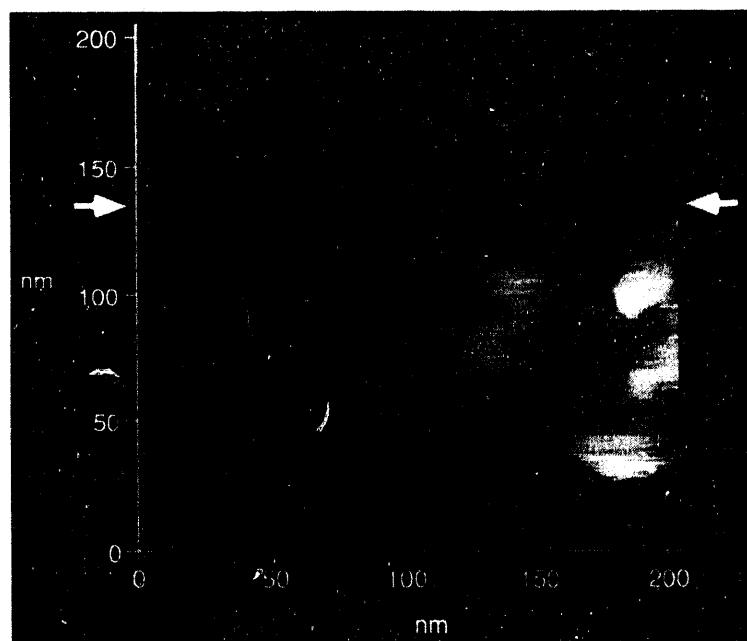


Figure 9-1. The "streakiness" in many images is related to tip contamination. This is demonstrated here by the fact that the image quality abruptly *improves* in mid scan (arrows), presumably because contamination was removed from the tip.

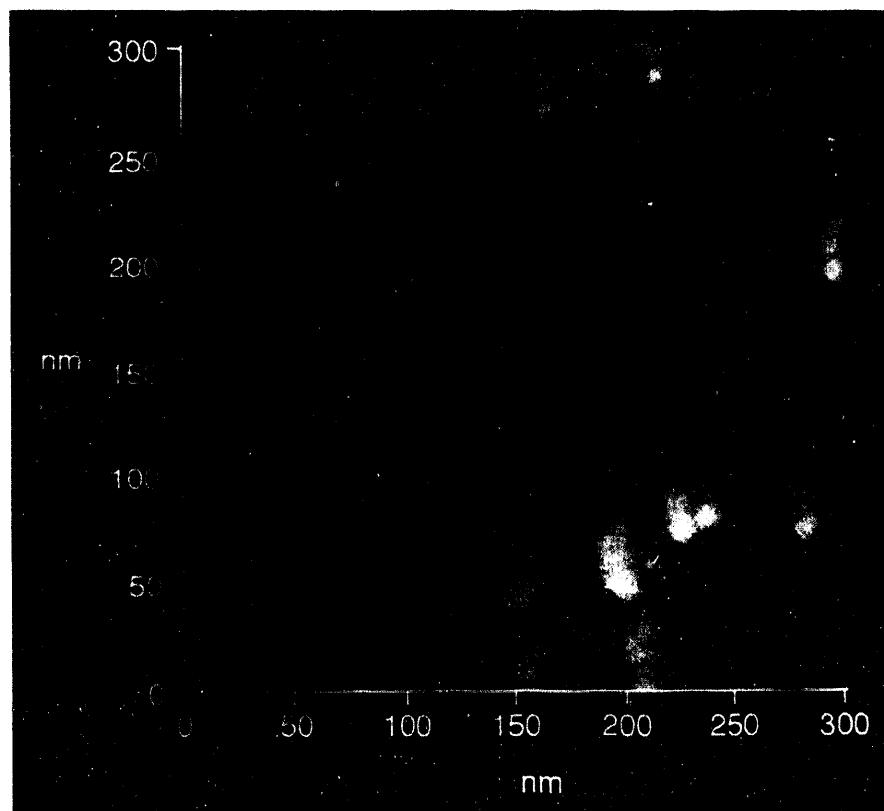


Figure 9-2. Ferritin in hexadecane at room temperature, scanned with a "bad" tip. The fact that the same pattern is repeated many times indicates that it represents the shape of the tip, not of the sample. At the same time, the narrow radius of the highest peak of each instance implies that it is possible to get high resolution images of ferritin at room temperature, given a sufficiently sharp tip.

molecule as the "tip," the latter in some sense being sharper.) However, the central maximum of each instance of the pattern is quite sharp, with a diameter of 13-15 nm, to be compared to the ferritin molecular diameter of 12 nm. These sharp maxima tell us that there was a sharp protrusion at the apex of the scan tip, but also that with a tip as sharp as this protrusion, ferritin molecules can be imaged at room temperature with only a marginally exaggerated radius. This implies that the larger apparent radius seen in other ferritin images can be ascribed to imperfect tips. It thus appears that the resolution of SFM on ferritin is at present limited by the tip shape even at room temperature. My impression is that this is the case for all the biological samples described in chapter VIII.

This does not necessarily imply that there is no improvement in sample rigidity at the lower temperature, only that sample rigidity is not the limiting factor on the samples we have studied, with the tips that are available today. (We have used several types of tips: our own fractured silicon nitride tips described in chapter III, e-beam deposited gallium-doped tips¹, and commercial etched silicon tips², all with similar limiting resolution.) There are samples which we have not been able to image at all at room temperature — such as the protein alpha-lytic protease and individual molecules of alcian blue — but we have been equally unable to image these at 143 K.

Leaving resolution aside, one can try to evaluate the sample resistance to damage by subjecting it to successively higher forces, and seeing when drastic sample damage occurs. Such a "force

series," applied to collagen at room temperature in isopentane, is shown in Fig. 9-3. Figure 9-3(a) is the "before" image, taken at 0.25 nN. As the force is increased, the images get streakier [Fig. 9-3(b), at 4 nN], but the molecules remain unchanged. At 16 nN [Fig. 9-3(c)], the image is nearly destroyed, but when returning to low force [Fig. 9-3(d)] it is seen that the molecules still remain (though the tip has now deteriorated). Only at 29 nN [Fig. 9-3(e)] can actual damage to the collagen be detected (the missing fragment of a molecule at the arrow). A rescan at low force [Fig. 9-3(f)] confirms this damage (white arrow) and the deposition of debris (black arrow). A second example of such a force series, on a different sample but still at room temperature, is shown in Fig. 9-4. At 32 nN [Fig. 9-4(b)], the image has turned streaky but the collagen molecules are unchanged. The first signs of damage to the collagen are seen only after a 64 nN [Fig. 9-4(c)] scan. When a larger surrounding area is then imaged at low force [Fig. 9-4(d)], the area of the high force scan can be distinguished by the partial removal of the collagen, and the concentration of debris at the edges. After a scan at 125 nN, the collagen is totally removed [Fig. 9-4(e)].

Figure 9-5 shows the effect of such a force series at 143 K. [The doubled appearance of the molecules in Fig. 9-5(a) is due to the scan tip having two points.] Again the image turns streakier as the force is increased [Fig. 9-5(b), at 22 nN], but damage to the collagen is not detected until the force has been raised to 44 nN [Fig. 9-5(c)]. As the low force rescan shows [Fig. 9-5(d)], a collagen fragment has been entirely removed (arrows). An overview scan [Fig. 9-5(e)] shows

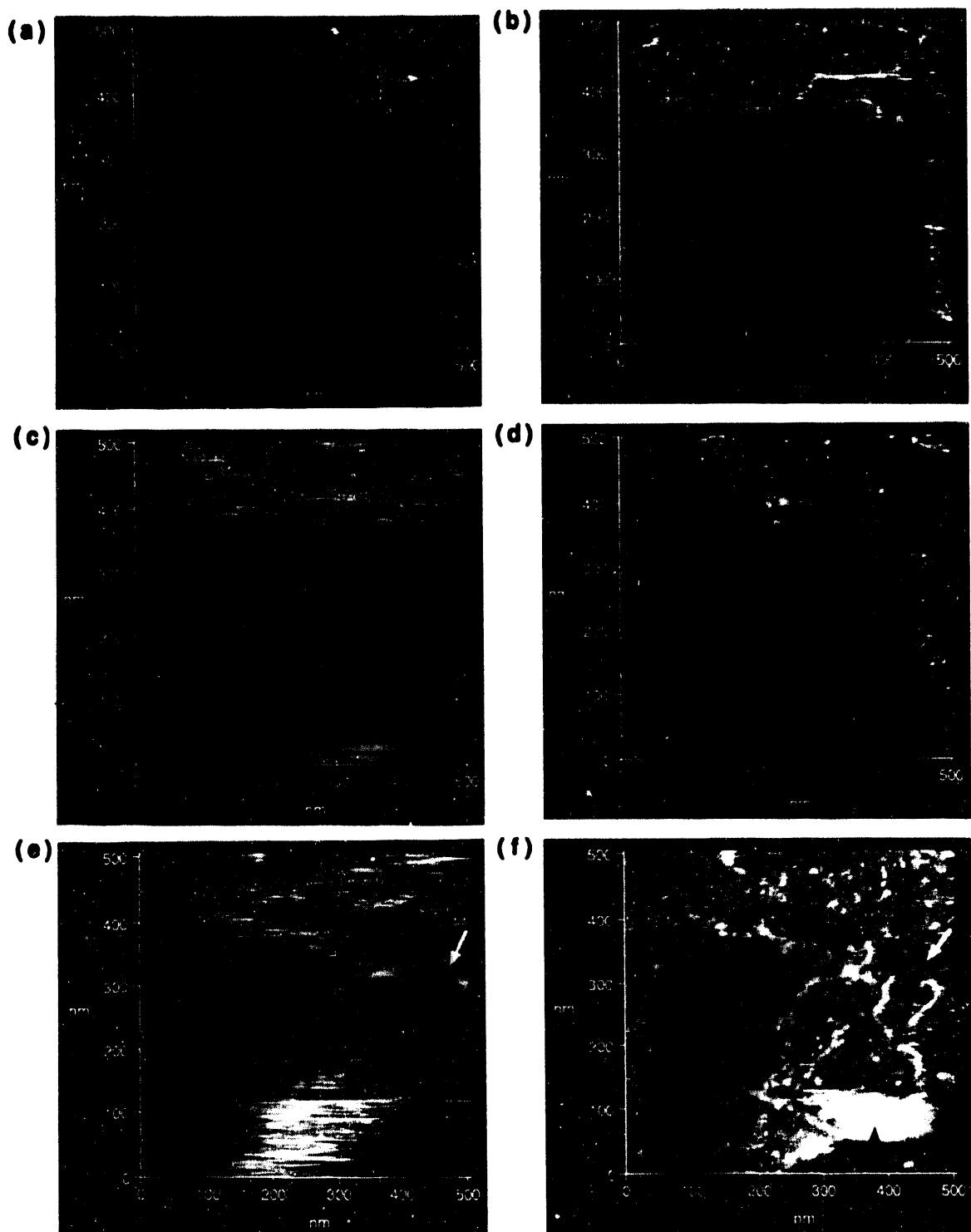


Figure 9-3 Sample images from a force series on collagen in isopentane at room temperature. The same area scanned at: (a) 0.25 nN, (b) 4 nN, (c) 16 nN, (d) 1 nN, (e) 29 nN, (f) 1 nN. Molecular damage was first seen at 29 nN [white arrows in (e) and (f)], accompanied by the deposition of debris [black arrow in (f)].

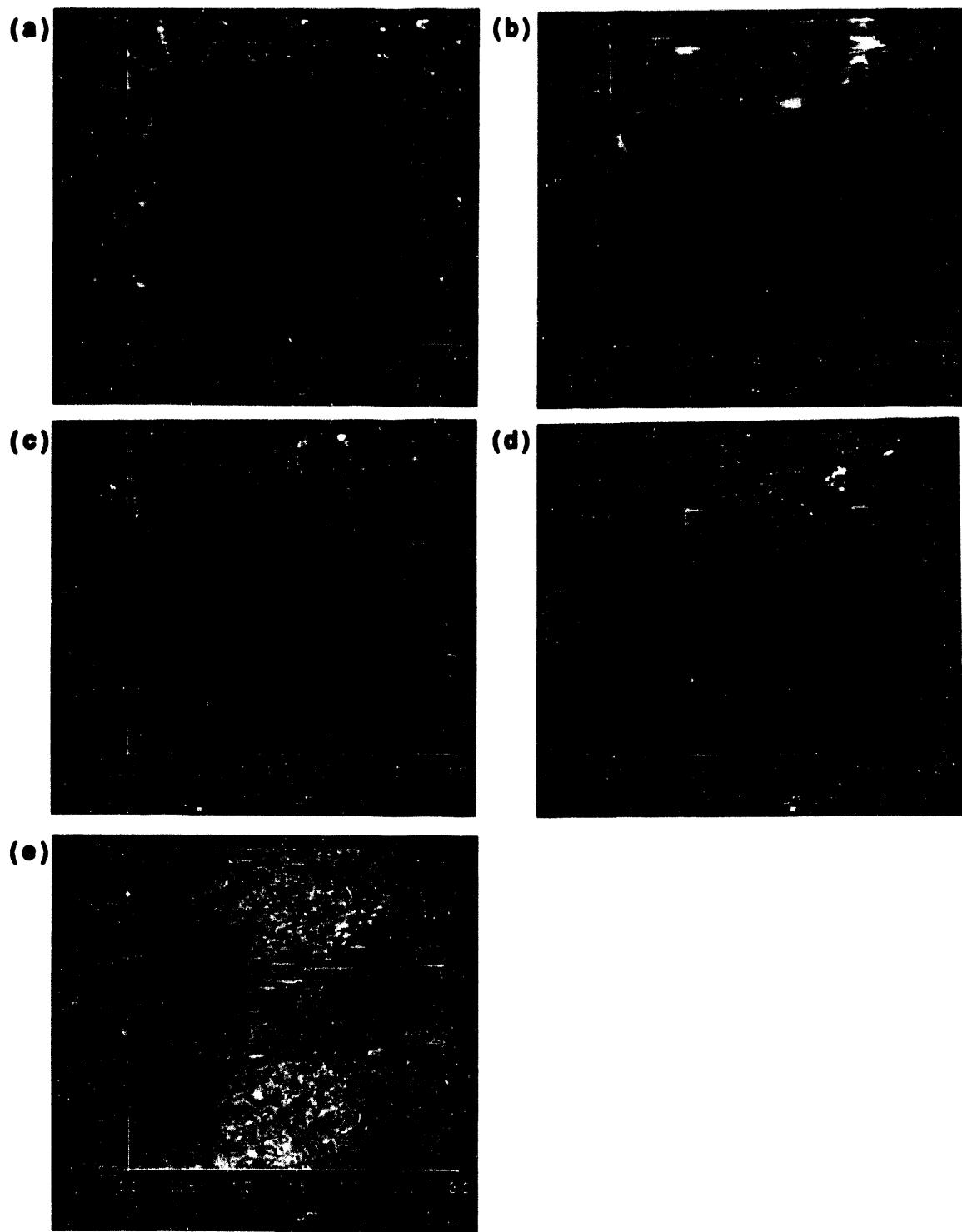


Figure 9-4 A second force series on collagen in isopentane at room temperature. The same area scanned at: (a) 0.5 nN, (b) 32 nN, (c) 0.5 nN after the area had been scanned at 64 nN, note the damage. The scanned region is visible in a larger area low force scan [(d), arrows] because of the damage. After scanning the original area at 125 nN, the damage is severe (e).

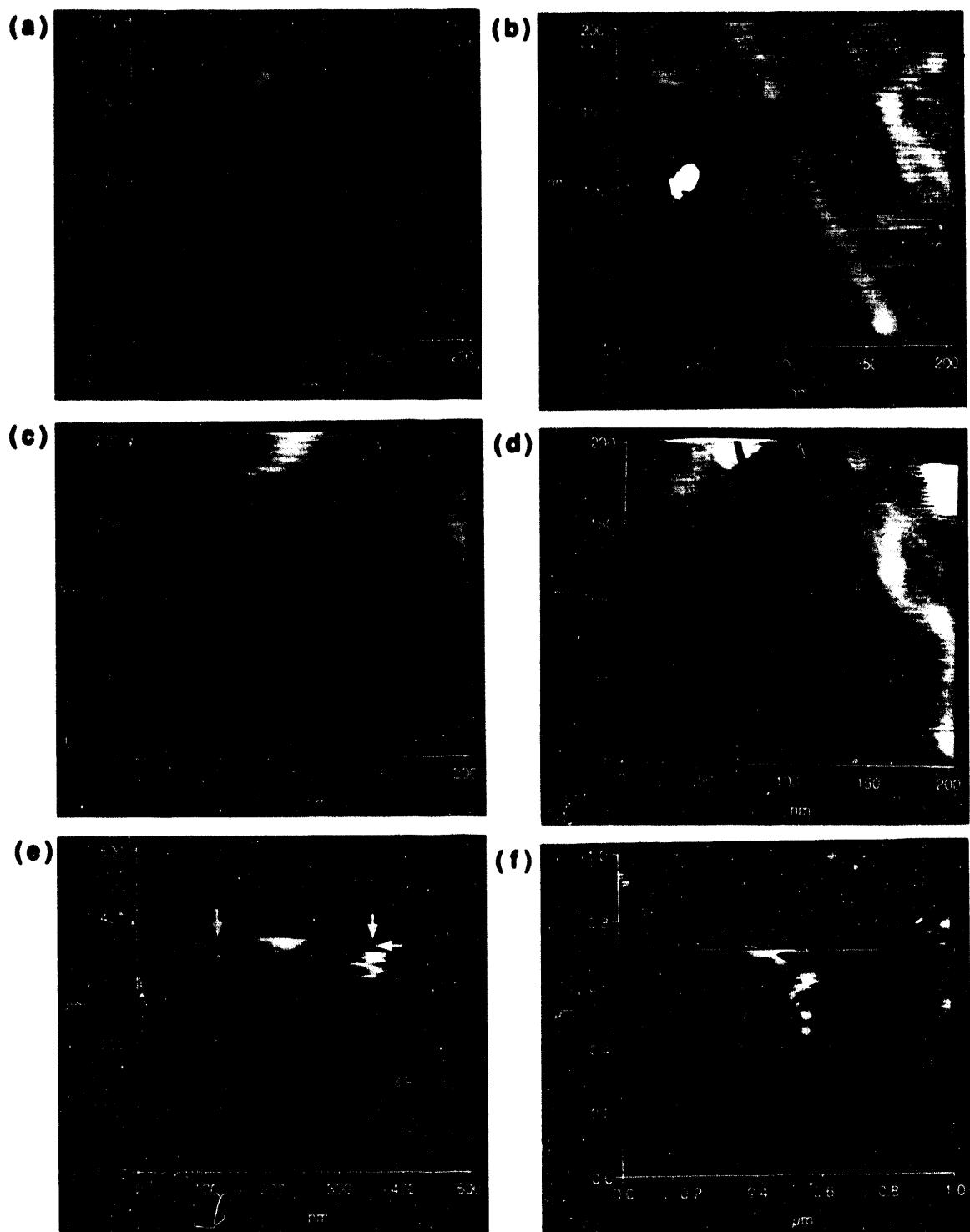


Figure 9-5 A force series on collagen at 143 K in isopentane. The same area scanned at: (a) 0.35 nN. (b) 22 nN, and (c) 44 nN (one molecule is being removed during this scan). Low force rescans confirm that the molecule was completely removed (d) and debris was moved to the edges of the scan (e). After a scan at 125 nN, all collagen is removed (f).

debris accumulation at the edges. After a scan at 170 nN, removal of the collagen is total [Fig. 9-5(f)].

We see that at both temperatures, the force needed to start breaking or removing molecules is on the order of 50 nN. If there is a difference between the temperatures, it is not dramatic.

A more direct comparison is done in Fig. 9-6 (note that the figure continues over two pages). Here the *same* sample has been imaged with the *same* tip at the two different temperatures, as part of a single variable-temperature run. The sample in this case is ferritin, prepared as described in the ferritin section (except using alcian blue instead of polylysine — this difference is not important here). The force is here not driven to the point of irreversible sample damage (to avoid damaging the tip, which would make the comparison biased), the focus instead being on the effect of increased force on image quality and reversible sample deformation. The left image of each pair is acquired at 143 K, the right image at room temperature. The force is increased from 0.18 nN in (a),(b) to 2.0 nN in (g),(h), then returned to 0.2 nN in (i),(j). As seen, the behavior is qualitatively similar at the two temperatures.³

[A quantitative way to assess the amount of sample deformation under the different conditions of temperature and scan force would be to measure the apparent height of the molecules in the images. Unfortunately, this run was blighted by tip-sticking artifacts (appearing as bright "flags" stretching from the top of the molecules), probably indicative of a contaminated tip, and these artifacts make it impossible to make an accurate measurement of

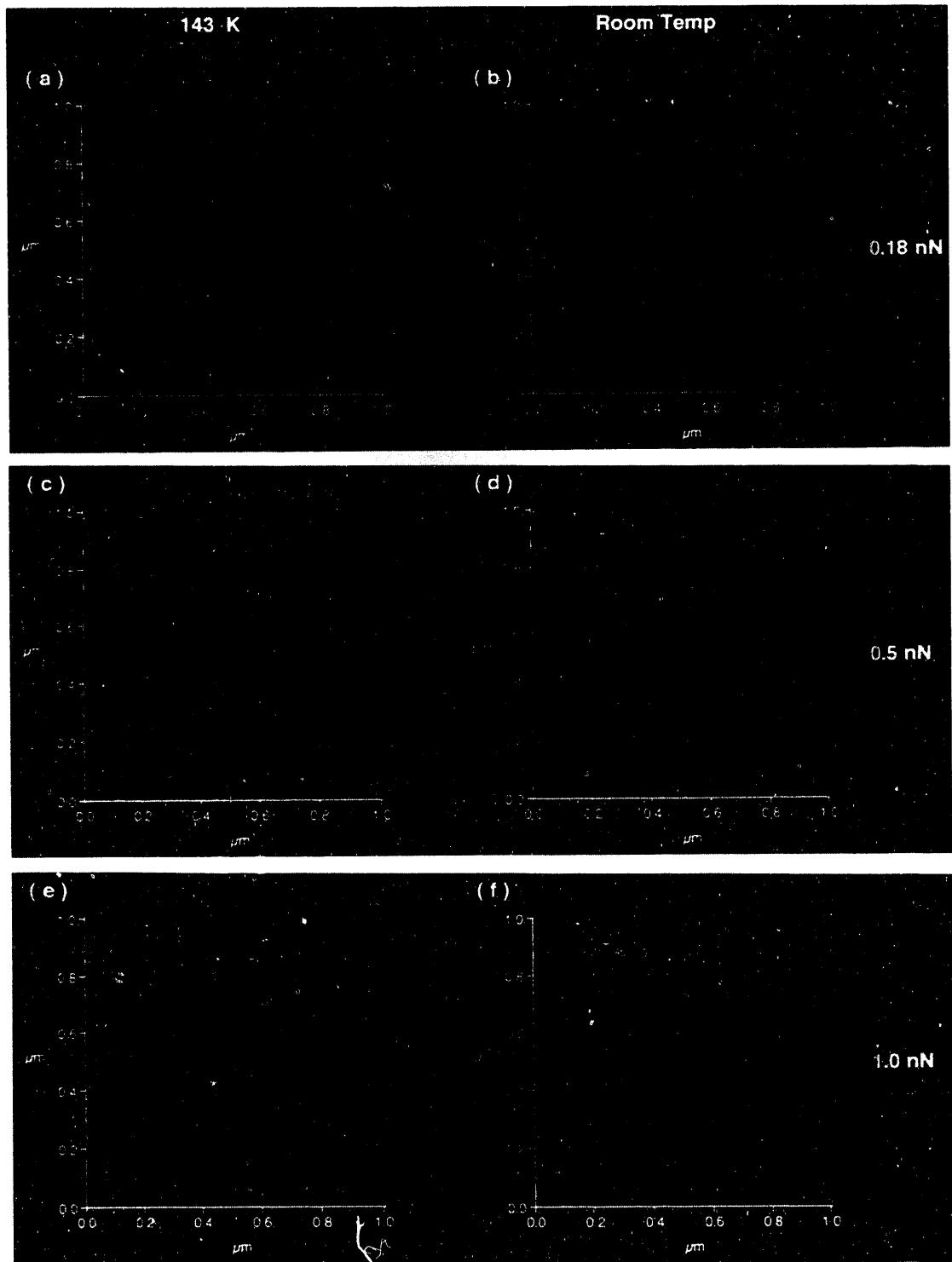
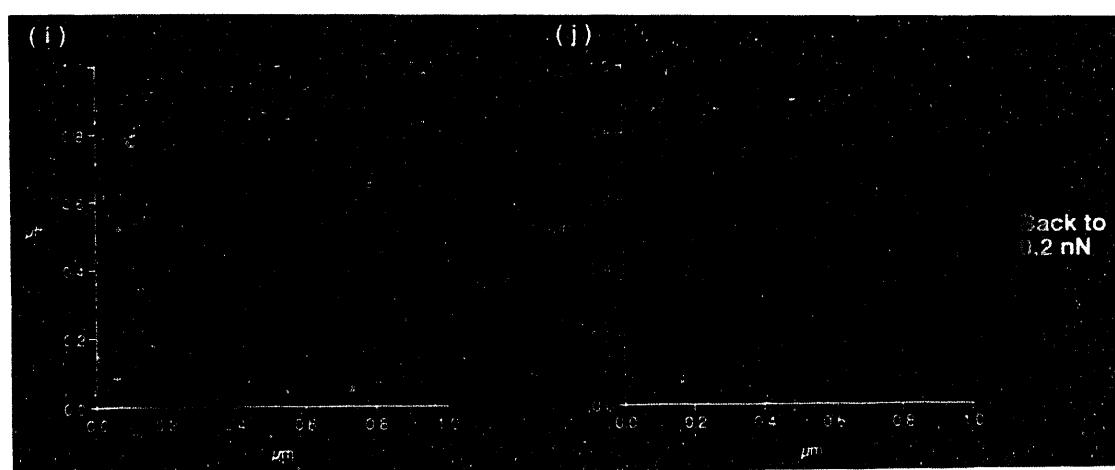
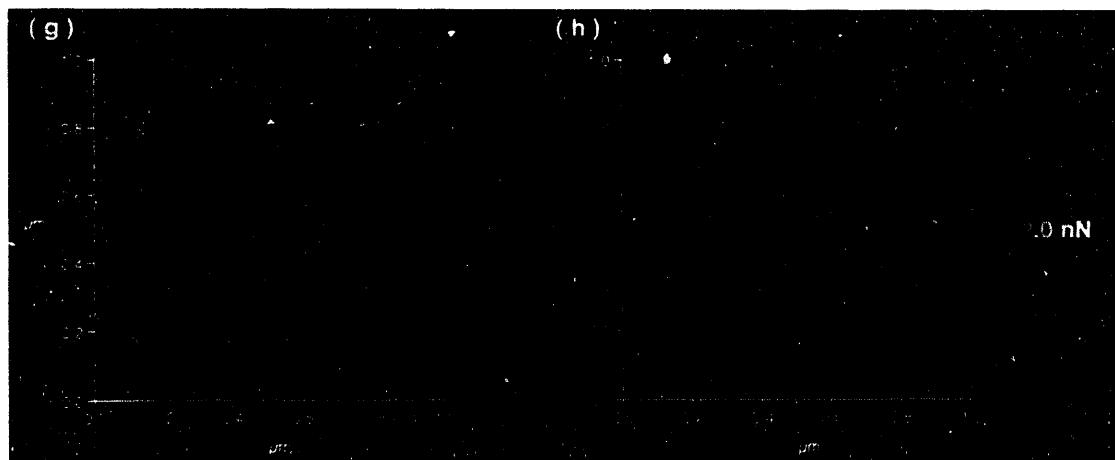


Figure 9-6. (This figure continues on the following page.) Parallel force series on ferritin in isopentane at 143 K and at room temperature, on the same sample using the same tip.



(Fig. 9-6. continued from the previous page.)

the height of many molecules. One might restrict the measurement to those molecules which do not show such artifacts, but this may be a biased subset, since the original population of apparent heights is nonuniform. Therefore we have not attempted such a comparison.]

In summary, none of our different measurements have established a clear difference in the behavior of the biological samples at 143 K compared to room temperature. Indeed, the results are consistent with there being no difference whatsoever.

This raises two questions, the first being "Does this mean the end of cold SFM for biology?" Not necessarily. Our conclusion was that the resolution is tip-limited at both temperatures, with the tips at our disposal and at a scan force of 0.5 nN or higher. It is likely that as tip technology is improved, a deformation-limited regime will soon be reached. Only in such a regime can one judge to what extent the ultimate resolution is affected by the temperature.

There is also the possibility that a temperature of 143 K is simply not low enough to improve sample rigidity significantly, but that such an improvement does take place at lower temperatures. One certainly should not attempt to pass judgement on low temperature SFM as a concept before it has been tried in the liquid helium temperature range. As I mentioned in the introduction, superfluid helium in particular is an attractive cryogen for SFM purposes since it does not bubble. Furthermore, an SFM operating at helium temperatures would benefit from a great reduction in thermal drift as well as of the thermally induced vibrations of the SFM spring. Under such conditions, it may be possible to operate at

lower forces, which would contribute to reducing sample deformation.

The SFM design described in chapter VI could be used in liquid helium as it is; the only necessary modifications would be to the sample manipulation and transfer tools.

Independently of the question of resolution, the advent of low temperature SFM opens to study a new class of experiments which by their nature can only be performed at low temperature. Examples of this will be discussed in chapter XI.

The second question is: "Is this result in conflict with the biological measurements on which we based our predictions?" It is not necessarily in conflict with the measurements themselves, only with our initial interpretation that these results were relevant to the question of deformability in SFM samples. Actually, the systems addressed by each of those experiments are all quite different from the situation encountered in SFM, namely a single, isolated molecule subjected to a force. In particular, they all address only thermal motion, not deformability. Furthermore, the molecules were studied not in isolation, but in intimate contact with an extremely viscous solvent (in the IR absorption measurements), and/or with each other in a crystal (in the X-ray crystallography), both environments where intermolecular interactions are likely to be as important as intramolecular ones.⁴

References:

- ¹Grown by P. E. Russell, whose help is gratefully acknowledged.
- ²Ultralevers from Park Scientific Instruments, 1171 Borregas Ave., Sunnyvale, CA 94089, USA.
- ³Note, however, that this sample was imaged at room temperature first, then cooled slowly, instead of being plunged into the cold liquid, which is our usual procedure for low temperature imaging.
- ⁴A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter J, and W. A. Eaton, *Science* **256**,1796 (1992).

Chapter X: Sample Modification

In previous chapters we have encountered the fact that the forces applied by the tip can alter the sample. There we referred to this as "sample damage," a nuisance to be avoided. In this chapter we take the opposite attitude: the ability of the SFM to apply mechanical force selectively to nanometer-scale areas makes it a unique tool with which to manipulate the sample deliberately. As we will see, it is even capable of controlled surgery on a single molecule at a time.

Removal of Single Molecules

The first example of sample alteration (Fig. 10-1) belongs to the non-deliberate category: a severe tip crash during walkup (with the early, inertial walker system, which was quite unreliable at this time), on a ferritin sample¹ at 143 K, has plowed a "trench" by removing ferritin molecules. The detached molecules are the cause of the streaks and glitches in the image — this streakiness abates in successive images as the molecules disperse in the liquid.

Figure 10-2 shows an attempt to do the same thing in a more controlled fashion. After Fig. 10-2(a) [which is recognized as a subset of image 8-8(a), from the ferritin section] was acquired, the tip force was briefly increased to about 100 nN as it was scanned back and forth along the bottom of the image (gray arrows). The result [Fig. 10-2(b)] was, in addition to clearing the area directly under the path of tip, to remove cleanly several ferritin molecules outside this path, exposing previously hidden molecules to view (white arrow). I want to emphasize that these exposed sites, as

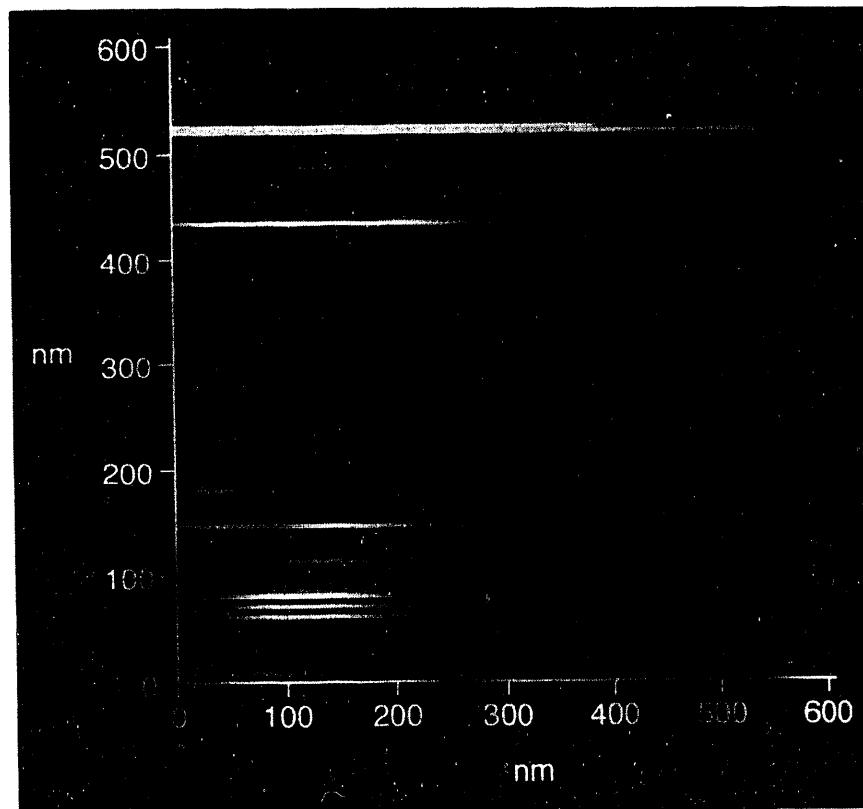


Figure 10-1 Unintentional removal of ferritin molecules by a tip crash. 143K in n-pentane.

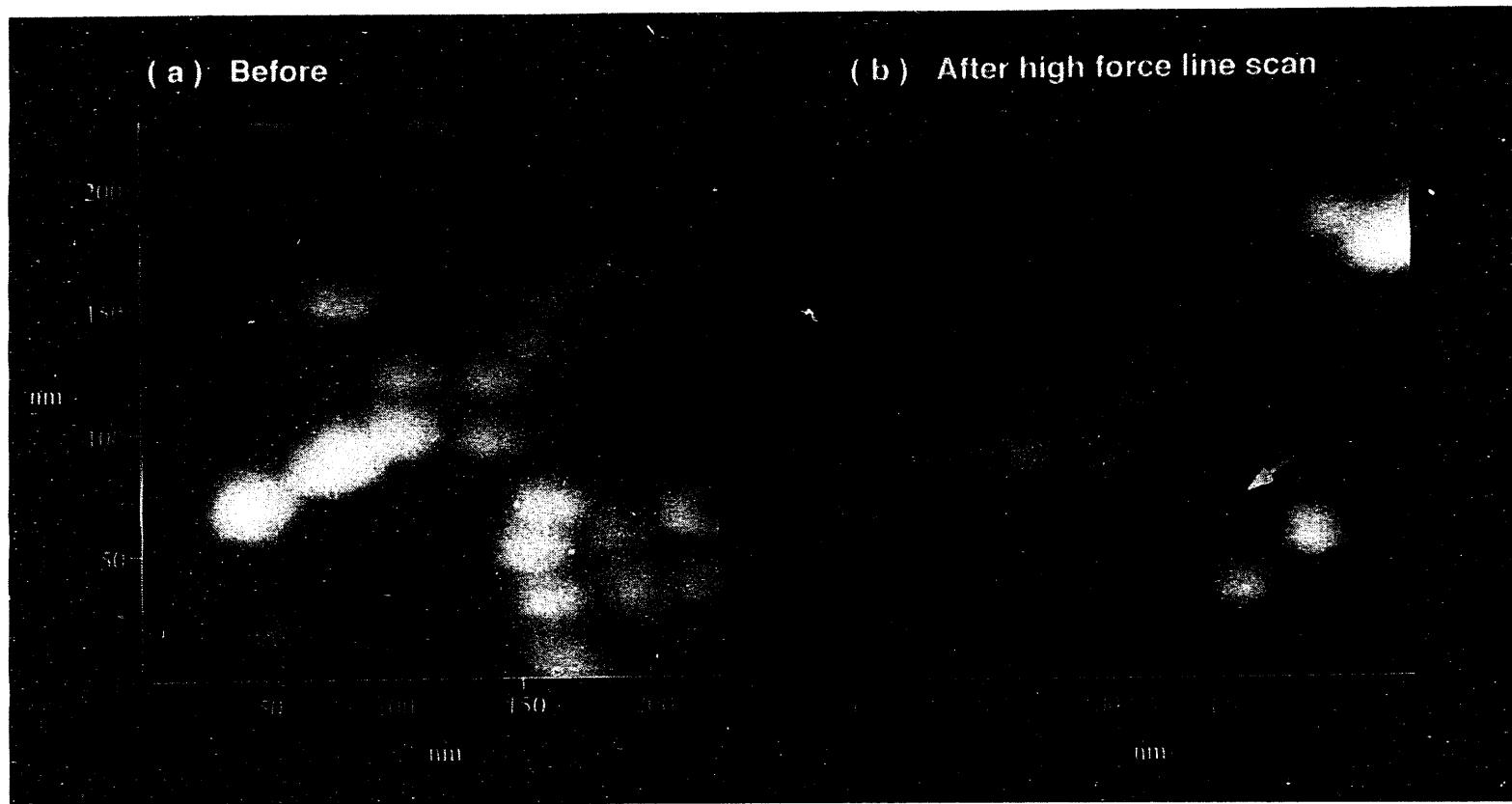


Figure 10-2. Intentional removal of ferritin molecules (at 143K in n-pentane). Between the images the tip was scanned at 100 nN along the bottom of the image (gray arrows). Some molecules were removed outside of the scan line, exposing previously hidden molecules which have not been exposed to the high force (white arrow).

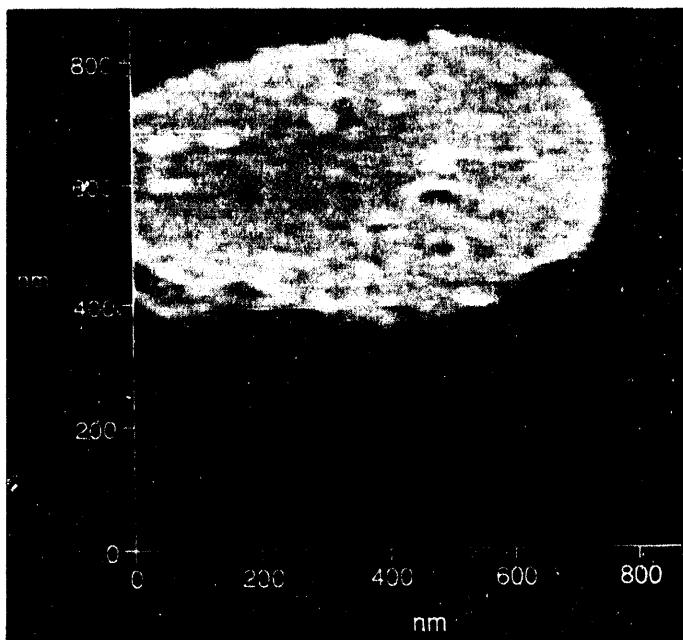
opposed to those in the tip path, have *not* been subjected to the high force scan, with any artifacts that it may produce. One can easily imagine the potential of such selective removal of single molecules if the technique could be applied to, say, neurotransmitter molecules, which could be imaged, then removed to expose their binding site.

Splitting / Removal of Membranes

Cell membranes are natural candidates for mechanical alteration by the tip, since they consist of a lipid bilayer, the two sheets of which are weakly bound and could be separated. Any cell membrane contains and/or is pierced by membrane proteins, and removal of the upper lipid layer could expose previously near-hidden proteins. Removal of the proteins themselves would leave a set of holes which would contain information about the arrangement of the proteins, possibly in a clearer form than was visible when the proteins were present.

We have attempted to demonstrate the feasibility of such cleavage, using purple membrane as the sample. With the SFM operating in n-pentane at 143 K, we scanned the tip at a force of about 100nN over a region of the PM patch shown in Fig. 10-3 (a). As we see in Fig. 10-3(b), this resulted in the removal of part of the membrane. The area exposed by this is still higher than the background (by about 0.8 nm), implying that either the lower lipid layer, or possibly a layer of water ice which had been trapped under it (and which would presumably constitute a cast of the bottom of the membrane) remains. Unfortunately, the tip was damaged and/or contaminated during the high force scan, so as to diminish the

(a)



(b)

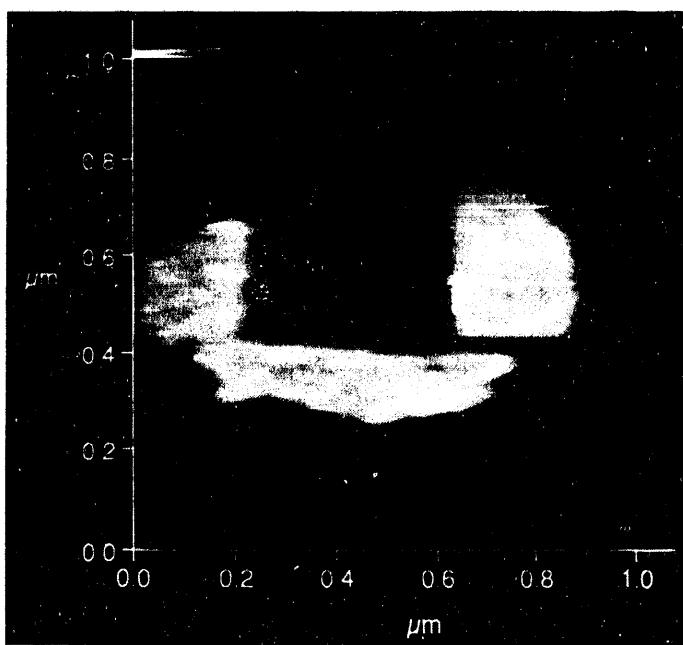


Figure 10-3 PM at 143 K in n-pentane. A section of the membrane shown in (a) was scanned at about 100 nN. as seen in (b), this removed all or part of the membrane.

resolution to the point where we could not judge whether any trace of the protein lattice was present in this exposed layer.

In the procedure in the previous paragraph, the area exposed by the removal of the membrane was itself subjected to the high tip force, and was thus at risk of being damaged. Other structures can be removed with much weaker forces. Figure 10-4 shows an object which could be fractured and partly removed by a force of only 3 nN, 40 times lower than that used to remove the PM. The nature of this object is unknown —it occurred as a contaminant in a DNA sample on silanated mica, similar to the RNA samples described in chapter VIII. Its thickness was about 5 nm; possibly it was a lipid membrane. The first image, Fig. 10-4 (a), was taken in the error signal mode, at a force of about 1 nN. The same area was then rescanned in the open loop mode [Fig. 10-4 (b)], at a force which increased from 0.6 to 5 nN from the bottom to the top of the image (due to thermal drift — the temperature was still settling toward the equilibrium at this point). This second scan damaged the object in two ways: a crisp brittle crack appeared (in the lower part of the image), and the tip appears to have cut through the object to the underlying substrate during some scans. Note that during these scans, the tip is tracking well and is imaging the texture of the substrate correctly — it appears that the object could be sliced away with little effort. During subsequent scans, the entire part of the object beyond the crack was successively removed [Fig. 10-4 (c), (d)]. The underlying area thus exposed (Fig. 10-5) has a distinctly different, more fine-grained texture than the silane background, but its height is very similar. It

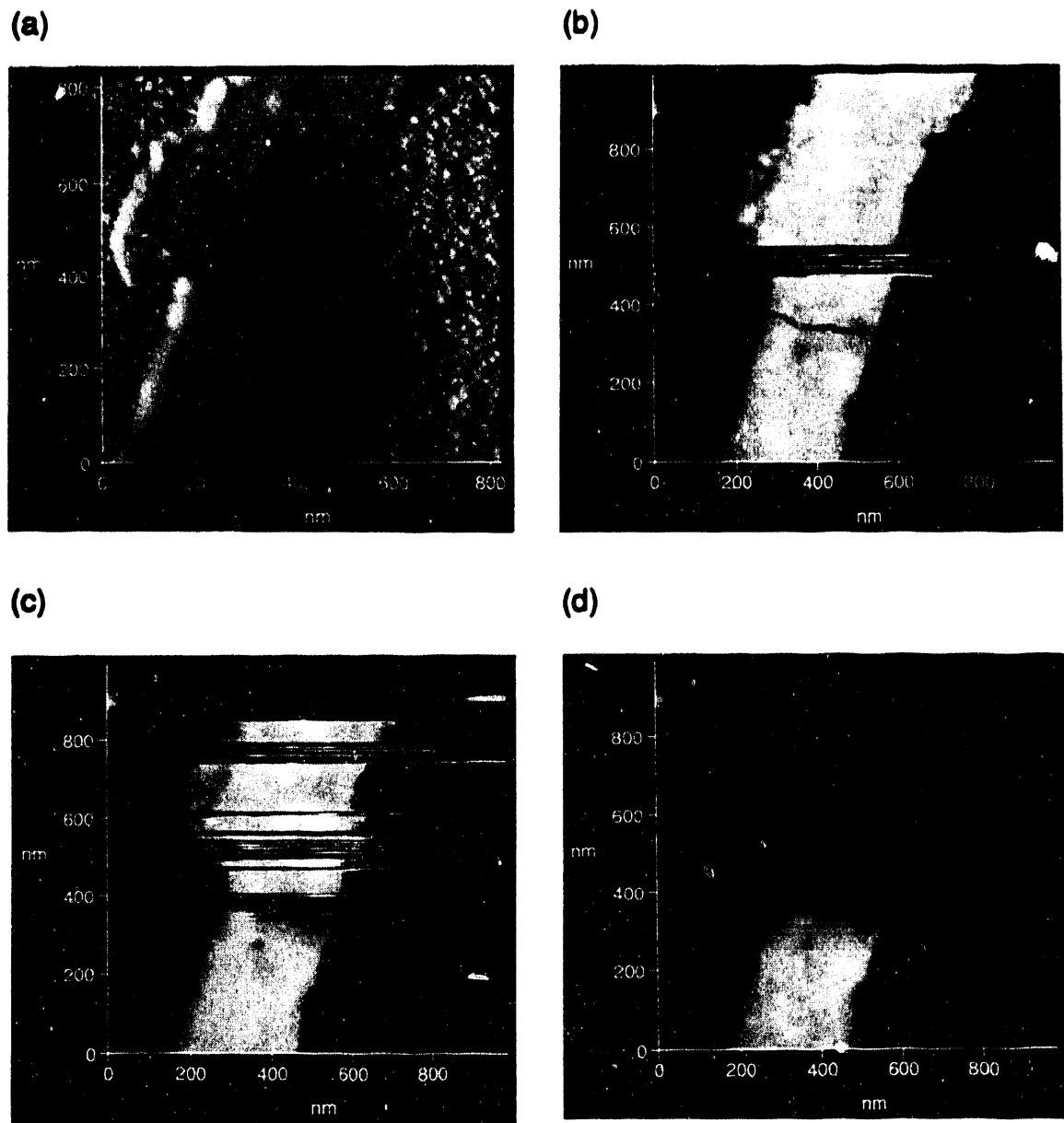


Figure 10-4 An object which could be nanomachined at a low scan force. This object was a contaminant in a DNA sample on silanated mica, and its nature is unknown. (a) "Before" image (in the error signal mode). (b)-(d) Rescans in the open loop mode. The force is increasing from 0.6 nN at the bottom of the images to 3-5nN at the top due to drift.

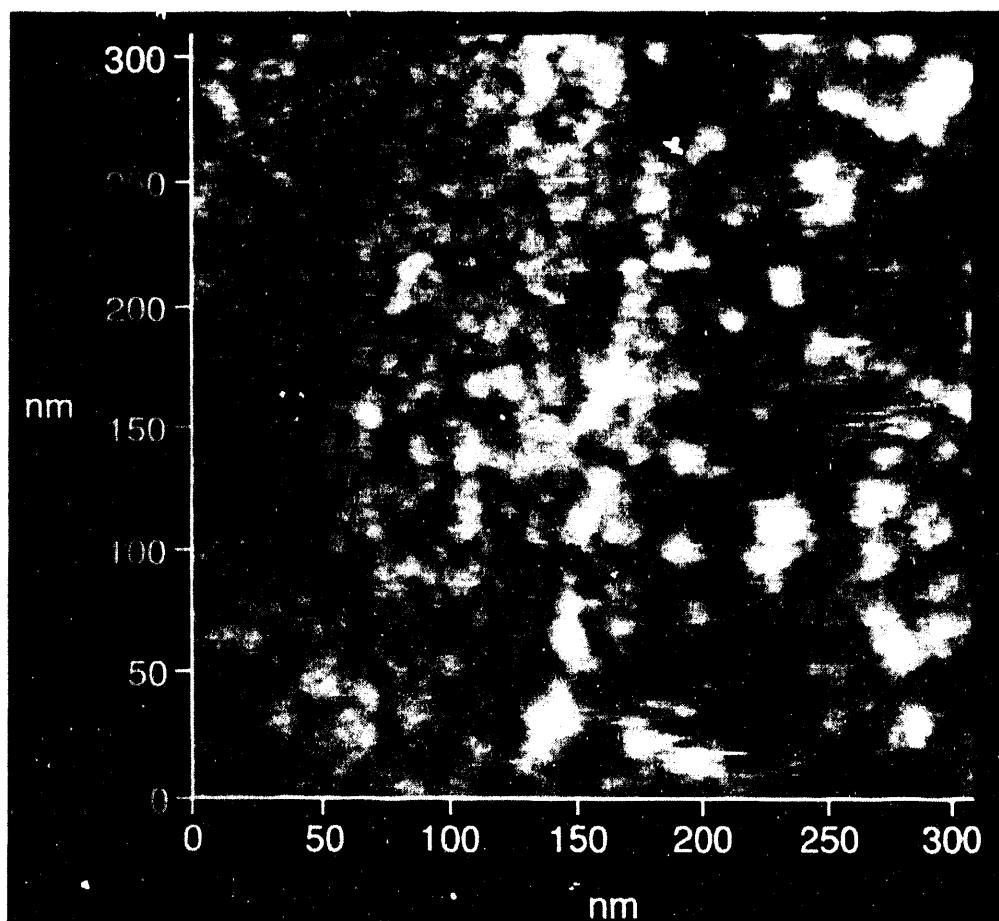


Figure 10-5 A closeup of the edge between the area exposed by the operation in the previous figure (left half of the image) and the background (right half). The difference in texture is distinct.

may be simply the silane layer, but protected from dehydration effects and contamination by the now removed upper layer. Note how the raised humps in Fig. 10-4 (a) were caused by covered-over objects which are exposed in Fig. 10-4 (d). Also note that the tip resolution is of unchanged (and comparatively high) quality after the removal operation.

Surgery on Single Molecules

So far we have described how the SFM can be used to remove molecules or larger parts of the sample. Its abilities extend further than this: it can actually section a single molecule at a desired location. Figure 10-6 shows an example using type I collagen molecule. After Fig. 10-6 (a) was acquired, the tip was placed to scan back and forth at the level shown by the arrow in Fig. 10-6 (b), and the tip force was briefly increased to about 100 nN. As seen in Fig. 10-6 (b), this severed the collagen molecule (all three strands). To demonstrate the reproducibility with which this kind of molecular surgery can be done, the procedure was repeated at a second place [Fig. 10-6 (c)], with similar result. Remarkably, the binding that attaches the molecule to the substrate is strong enough that all parts of the molecule that are 15 nm or more away from the cut appear entirely unaffected, while covalent bonds are being severed a few nm away! The bright, streaky areas near the cut are caused by the "frayed ends" of the molecule which are no longer tightly bound to the substrate, but are "swaying in the breeze." The apparent width of the gap in the molecule being about 13 nm, we estimate that the position of the cut can be chosen with a precision of 10 nm or better.

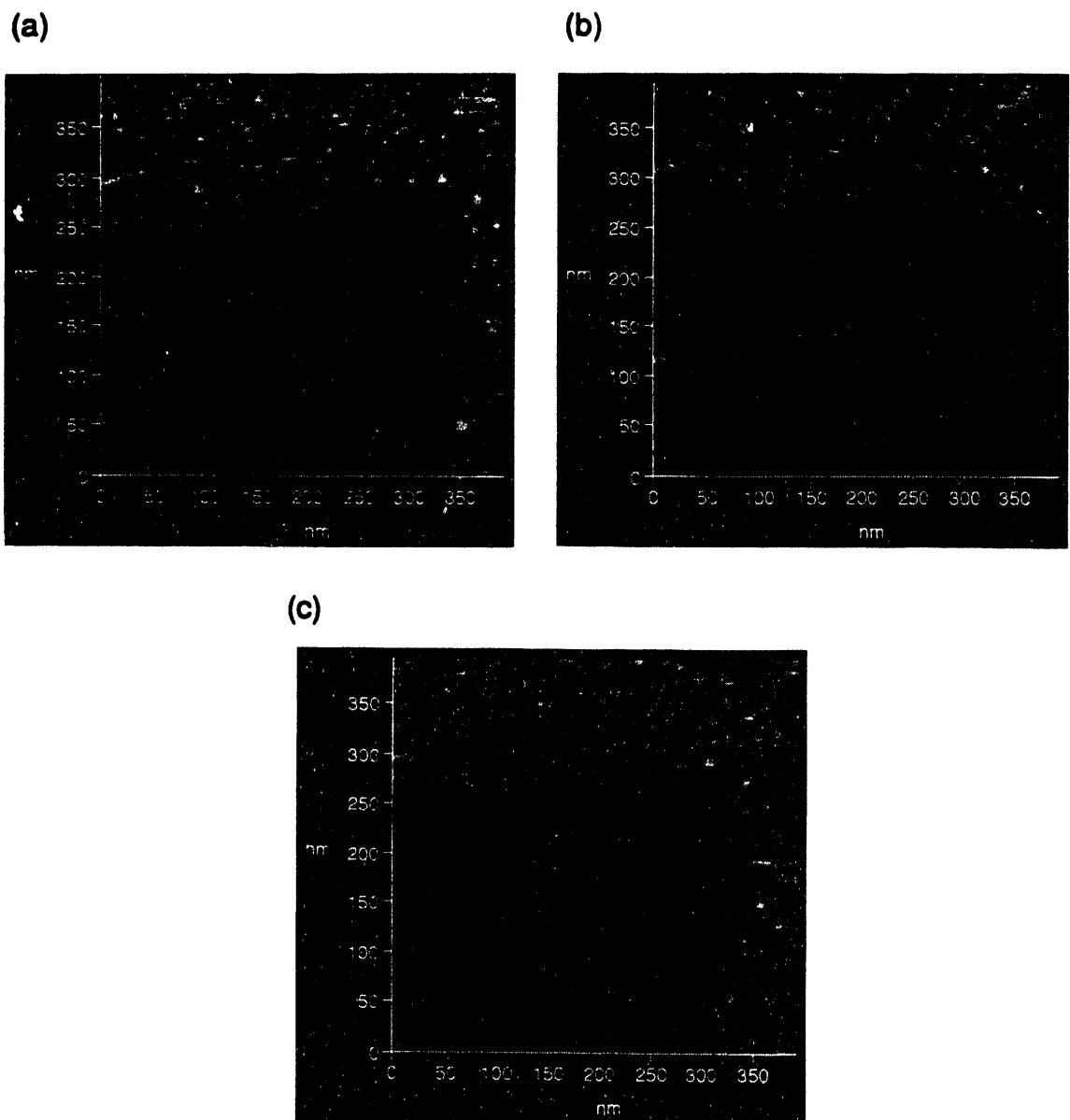
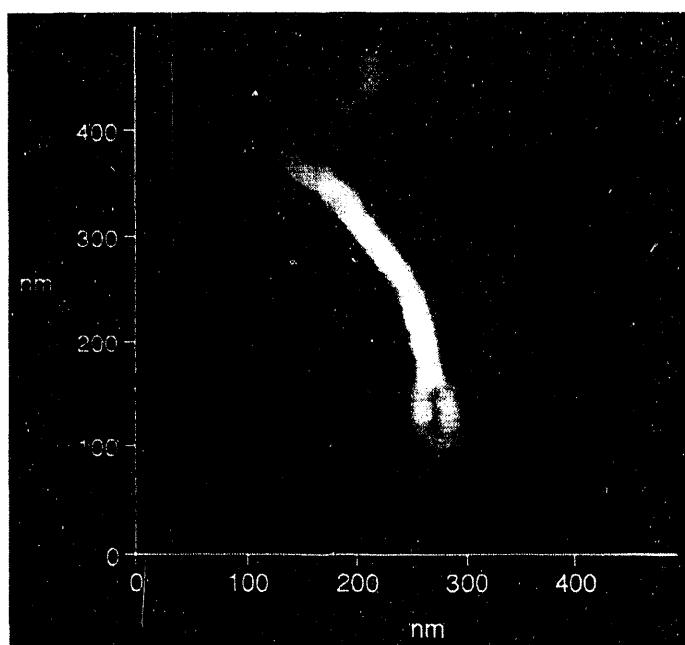


Figure 10-6 Surgery on a single molecule. The collagen molecule in (a) was cut into two (b) and then into three (c) pieces by scanning the tip across it at 100 nN (arrows). (At 143 K in isopentane.)

This single-molecule surgery is not limited to collagen. Figure 10-7 shows the same process applied to a molecule of double-stranded RNA — actually to a molecule which had twisted back upon itself to form a 4-stranded structure. Apparently, all four molecular strands were severed simultaneously, since the apparent height in the cut has decreased back to substrate level.

While the ability to segment a single molecule admittedly is an attribute unique to the SFM, one may ask of what use a single custom sliced molecule could possibly be? Actually, it is not difficult to envision scenarios where it may be extremely useful, particularly scenarios involving DNA or RNA, because of the ability of the polymerase chain reaction (PCR) to amplify a single nucleic acid fragment to an arbitrarily large quantity of identical copies. Consider, for an example, a DNA defect that causes a "kink", "bump" or other recognizable change in the shape of the DNA double spiral². A preparation of DNA known or suspected to contain such defects could be scanned with an SFM, which would be able to spot such defects if any were present. Once identified, a molecular segment containing the defect could be excised with the tip. If the DNA was then redissolved off the substrate and separated by size using standard molecular biology methods, the excised fragment could be isolated from all other DNA (assuming that the original preparation contained no fragments this small). It could then be amplified, using PCR, and sequenced, and the relation between the sequence and the defect could be elucidated. In the same way one could add a DNA-binding protein to a DNA preparation, and locate, excise, amplify and sequence the particular regions of DNA to which it bound.

(a)



(b)

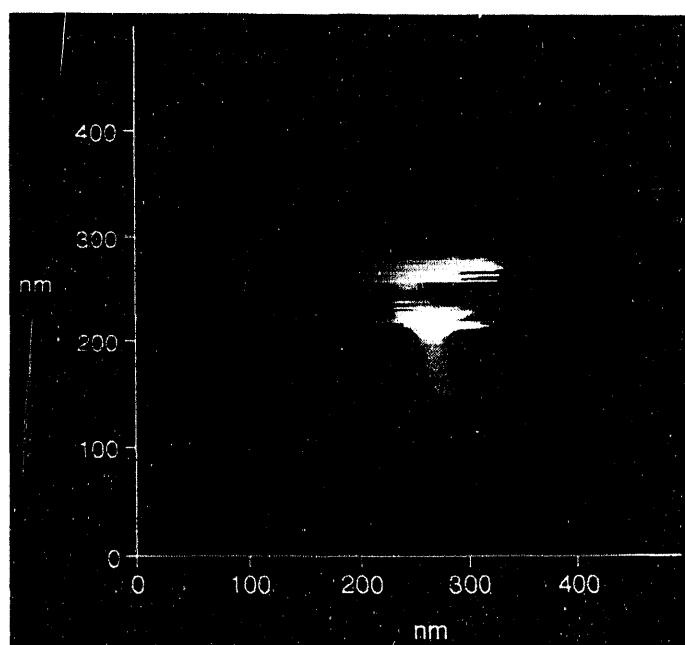


Figure 10-7 (a) A 4-stranded RNA structure (b) All four strands could be severed by a line scan at 100 nN. (At 143 K in isopentane.)

References:

- 1**All images in this section are acquired at 143 K unless noted otherwise.
- 2**H. Htun and J. E. Dahlberg, *Science* **241**, 1791 (1988).

Chapter XI: Future Possibilities

In this chapter, I will discuss applications which are only possible with a low temperature SFM, not because of any resolution improvement, but by their nature.

Freeze Fracture

Even if it were to turn out that low temperature operation of an SFM confers only a negligible resolution advantage, or none at all, it will still have opened to SFM study a whole new class of samples, including those prepared by rapid-freezing and freeze fracture. As discussed in earlier chapters, these methods allow one to lock a sample into a fixed configuration, which, if the freezing is rapid enough, will be very close to the native one. Freeze-fracture allows one, for instance, to image the interior of cells that have been broken open, and which would disintegrate if anything similar were to be attempted at room temperature.

For this reason it was one of our major objectives, in designing our low temperature SFM, to make it possible to study freeze fractured samples. The SFM system, described in chapter V, does allow us to transfer frozen samples via the airlock into the chamber (and into the microscope well) without contaminating the atmosphere in the dry chamber. Unfortunately, our first procedure for sample preparation did not achieve a fast enough freezing rate, and we had to abandon the effort for lack of time before having developed a successful preparation protocol. Since many parts of the procedure do work well, I will describe it here: The sample, red blood cell ghost vesicles, was applied to a polylysine coated cover

glass (which was already epoxied to one of our sample carriers) and rinsed with distilled water. With a drop of water remaining, the sample was covered with a thin (0.1 mm) copper "hat"¹ so that only a thin water film remained. The whole sandwich was plunged into isopentane at its melting point (113 K), then moved into liquid nitrogen while still covered with isopentane. The sample carrier was put into the transfer airlock in a small glass dewar, still under liquid nitrogen. There the nitrogen was pumped off, together with the air, and the transfer chamber backfilled with dry nitrogen gas. An Al block in the dewar supplied heat capacity to evaporate the nitrogen, and a Cu transfer boat covered the sample to shield it from radiant heating and contamination. Note that the sample surface to be imaged was at this time imbedded in ice, and thus not itself subject to contamination. The airlock was opened to the main chamber, and the sample is transferred into the microscope well. There, protected from frost by the immersion liquid, the Cu hat was cleaved off, using tools from above, at which point the sample was ready to be scanned. For freeze-fractured samples, we use an immersion liquid consisting of 85% isopentane and 15% n-propanol by volume. We have found this mixture to posses a small but non-zero electrical conductivity, which allows the triboelectric charges generated by the fracturing to dissipate from the sample surface, eliminating unwanted electrostatic tip-sample attraction.

The transfer procedure does work, and the samples can be imaged. The appearance of such a sample is shown in Fig. 11-1. The mountainous landscape represents micron-sized ice crystals, indicating that the freezing rate was insufficient. It is possible that

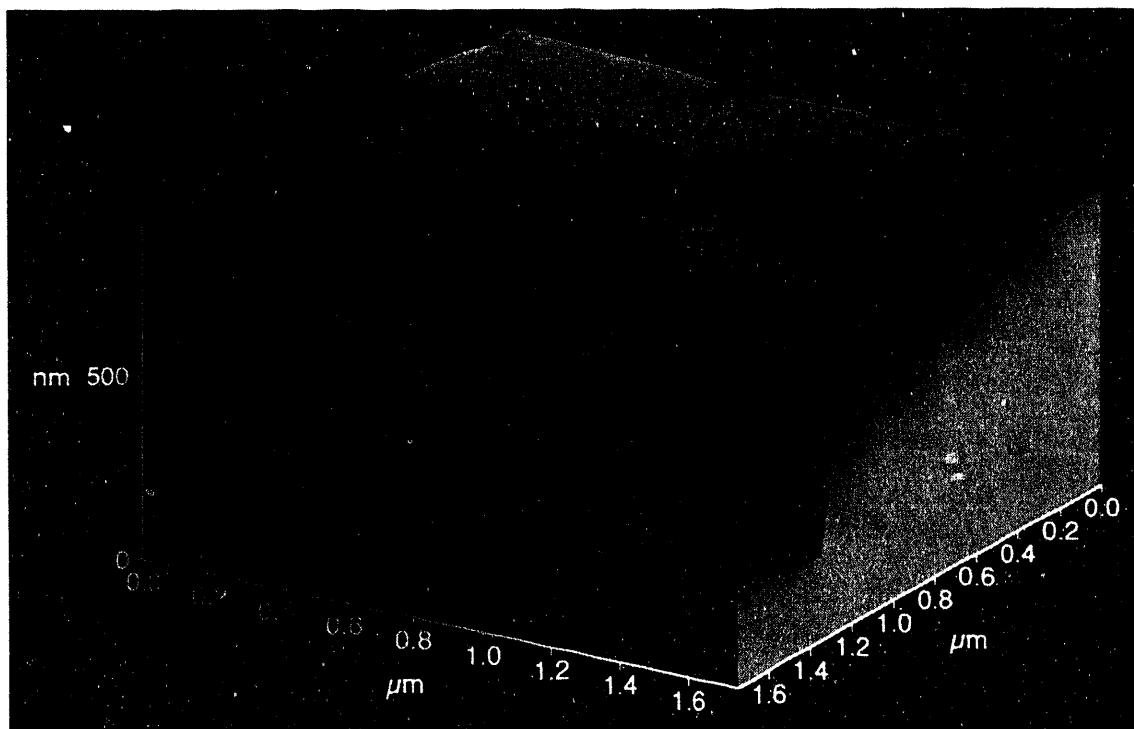


Figure 11-1. A freeze-fractured sample imaged at 143 K. The mountainous structures are micron-size ice crystals. The existence of these indicates that the freezing rate in our sample preparation procedure was insufficient. (Note that the z scale in this image, unlike all other images in the thesis, has not been enhanced relative to the xy scale. The steep hole in the foreground is an artifact caused by the sample actually exceeding the z range of the scanner.)

a fairly minor modification to the procedure up until the freezing step could remedy this. For instance, the sample could be prepared between two copper hats, eliminating the heat capacity of the carrier. (Our theory, when designing the first procedure, was that the low thermal conductivity of the cover glass would make its thermal contact with the carrier small compared to the thermal contact to the bath through the copper hat.) The plunging into isopentane can also be replaced by an even more effective cooling method, such as subjecting the copper hat sandwich to opposing jets of propane at its melting point.² We have manufactured sample carriers with the same outer dimensions as those described in chapter VI, in which such sandwiches can be mounted and fractured.

Elasticity Measurements

In addition to being an imaging tool and a sample manipulation tool, The SFM has a third and unique ability: it can be used to measure the mechanical properties^{3,4} of the sample on the nm size scale, even on single molecules. The low temperature SFM adds the opportunity to study these properties as a function of temperature over a wide temperature range. The stiffness of a particular deformation mode of a biomolecule depends on the strength of the interactions which stabilize the molecule in its conformation, and the temperature dependence of the stiffness contains information about the energetics of these interactions.

References:

- 1 K. Mühlenthaler, E. Wehrli, and H. Moor, Proc. 7th Intl. Congr. Electr. Microsc., Grenoble, 1970, p. 449.
- 2 A. W. Robards and U. B. Sleytr, "Low Temperature Methods in Biological Electron Microscopy," Elsevier, Amsterdam, 1985.
- 3 P. Maivald, H. J. Butt, S. A. C. Gould, C. B. Prater, B. Drake, J. A. Gurley, V. B. Elings, and P. K. Hansma, Nanotechnology **2**, 103 (1991).
- 4 N. J. Tao, S. M. Lindsay, and S. Lees, Biophys. J. **63**, 1165 (1992).

Chapter XII: Summary

In summary, I have described (in addition to two improvements to the technology of SFM springs) a force microscope capable of operating at 143 K under liquids, and have presented images of biological samples acquired with this system both at 143K and at room temperature. We have not observed any dramatic resolution advantage from low temperature operation, though such an advantage may become apparent as sharper tips are made available. Regardless of the question of resolution, the low temperature SFM allows the study of new categories of samples, such as those prepared by freeze-fracture.

Further, we have demonstrated that the SFM can be used to manipulate samples on a nm scale, and in particular to perform surgery on single molecules of proteins and nucleic acids with a precision of 10 nm.

The low temperature SFM has the potential of measuring the mechanical properties of single molecules, and the temperature dependence of these properties. In the future, it may be possible to use such measurements to gain new insights into the energetics of molecular stability and binding.

**DATE
FILMED**

6/3/94

END

