SPECIAL TOPIC: NEAR-FIELD MICROSCOPY AND SPECTROSCOPY

Scanning near-field optical microscopy with aperture probes: Fundamentals and applications

Bert Hecht,[a] Beate Sick, and Urs P. Wild
Physical Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Z, Universitätstr. 22, CH-8092 Zurich, Switzerland

Volker Deckert and Renato Zenobi
Organic Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Z, Universitätstr. 16, CH-8092 Zurich, Switzerland

Olivier J. F. Martin
Electromagnetic Fields and Microwave Electronics Laboratory, Swiss Federal Institute of Technology, Gloriastrasse 35, ETZ ETH-Zentrum, CH-8092 Zurich, Switzerland

Dieter W. Pohl
Institute of Physics, University of Basel, Klingelbergstr. 82, CH-4056 Basel, Switzerland

(Received 13 October 1999; accepted 1 February 2000)

In this review we describe fundamentals of scanning near-field optical microscopy with aperture probes. After the discussion of instrumentation and probe fabrication, aspects of light propagation in metal-coated, tapered optical fibers are considered. This includes transmission properties and field distributions in the vicinity of subwavelength apertures. Furthermore, the near-field optical image formation mechanism is analyzed with special emphasis on potential sources of artifacts. To underline the prospects of the technique, selected applications including amplitude and phase contrast imaging, fluorescence imaging, and Raman spectroscopy, as well as near-field optical desorption, are presented. These examples demonstrate that scanning near-field optical microscopy is no longer an exotic method but has matured into a valuable tool. © *2000 American Institute of Physics.* [S0021-9606(00)70316-3]

I. INTRODUCTION

Scanning near-field optical microscopy (SNOM or NSOM) is at the forefront of science and technology today because it combines the potentials of scanned probe technology with the power of optical microscopy. SNOM provides us with *eyes* for the nanoworld. Among the main parameters that might be of interest for a nano structure under investigation are, besides shape and size, its chemical composition, molecular structure, as well as its dynamic properties. In order to investigate such properties, microscopes with high spatial resolution as well as high spectral and temporal resolving power are required. The classical optical microscope excels with respect to spectroscopic and temporal selectivity, although its resolution is restricted by diffraction to about half the wavelength, i.e., to 0.2–0.5 micrometer for visible light. Present-day science and technology, however, have an increasing need for tools that allow to characterize, generate, and manipulate structures as small as a few nanometers in size. Examples are readily found in biology, microelectronics, and medical sciences. Electron microscopes as well as scanning tunneling and atomic force microscopes easily achieve 10 nm spatial resolution and beyond, but they are relatively poor performers with respect to spectral and dynamic properties. Electron microscopes have to be operated in vacuum, which limits their application in life sciences, requires special sample preparation, and complicates sample manipulation. SNOM combines the excellent spectroscopic and temporal selectivity of classical optical microscopy with a lateral resolution reaching well into the sub-100 nm regime. However, for a long time technical problems jeopardized the widespread application of SNOM. Near-field optics (NFO) therefore became a focus of research and development in the field of optical microscopy in recent years.[1–9] Today, we have reached the point where SNOM represents a powerful tool for surface analysis that is technically and theoretically well understood. It is ready to be applied to a large variety of problems in physics, chemistry, and biology.

In a simplified view of classical far-field optical microscopy the object is illuminated by a monochromatic plane wave. The transmitted or reflected light, scattered by the object in a characteristic way, is collected by a lens and imaged onto a detector. For practical reasons, the lens is placed at least several wavelengths λ of the illuminating light away from the object surface, i.e., in the far field. Since high spatial frequencies corresponding to the fine details of the object generate Fourier components of the light field that decay exponentially along the object normal,[10] they cannot be col-

[a]Electronic mail: hecht@phys.chem.ethz.ch

lected by the lens. This effect leads to the well known Abbé diffraction limit $\Delta x = \lambda/(2\pi \mathrm{NA})$,[11] where NA is the numerical aperture of the lens.

This resolution limit can be slightly pushed by scanning confocal optical microscopy where a sharply focused spot of light replaces the wide-field illumination.[12,13] Using special geometries of illumination and detection, sometimes in combination with nonlinear effects such as multiphoton excitation,[13] further progress in this field is still being made; however, steps are small and no major breakthrough is to be expected.

In 1928, Synge, an Irish scientist, described an experimental scheme that would allow optical resolution to extend into the nanometer regime:[14] He proposed to use a strong light source behind a thin, opaque metal film with a 100 nm diameter hole in it as a very small light source. The tiny spot of light created this way should be used to locally illuminate a thin biological section. In order to guarantee the local illumination, he imposed the condition that the aperture in the metal film be no further away from the section than the aperture diameter, i.e., less than 100 nm. Images were to be recorded point by point detecting the light transmitted by the biological section by means of a sensitive photo detector.

With his proposal, Synge was well ahead of his time. Unfortunately, he never tried to realize his idea. In 1932, he actually proposed an alternative, discarding his original scheme because of the difficult sample approach to a planar screen.[15] He therefore suggested to use the image of a point light source as an optical probe instead of an aperture. The image was to be generated by an ellipsoidal mirror which would provide, in modern words, the largest possible numerical aperture. Ironically, this second scheme never would have achieved SNOM-type resolution since it ignored the role of near fields which cannot be recovered by any conventional imaging scheme, no matter what the numerical aperture is.

In 1984, shortly after the invention of the scanning tunneling microscope,[16] nanometer-scale positioning technology was available and an optical microscope similar to Synge's proposed and forgotten scheme was re-invented by Pohl and demonstrated together with Denk and Duerig at the IBM Rüschlikon Research Laboratory.[17–19] Independently, a similar scheme was proposed and developed by Lewis and his group at Cornell University.[20–22] The key innovation was the fabrication of a subwavelength optical aperture at the apex of a sharply pointed transparent probe tip that was coated with a metal. In addition, a feedback loop was implemented maintaining a constant gapwidth of only a few nanometers while raster scanning the sample in close proximity to the fixed probe.

The promise to extend the power of optical microscopy beyond the diffraction limit triggered the development of several experimental configurations that are able to generate optical images with nanometer resolution. Figure 1(a) shows the classical aperture SNOM configuration in which an aperture probe is illuminating a small area of a sample surface.[4,5,17–23] In the most general configuration with respect to light detection, the sample is placed on top of a hemispherical substrate which allows the capture of all the radiation emerging from the probe sample interaction zone into the far field.[24,25]

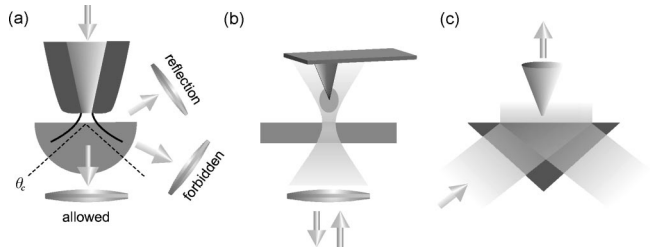

FIG. 1. Different types of scanning near-field optical microscopes: (a) aperture SNOM with angular resolved detection, (b) apertureless configuration, and (c) scanning tunneling optical microscope.

Figure 1(b) shows the intriguing ''apertureless'' NFO techniques.[26–34] In this configuration a strongly confined optical field is created at the apex of a sharply pointed probe tip by external (far-field) illumination. The relevant NFO signal therefore often has to be extracted from a huge background of far-field scattered radiation. Resolutions ranging from 1–20 nm have been reported in laboratory experiments. The general applicability of the technique to a wider range of samples is currently being investigated.

The third NFO microscopy configuration [see Fig. 1(c)] relies on the direct detection of localized fields in the near field above a sample by uncoated dielectric tips. The so-called scanning tunneling optical microscope (STOM),[35] also called photon scanning tunneling microscope (PSTM),[36] is schematically depicted in Fig. 1(c). In selected experiments the STOM led to results that could not have been achieved by other techniques.[37–40]

II. THE APERTURE SNOM

A. General setup

In this article we will concentrate on the aperture SNOM. Reaching a resolution of 50–100 nm on a routine base and having a potential down to 10–30 nm, SNOM is at least a factor of 5 to 10 better in resolution than a standard scanning confocal optical microscope with 1.4 NA. This represents an enormous progress in view of the relevance of the sub-100 nm regime today, and certainly, aperture SNOM is presently the most widely used and most developed NFO technique. This is also illustrated by the fact that all the commercial instruments currently available employ the aperture SNOM technique.

In our opinion the reasons for this situation are obvious. From a simplified point of view, imaging with SNOM is similar to scanning confocal optical microscopy except that the illumination spot is smaller. Of course, one has to keep in mind that SNOM is a surface selective technique, whereas scanning confocal optical microscopy can advantageously be applied to thick samples that require optical sectioning.[12,13] Nevertheless, all the contrast mechanisms such as absorption, phase, and fluorescence contrast, well-known in conventional optical microscopy, can be transferred more or less directly to SNOM applications. An aperture represents a very confined light source without any background. This is in contrast to the apertureless techniques where a rather large,

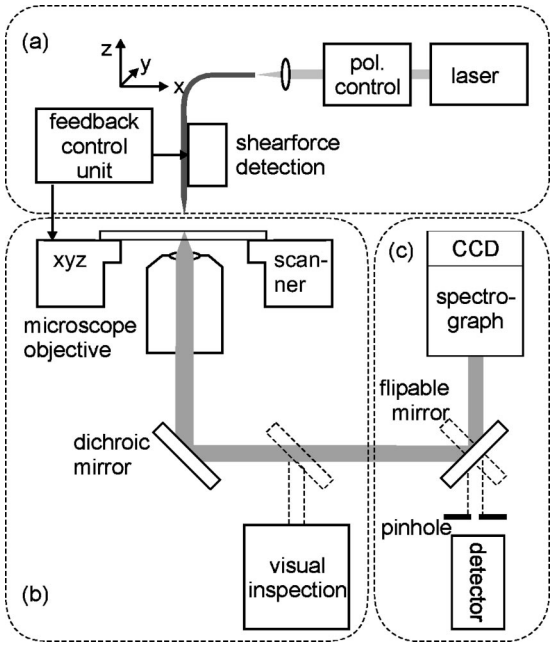

FIG. 2. Standard SNOM setup consisting of (a) an illumination unit, (b) collection and redistribution unit, and (c) a detection module.

intensive laser spot is focused onto a tiny tip apex. For fluorescence studies, which are among the most relevant applications of SNOM, a localized illumination is important because excessive photobleaching of the sample has to be avoided.

It is straightforward and tempting to extend existing conventional optical microscopes by a SNOM zooming stage. A typical SNOM setup, as it is realized in many laboratories and commercial instruments, is sketched in Fig. 2. Laser light of a suitable wavelength is coupled into an optical fiber that has an aperture probe at its far end. Control of the polarization and spectral filtering of the light is necessary before coupling into the fiber. The tip is mounted onto a mechanical support that includes adjustment screws to facilitate a coarse approach of the tip to the sample (into the range of the z-piezo displacement). Furthermore, it is often necessary to align the tip laterally onto the optical axis of the conventional microscope's objective or move it to specific areas of interest on the sample.

In order to measure and control the gapwidth between tip and sample, a strongly gapwidth-dependent signal is needed. Early SNOMs relied on electron tunneling feedback.[17,18] Today, the majority of SNOMs utilizes a method similar to noncontact atomic force microscopy (AFM) which is called shear force feedback.[41,42] The fiber probe is vibrated at one of its mechanical resonances parallel to the sample surface, ideally at amplitudes below 1–5 nm. Amplitude and phase of this minute fiber oscillation are monitored by a suitable displacement sensor.[41–46] During the final approach ($\approx 0-20$ nm) due to the action of shear forces, the resonance frequency is detuned with respect to the driving oscillator leading to a decrease in amplitude and to a phase shift. The origin of the shear force effect is not entirely clear. There might be a range of mechanism involved, including viscous damping in thin water films,[47,48] intermittent contact,[49] and electrostatic image forces.[50] The phase signal reacts faster to changes in the state of oscillation, because it is independent of the dissipation of kinetic energy stored in the resonance. Therefore, the use of pure phase feedback[51] or a combination of amplitude and phase feedback[44,46,52] is advantageous in fast shear force feedback systems. Light source, fiber, optical probe, shear force sensor, and mechanical support often make up an independent unit in an SNOM device as sketched in Fig. 2(a).

Light that is emitted by the aperture locally interacts with the sample. It can be absorbed, phase shifted, or locally excite fluorescence, depending on the sample and the contrast mechanism of interest. In any case, light emerging from the interaction zone has to be collected with the highest possible efficiency. For this purpose, high NA (oil immersion) microscope objectives or mirror systems are often used. A distinct advantage of an oil immersion objective when using a plane parallel substrate is that it can collect light that would otherwise undergo total internal reflection and thus fail to reach the detector. This is the so-called ''forbidden light.''[25,44,52,53] Forbidden light detection can lead to enhanced contrast and resolution in phase and amplitude contrast images.[54] In the case of mirror-based detection, hemispherical substrates providing a ''solid immersion'' are sometimes used to capture the forbidden light.[25,44,52,53]

The collected light is directed via a (dichroic) mirror either to a visual inspection port of the microscope or to a suitable detector [see Fig. 2(c)]. Filters can be inserted to remove unwanted spectral components. Light collection, redistribution, and filtering is advantageously done by means of a standard (inverted) optical microscope depicted in Fig. 2(b).

B. Fabrication of near-field optical probes

Great efforts have been devoted to the microfabrication of NFO probes.[55,56] However, few of the present designs are commercially available or have actually been used in NFO experiments. Therefore, the NFO community still relies heavily on aperture probes based on tapered optical fibers. Since fiber probes of sufficient quality are still expensive or even impossible to obtain commercially, tip fabrication remains of major interest to scientists working in the area of near-field optics.

The fabrication process for fiber-based optical probes can be divided in two main steps: (a) the creation of a transparent taper with a sharp apex, and (b) subsequent coating with aluminum to obtain an entirely opaque film on the cone walls and to form a transmissive aperture at the apex. There are two methods for the preparation of tapered optical fibers with a sharp tip and reasonable cone angle: (i) the ''heating and pulling'' method, and (ii) chemical etching.

The heating and pulling method is based on local heating using a CO_2 laser or a filament and subsequently pulling the fiber apart. The resulting tip shapes depend heavily on the temperature and the timing of the heating and pulling, as well as on the dimensions of the heated area.[44,57,58] The pulling method has the advantage that the glass surface on the taper is very smooth, which positively influences the quality of the evaporated metal layer. Furthermore, pulled tips often

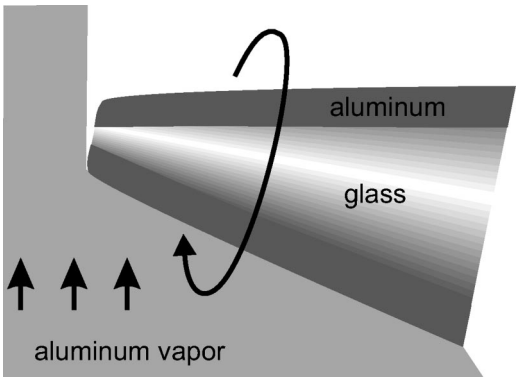

FIG. 3. Evaporation geometry of the aluminum coating process: evaporation takes place under an angle slightly from behind while the tip is rotating. The deposition rate of metal at the apex is much smaller than on the side walls.

exhibit flat facets at the apex due to a final fracture during pulling which facilitates the formation of an aperture during evaporation. It is, however, difficult, though not impossible, to obtain a large cone angle; correspondingly the transmission coefficient for a given aperture size is low, as to be discussed in Sec. III A 1.

Chemical etching of optical fibers in HF solutions became popular since Turner's etching technique[59] was introduced.[60] It comprises dipping the fiber into an HF solution covered with an overlayer of an organic solvent. Tip formation takes place at the meniscus that forms at the fiber at the interface between HF and solvent. The mechanism is based on the fact that the meniscus height is a function of the remaining fiber diameter. Chemical etching allows more reproducible production of larger quantities of probes in a single step, thus opening the road towards a laboratory-scale ''mass production'' of optical probes. A specific advantage of Turner's method is that the taper angle can be tuned by varying the organic solvent used as the overlayer. Accordingly, optical probes with correspondingly large transmission coefficient can be produced.[61,62] A major disadvantage of Turner's technique turned out to be the microscopic roughness of the glass surface on the taper walls. It leads to pinholes and imperfections in the metal coating that interfere with the true NFO signals.

This problem was solved only recently by introducing a method called tube etching.[63,64] In tube etching the fiber is not stripped before dipping it into HF. Tip formation takes place inside a small volume defined by the polymer coating due to convective flow of HF. The delicate meniscus between HF and the organic overlayer no longer plays a direct role in the tip formation process. The method is thus less prone to produce roughness on the taper while still providing large cone angles.

In both techniques, the aperture is formed during the evaporation of aluminum. Since the evaporation takes place under an angle slightly from behind, the deposition rate of metal at the apex is much smaller than on the sides. This geometrical shadowing effect leads to the self-aligned formation of an aperture at the apex, as is illustrated in Fig. 3.

In Fig. 4 we compare typical fiber probes produced by heating and pulling (left side) and tube etching (right side),

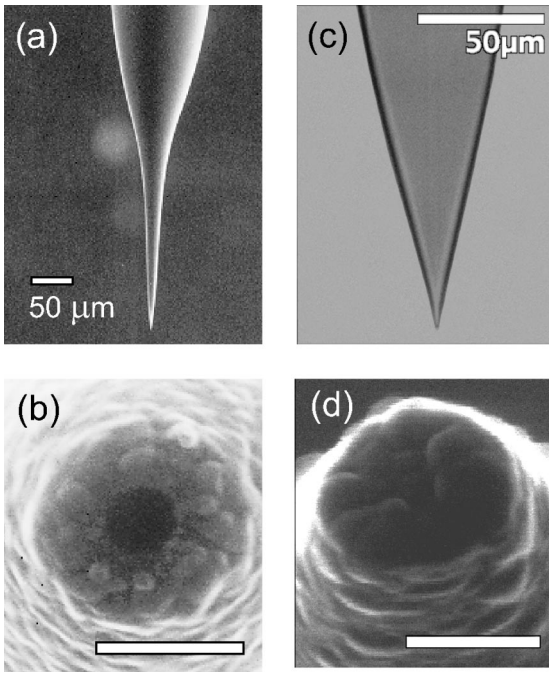

FIG. 4. Aluminum-coated aperture probes prepared by pulling (a),(b) and etching (c),(d): (a),(c) macroscopic shape, SEM and optical image. (b),(d) SEM close-up of the aperture region, scale bar corresponds to 300 nm.

respectively. A comparison of the different global shapes of the probes in Figs. 4(a) and 4(c) before aluminum coating evidences the improved cone angle and overall shape of the tube etched tips [Fig. 4(c)]. Figures 4(b) and 4(d) show scanning electron microscopy (SEM) images of the apex region of pulled and tube etched tips after aluminum coating. It is obvious that both tips exhibit a comparably high quality of the aluminum layer. The aperture can be observed for both tips.

III. SQUEEZING LIGHT THROUGH A SUBWAVELENGTH APERTURE

To achieve the ultimate performance possible in aperture SNOM, the optical probe should combine two main properties: (i) the spot size determined by the aperture diameter should be as small as possible. At the same time, (ii) the light intensity at the aperture should be as high as possible. This can be achieved either by optimizing the overall light throughput or by improving the damage threshold of the metal coating thus increasing the maximum input power. Finally, for practical reasons, delivering the light into the vicinity of the aperture should be as convenient as possible. This requirement is well met by aperture probes produced from standard optical fibers. Optical probes with small spot sizes and large transmission coefficients could open up fascinating new areas of research such as nanoscale nonlinear optics and optical surface modification on the nanometer scale. It is therefore of principal interest to understand the limitations and possibilities of aperture probes.

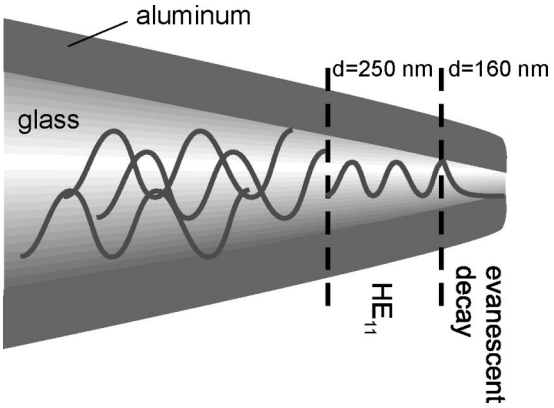

FIG. 5. Mode propagation in a tapered metal-coated optical fiber at a wavelength of 488 nm. Cutoff diameters taken from Ref. 66.

A. Transmission coefficient of aperture probes

1. The taper region

The transmission coefficient of an aperture probe at a given wavelength is defined as the power of the light coupled into the taper region divided by the light power emitted by the aperture. In the case of an optical fiber, the power coupled into the taper region is easily determined by cleaving the fiber behind the coupler and measuring the output. The measured value corresponds to the power in the fiber minus $\approx 4\%$ reflection loss.

The power emitted by the aperture is often measured in the far field. In principle, this is correct since the nonpropagating field components around the aperture do not contribute to the time-averaged Poynting vector.[65] However, in almost all experimental situations the *instantaneous* electric fields dominate light-matter interaction. Thus, the far-field transmission coefficient does not reflect the field enhancement effects in the near field. However, in comparing tips with the same aperture size, the far-field transmission coefficient can be valuable information.

The transmission coefficient of an aperture at the apex of a conical tip is determined by the characteristics of its two main structural components: (i) the part that guides light into the vicinity of (ii) the actual subwavelength aperture. In standard fiber probes, as depicted in Fig. 5, the guiding part corresponds to a tapered, metal-coated dielectric waveguide. The efficiency of guiding light to the aperture is determined by the distribution of propagating modes in the tapered waveguide. In the special case of a single-mode fiber with a taper produced by heating and pulling, the initial effect of the taper is a spreading of the guided mode since the guiding core is getting thinner and thinner. When the mode starts to ''feel'' the vicinity of the metal coating, the waveguide changes its character from a dielectric to a metallic hollow one filled with a dielectric. In the case of a taper produced by chemical etching, the mode spreading effect is absent because the core diameter is not affected by the chemical etching. As a result the influence of the metal coating starts to be important at smaller waveguide diameters closer to the aperture. In both cases, the mode structure in the metallic waveguide is completely different from the one in an unperturbed fiber leading to a reorganization of propagating modes accompanied by a rather strong back-reflection of light as a first source of attenuation.

The mode structure in a metallic waveguide at optical frequencies was calculated by Novotny and Hafner[66] as a function of the dielectric core diameter. The results can be interpreted such that in a tapered waveguide, where the core diameter decreases gradually, one mode after the other runs into cutoff until only the final HE_{11} mode is still propagating. An aluminum-coated ($\epsilon_{alu}=-34.5+i8.5$) dielectric ($\epsilon_{core}=2.16$) waveguide supports the pure HE_{11} mode at a wavelength $\lambda=488$ nm for inner diameters between ≈ 250 and 160 nm as sketched in Fig. 5. The transmission coefficient of light up to this point is mainly determined by the fraction of power contained in cutoff modes as compared to the power in the still-propagating HE_{11} mode. The magnitude of this factor is likely to depend on the geometry of the taper. This connection has been poorly understood up to now. Power not comprised in the propagating modes is either reflected back into the waveguide or absorbed in the metal coating, leading to a considerable heating of the metal.[67–70]

Below an inner diameter of 160 nm, even the HE_{11} mode runs into cutoff. Cutoff means that the wave vector becomes imaginary and thus the mode field decays exponentially. The power that is actually delivered to the aperture depends on the distance between the HE_{11} cutoff diameter and the aperture plane which is determined by the cone angle of the taper. This fact was first noticed by Novotny *et al.*,[71] who predicted that larger cone angles would drastically improve the overall transmission coefficient of a tapered waveguide structure.

2. The aperture

The transmission coefficient of a subwavelength hole in an infinitely thin perfectly conducting screen was first calculated rigorously by Bethe.[72] Errors in his expression for the near field were corrected in a paper by Bouwkamp.[73] The so-called Bethe/Bouwkamp model of a subwavelength aperture is important because it provides closed analytic expressions for the resulting electric and magnetic field. Although the model fails to accurately describe reality, the expressions still include the most characteristic features of the near- and far-field distribution of a realistic subwavelength aperture.

Within the Bethe/Bouwkamp model, it is expected that the transmission coefficient of a subwavelength hole should scale as a^4, where a denotes the aperture diameter.[72–74] An increase of the aperture diameter from 20 to 100 nm hence results in a transmission coefficient increased by a factor of 625. In order to illustrate this, we use transmission data calculated by Novotny *et al.*[71] for a realistic tip structure with different full taper angles α and aperture diameters of 10 and 20 nm. The calculations employed the multiple multipole method[75] to solve Maxwell's equations.

Scaling the numerical data according to the Bethe/Bouwkamp model provides a simple approximation of the transmission of a 100 nm aperture as a function of the taper angle (see Fig. 6). For values up to $\alpha=30°$, where the two curves run parallel, scaling the 10 nm curve by 10^4 and the 20 nm curve by 5^4 leads to transmission coefficients that

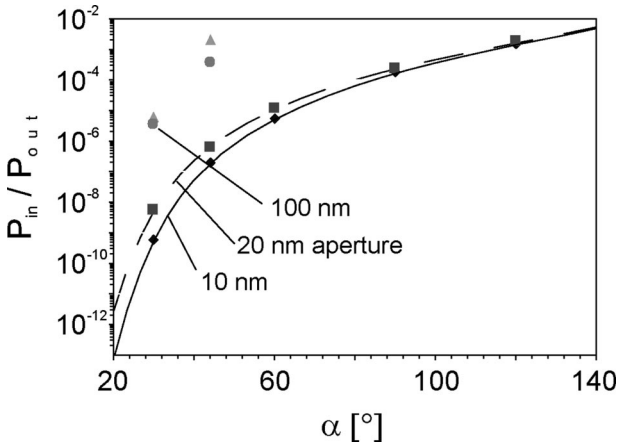

FIG. 6. Transmission coefficient of an aperture probe as a function of the full taper cone angle α. Squares and rhombs: calculated values from Ref. 71 for a 20 and 10 nm diameter aperture, respectively. Circles and triangles: interpolated values determined by scaling the numerical data according to the Bethe/Bouwkamp a^4 law by a factor of 10^4 and 5^4, respectively. This provides a simple approximation for the transmission coefficient of a 100 nm aperture.

nearly coincide. At larger taper angles the interpolated data points start to diverge. This is likely due to the breakdown of the Bethe/Bouwkamp model because the finite penetration depth of light into the aluminum (≈ 7 nm) starts to wash out the difference in diameter between the two apertures and the transmission through the metal screen becomes dominant.

From Fig. 6 it can be deduced that for a 100 nm aperture and a taper angle of about 30°, a transmission between 10^{-6} and 10^{-5} may be expected. For even larger taper angles around 42°, as they occur in etched fiber probes and in anisotropically etched silicon structures, the transmission coefficient should reach values around 10^{-3}. These numbers are in reasonable agreement with transmission coefficients reported in the literature and measured in our laboratories.[61,63,76] It should even be possible from such considerations to estimate the effective aperture size from transmission measurements if the taper angle at the apex is known from an independent (e.g., SEM) measurement.

Working with really small apertures ($<40-50$ nm) is often impossible since the transmission coefficient decreases dramatically with decreasing aperture size. This cannot be overcome by an increased input power because of the generally low damage threshold of the metal coating (≈ 10 mW). The relevant NFO signals becomes so small that the signal to noise ratio is no longer sufficient to provide reliable results. This is the reason why most NFO studies today are carried out with apertures of 80–100 nm.

B. Field distribution

Since we focus this article on aperture SNOM, it is important to investigate the character of the electric field distribution behind a small aperture. Let us consider a geometry that is even simpler than the Bethe/Bouwkamp configuration, namely a slit in a thin, good-conducting metallic screen (Fig. 7). The screen is illuminated at normal incidence by a plane wave with a wave number $k_0 = 2\pi/\lambda$, where λ is the wavelength [Fig. 7(a)]. Right behind this aperture, the transmitted field will be more or less confined to its width.[77]

The character of this field is best understood by considering its angular (Fourier) spectrum.[10] The electric field distribution behind the aperture can be obtained by convoluting the angular spectrum of the incident field with the Fourier transform of the aperture.[10] The Fourier transformation must be performed in the plane of the screen and provides an angular spectrum of plane (and evanescent) waves with field amplitudes $\gamma(k_\parallel, k_z)$. Here, $k_\parallel$ and k_z are respectively, the transverse and longitudinal components of the **k** vector.

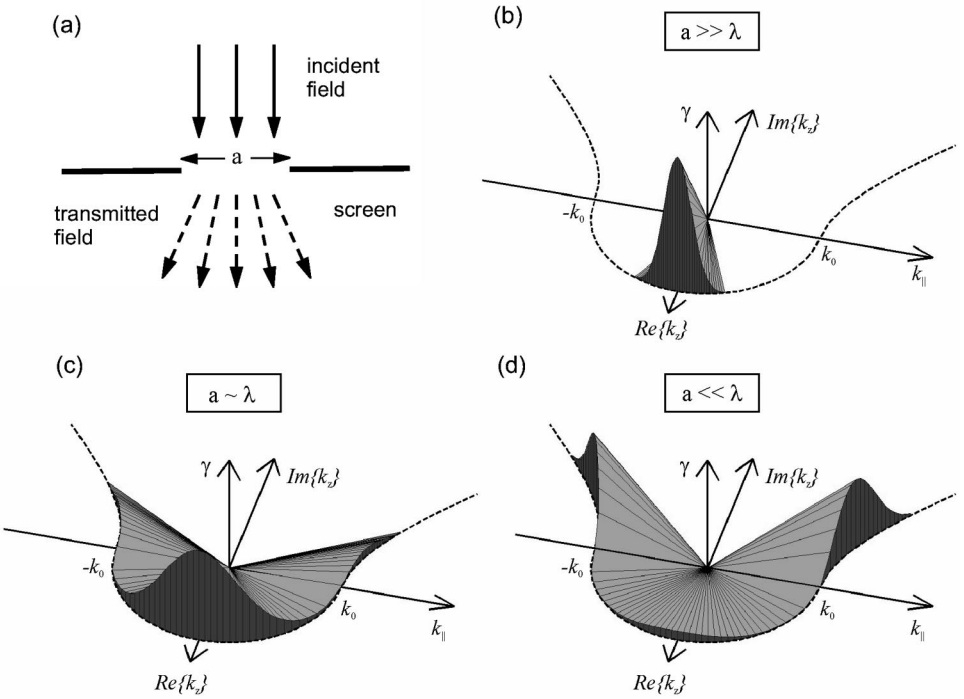

FIG. 7. Character of the electric fields behind an aperture: (a) A small aperture of size a in an infinitely thin and perfectly conducting screen is illuminated at normal incidence with a plane wave propagating in the z direction with a vector $k_z = k_0 = 2\pi/\lambda$. (b)–(d) The spectrum of the transmitted field is given for three different aperture sizes. In each case the vertical axis represents the field amplitude γ of the **k** vector with components $(k_\parallel, k_z)$, where the transverse component $k_\parallel$ is real and the longitudinal component k_z can be either real or imaginary. The latter corresponds to an evanescent field strongly localized at the vicinity of the aperture. The ratio of propagating and evanescent fields strongly depends on the aperture size a relative to the wavelength λ.

Since our aim here is not to give an exact treatment of the scattering by the slit but rather to illustrate the physical phenomena relevant to aperture SNOM, we will assume that the transverse component $k_{\|}$ is real. However, the longitudinal component k_z can be either real or imaginary. It is related to $k_{\|}$ via

$$k_0^2 = k_{\|}^2 + k_z^2. \tag{1}$$

Together, $k_{\|}$ and k_z describe the radiation properties of the plane and evanescent waves that constitute the field behind the aperture. The angular spectra for different aperture sizes are depicted in Figs. 7(b)–7(d). Field amplitudes $\gamma(k_{\|},k_z)$ concentrated at small $k_{\|}$ correspond to large apertures, whereas a distribution of $\gamma(k_{\|},k_z)$ over a broad range of $k_{\|}$ values corresponds to small apertures.

From Eq. (1) we see that, depending on the magnitude of $k_{\|}$, two different types of solutions are obtained for k_z

$$k_z = \begin{cases} \sqrt{k_0^2 - k_{\|}^2} & k_{\|} \leqslant k_0 \\ i\sqrt{k_{\|}^2 - k_0^2} & k_{\|} > k_0 . \end{cases} \tag{2}$$

As an example, let us consider an aperture of width a. The angular spectrum of this aperture will contain significant field amplitudes $\gamma(k_{\|},k_z)$ at $k_{\|} = 2\pi/a$. For this particular $k_{\|}$, the expression for the longitudinal wave vector component reads

$$k_z = \sqrt{\left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{2\pi}{a}\right)^2}. \tag{3}$$

It is obvious that for $a < \lambda$, k_z becomes imaginary.

The two sets of solutions described by Eq. (2) are depicted as a dashed line in Figs. 7(b)–7(d): for $|k_{\|}| \leqslant k_0$ the extremity of the **k** vector describes a circle in the $(k_{\|},k_z)$ plane, the k_z component being real. On the other hand, for $|k_{\|}| > k_0$ the solutions follow a square-root curve, with an imaginary k_z component. A field component with an imaginary wave vector in the propagation direction corresponds to an evanescent field. It hence remains strongly bound to the aperture and does not propagate into the far field. The strong localization is responsible for the resolution that can be achieved in scanning near-field optical microscopy.[78]

For a given wavelength, the ratio of propagating (k_z real) and evanescent (k_z imaginary) field amplitudes in the angular spectrum strongly depends on the size a of the aperture as compared to the wavelength λ [see Figs. 7(b)–7(d)]. When $a \gg \lambda$, the angular spectrum of the transmitted field is very similar to that of the incident field, i.e., most of the transmitted field propagates in the forward direction with a propagation vector $k_z \approx k_0$, $k_{\|} \approx 0$ [Fig. 7(b)]. Few field components with a small $k_{\|}$ can appear, leading to a slight divergence of the beam transmitted through the aperture.

When the aperture size becomes comparable to the wavelength, the lateral confinement increases and the spectrum broadens. As a result the transmitted field diverges strongly [Fig. 7(c)]. The angular spectrum of the transmitted field now also includes a small range of finite amplitudes with $|k_{\|}| > k_0$.

Finally, when the aperture is much smaller than the wavelength, the spectrum of the transmitted field in the z direction becomes dominated by the evanescent components; i.e., for the vast majority of the transverse components, $k_{\|}$ is larger than k_0 [Fig. 7(d)]. This means that the transmitted field is strongly localized at the vicinity of the aperture and decreases rapidly away from the aperture. As a result, the far-field power emitted by the aperture decreases, while strongly confined and enhanced fields appear at the vicinity of the aperture. Experimental evidence for the existence of nonradiative fields close to an aperture are the ''forbidden light'' emission,[24,44,54] surface plasmon excitation,[79] as well as an increased fluorescence from single molecules close to an aperture.[80]

IV. TIP–SAMPLE INTERACTION AND IMAGE FORMATION

We previously emphasized the importance of a localized light source for SNOM. However, in most SNOM experiments the field is detected at a large distance of the tip, in the far field. The mechanisms by which some components of the evanescent illuminating field can be transformed into propagating field components that carry information about the sample are at the core of image formation in SNOM.

Accurate and versatile computational techniques are mandatory to understand these intricate contrast mechanisms. One of the intrinsic difficulties in the simulation of near-field optical images is related to the fact that a practical experiment contains elements that have very different sizes: The sample is usually quite small, but it is deposited on a substrate with a quasi-infinite extension. The tip is very large too, but its interaction with the sample quite subtle. Finally, as already mentioned, while the relevant light–matter interaction takes place in the near field close to the sample, the image is formed by detecting intensity changes in the far field, at very large distances from the sample.

A great effort has been devoted, over the last 10 years, to the theoretical understanding of image formation in near-field optical microscopy, and two excellent review articles have been published by Girard and Dereux[81] and Greffet and Carminati.[82] However, it is only recently that self-consistent calculations taking into account these different length scales, and in particular the interaction between the tip and the sample in a fully three-dimensional manner, have been presented.[83,84] These results were obtained with the Green's tensor technique.[85] The decisive advantage of this approach to simulate SNOM experiments lies in the fact that it can accommodate complex systems that include both infinite substrates and strongly localized samples. Furthermore, it allows the computation of the measured far field while taking into account the near-field interaction in a fully self-consistent manner. Unfortunately, this approach is not very accurate when high permittivity materials are in play. This can, however, be superseded by using the filtered Green's tensor technique.[86]

Figure 8 presents results obtained with this approach. The geometry under study is depicted in Fig. 8(a), which shows a relatively large tip that is used to illuminate a small glass cube on a glass surface. In Fig. 8(b) we show an image calculated when the tip is raster-scanned over the cube at a constant height above the surface. The field transmitted

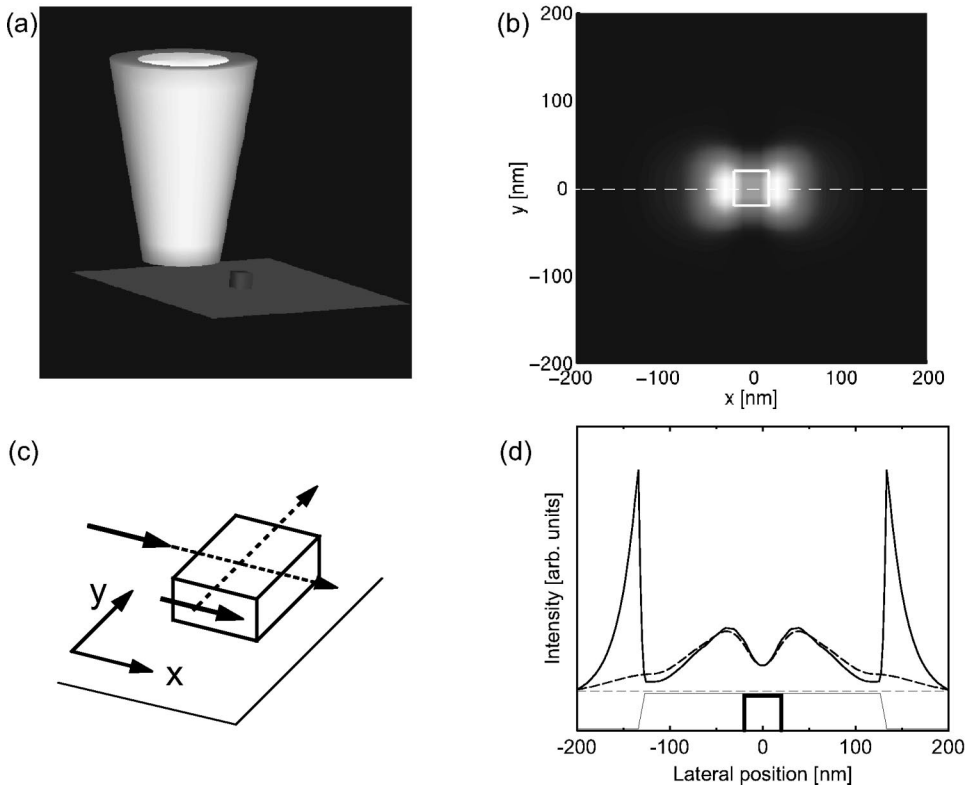

FIG. 8. Simulation of image formation in near-field microscopy. (a) An accurate model for an SNOM experiment should include an aperture tip with a reasonably large extension, a sample, and an infinite substrate supporting the sample. (b) Total transmitted electric field intensity computed for the geometry depicted in (a). The sample is made of glass ($\varepsilon=2$) and deposited on a glass substrate. The tip has a 50 nm opening with a 70 nm thick aluminum coating. The sample size is $40\times40\times40\ \mathrm{nm}^3$. The incident field is polarized in the x direction. To reproduce an experimental image, each pixel on this figure corresponds to a different calculation with a particular tip–sample position, the tip apex being kept at a constant height (2 nm) above the top sample surface. (c) The strong signals observed in (b) are related to depolarization fields created on the sample sides that are normal to the incident field. (d) Comparison of different scanning modes: constant height (dashed line) and constant gap (solid line) measured along the dashed line in (b). The corresponding tip motion is depicted at the bottom of the figure. When the sample is scanned in constant gap mode, i.e., when the tip motion follows the sample outline, the depolarization signal is much stronger as it is in constant height mode.

through the system is computed, as in a realistic experiment (see Fig. 2). We observe in Fig. 8(b) a strong asymmetry in the measured signal. In order to understand this effect, one must recall that when a system like the glass cube on the surface is illuminated with some external field $\mathbf{E}^0$, it generates a depolarization field $\mathbf{E}^{\mathrm{d}}$ such that the total field $\mathbf{E}^{\mathrm{t}}=\mathbf{E}^0+\mathbf{E}^{\mathrm{d}}$ fulfills the boundary conditions imposed by Maxwell's equations.[87] In our calculation, the illumination field $\mathbf{E}^0$ was chosen to be polarized in the x direction. Such an incident field is therefore parallel to the top and bottom sides of the sample [Fig. 8(b)]. Since an electric field parallel to an interface must be continuous across this interface, the incident field already satisfies the electromagnetic boundary conditions and no depolarization field is created along these sides [Fig. 8(c)].

On the other hand, this x-polarized incident field does not comply with the boundary conditions along both vertical sides, since it is normal to the sample interfaces. As a matter of fact, the total electric field $\mathbf{E}^{\mathrm{t}}$ should now have a discontinuity proportional to the sample permittivity at this interface [when the electric field is normal to an interface, it is the product of the field with the permittivity that must be continuous across the interface, Fig. 8(c)]. Therefore, the sample generates a strong depolarization field so that the total field fulfills the boundary conditions. It is this strong depolarization signal that is observed in the near-field image [Fig. 8(b)].

If instead of scanning at a constant height above the surface, a constant gapwidth is kept between the tip and the sample, one measures a very different signal [Fig. 8(d)]. In this scanning mode, the interaction between the incident field and the sample sides takes place during the entire upward motion of the tip necessary to follow the contour of the sample. This leads to the prominent depolarization peaks that are observed in Fig. 8(d) when the incident field is normal to the sample side. These peaks are located far away from the sample because they include the tip radius in addition to the sample side. The strong signal is hence merely related to the tip motion and constitutes a good illustration of a topographic artifact.[88] Such an artifact is very difficult to distinguish from the real near-field signal. Scanning at constant height therefore is the preferred mode of operation when working with slightly corrugated samples. On the other hand, our results also show that a careful control of the illumination polarization can provide another handle in order to keep these artifacts under control. In particular, a series of measurement of the same system with different illumination polarizations should allow discrimination between topography artifact and real near-field signal.

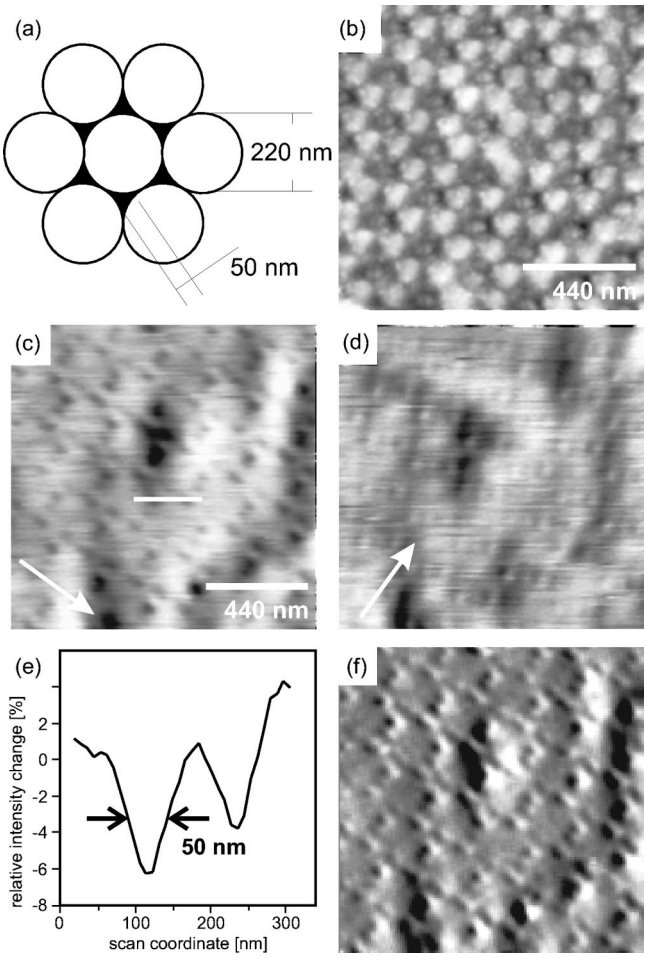

FIG. 9. Metal island film: (a) Sketch of metal patches arranged in hexagons in the interstices of densely packed 220 nm latex spheres. (b) Shear force image. (c), (d) Constant height forbidden light images taken at the same position as (b). The light emitted by the optical probe was linearly polarized in (c) and (d) in two orthogonal directions symbolized by the white arrows. (e) Intensity profile along the white line in (c). (f) High-pass-filtered version of (c) suppressing long-ranged intensity changes. The wavelength used for illumination was 633 nm.

V. SNOM APPLICATIONS

A. Amplitude and phase contrast

In order to obtain a phase or amplitude contrast image of a sample as discussed in Sec. IV, light emitted from the probe–sample interaction zone is collected and recorded directly by means of a suitable photodetector.

1. Metal island film

To demonstrate amplitude contrast, we chose a 15 nm thick aluminum island film produced by means of the latex sphere shadow mask technique.[89,90] The structure is sketched in Fig. 9(a). The objects of interest are the metal patches at the interstices between the spheres. They are formed by evaporation and subsequent dissolution of the spheres. To characterize their size, we propose to choose the altitude of the inscribable equilateral triangle as a standard. The length of this altitude is $h=(\sqrt{3}-3/2)d\approx 0.23d$, where d is the sphere radius. For the $d=220$ nm latex spheres used for the present test sample, the altitude hence is 50 nm. The nearest-neighbor distance between metal patches is $d/\sqrt{3}=127$ nm.

The calculated distance between the patches is in agreement with the shear force image [Fig. 9(b)] of the sample. The shear force image represents a convolution of a rather blunt protrusion on the aperture rim [cf. Figs. 4(c) and 4(d)] with the real sample topography. Therefore, the triangular shape of the metal patches, which was confirmed independently by SEM, is not visible.

Constant-height-mode optical images [Figs. 9(c) and 9(d)] were obtained with the same tip as used for Fig. 9(b) and at the same location on the aluminum island film, collecting the forbidden light emitted from the probe–sample interaction zone, only. The light emitted by the optical probe was linearly polarized [arrows in Figs. 9(c) and 9(d)] in two orthogonal directions. The extinction ratio measured was about 15. Figure 9(e) presents a cut through two neighboring dots at the position indicated by the white bar in Fig. 9(c). In Fig. 9(f) a high-pass-filtered representation of Fig. 9(c) is displayed achieving a better visibility of the small structures by removing long-ranged (far-field) modulations.

The optical scan images [Figs. 9(c) and 9(d) show a distinct fine structure. The hexagonal pattern, corresponding to the metal patches, is well visible. The structures in the optical images show systematic distortions. In particular, the patches are elongated in the direction of polarization. This might be attributed to the polar concentration of the electric energy on opposite sides of the rim of a small aperture.[71,73,91] A few nanometers behind the aperture plane this rather sharp double-peak structure is washed out, forming an approximately elliptic spot. The direction of strongest confinement hence corresponds to the direction perpendicular to the polarization in agreement with the observed elongation. It has to be pointed out that the metal patches strongly influence the field distribution at the aperture,[92] therefore, our discussion of spot shapes must remain qualitative if no realistic simulations are available.

The optical resolution in optical images [Figs. 9(c), 9(d), and 9(f), respectively] can be estimated from the line cut Fig. 9(e) to be at least 50 nm.

2. Dielectric grating

Phase contrast is obtained when imaging samples induce differences in the optical path. A typical example for such a sample is a dielectric grating. An AFM topograph of such a grating is shown in Fig. 10(a), accompanied by a height profile in 10(b). The grating structure has a period of 383 nm and a step height of 8 nm which was etched into a thin glass substrate.

Figure 10(c) shows a very high resolution constant height mode optical image recorded at a gapwidth below 1 nm. Figures 10(d) and 10(e) show constant height mode optical images recorded gap of a few nm and $\approx$100 nm, respectively. This series of images clearly evidences the strong dependence of the resolution on the gapwidth.

B. Fluorescence imaging

High resolution optical imaging of biological samples with fluorescent labels is among the most promising fields of application for SNOM. SNOM has the potential to image the

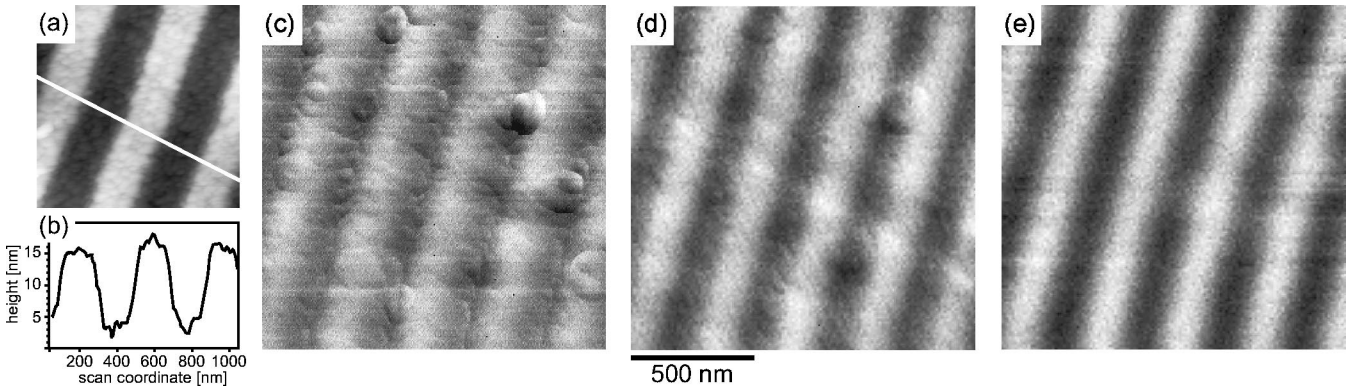

FIG. 10. Dielectric grating: (a) AFM topography. (b) Line cut through (b). (c) Very high-resolution constant height mode image close to contact. (d),(e) Constant height mode allowed images recorded at a gapwidth of a few, and ≈100 nm, respectively.

distribution of such labels down to the level of single fluorophores (see, for example, Refs. 4,5,93, and 94 and references therein).

Controllable deposition of well-defined species of biologically active molecules on surfaces is of interest in biology and biochemistry, since it opens up the possibility for nanometer-scale patterning and to carry out chemical assays on a nm scale with minute quantities of reagents.[95–97]

We investigated microcontact-printed patches of protein molecules, specifically chicken immunoglobulin labeled with tetramethyl-rhodamine-isothiocyanate (TRITC),[95] using SNOM. The patches consist of a monolayer of close-packed protein molecules, $\approx 4\times 10\times 14$ nm^3 in size. Each protein molecule is labeled by about 4 dye molecules. The fluorophores were excited at 514 nm and detection was performed above 570 nm, as described in Sec. II A using an oil immersion objective and a single photon counting avalanche photodiode. Figure 11(a) shows a near-field fluorescence image of two rectangular patches of proteins, nominally 500 $\times 3000$ nm^2 in size. Some areas of the sample show small irregularities. This feature makes the sample suitable for an SNOM resolution test. Figure 11(b) shows the simultaneously recorded shear force topographs of the structure. To estimate the optical resolution, intensity profiles were taken in the optical image [Fig. 11(a)] at two positions marked by arrows 1 and 2. The profiles are depicted in Figs. 11(c) and 11(d) as gray lines. The width of the small feature in Fig. 11(d) suggests a resolution of about 80 nm. To exclude the possibility of a z-motion artifact,[88] line cuts were taken at the same positions in the topographic image. The results are shown in Figs. 11(c) and 11(d) as black lines and prove that the feature in the optical image is not induced by a topographic crosstalk. A comparison of the line profiles taken in the optical and topographic image in 11(c) suggests a shift of

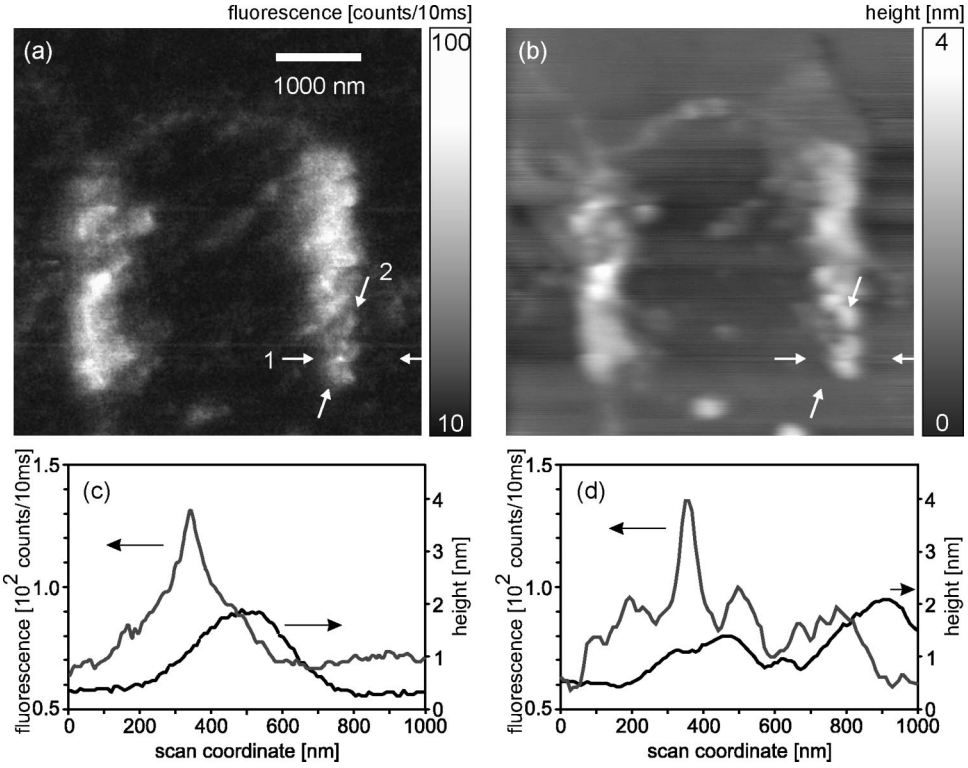

FIG. 11. (a) High-resolution (512 $\times$512 pixel) scanning near-field fluorescence image of a contact-printed pattern of TRITC-labeled chicken immunoglobulin molecules immobilized on a glass surface acquired simultaneously to the shear force image (b). The white arrows (1 and 2) mark the positions where the profiles shown in (c) and (d) were taken. (c) Fluorescence and topography profile along 1, and (d) Fluorescence and topography profile along 2.

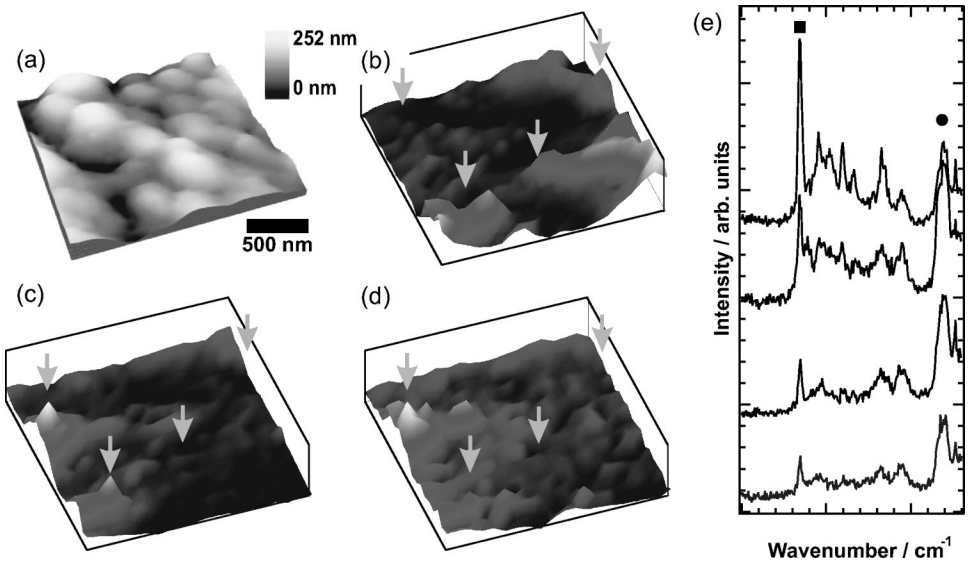

FIG. 12. SNOM–SERS imaging of brilliant cresyl blue labeled DNA fragments on an SERS substrate. Lateral dependence of (a) topography measured in shear force mode, (b) Raman intensity at 800 cm^{-1} corresponding to Raman scattering of the tip material (glass), (c) Raman intensity at 1641 cm^{-1} (BCB) before normalization, and (d) near-field surface enhanced Raman intensity at 1641 cm^{-1} after normalization with the glass signal. In (b), (c), and (d), the pixel size was only 100×100 nm, and for each graph the intensity of one line of the complete Raman spectrum was plotted. In (c) and (d), the contrast between the lowest and the highest Raman intensity was a factor of 3. The arrows in the spectral images correspond to the lateral positions of the Raman spectra shown in (e). The square marks a BCB band and the circle denotes a glass band. Note that the BCB signal never vanishes completely.

about 100 nm between the intensity and the topographic maximum. This is an effect that often occurs in SNOM images. It can be explained by assuming that a protrusion is located on the rim of the metal coating [see Figs. 4(c) and 4(d)] that actually is responsible for the shear force interaction. Thus, the topographic features are expected to be displaced from the center of the aperture.

C. Near-field Raman spectroscopy

In combination with vibrational spectroscopies, near-field optics can yield molecular contrast on a lateral scale determined by the aperture size. The main difficulties that are encountered applying this combination of techniques are the low efficiency of Raman scattering and the lack of suitable fiber material for infrared spectroscopy. Nevertheless, there are examples of both near-field Raman,[98–101] and near-field infrared spectroscopy.[33,102–104] A method to increase the small Raman scattering cross section to make it signals detectable in SNOM experiments is the use of rough noble metal substrates. Surface enhancement effects observed in the signals from samples deposited on such substrates can reach several orders of magnitude depending strongly on the morphology and shape of the metal features. The reason for the strong enhancement is still subject to some controversy.[105,106] Two different mechanisms seem to be involved. The so-called pure electromagnetic effect is due to plasmon resonances of the rough metal films or particles. The electromagnetic (near) field of a metal particle within this model can exceed the applied field by orders of magnitude. An additional enhancement can be caused by charge transfer or bond formation between the sample and the metallic substrate, which can strongly enhance the polarizability of the molecule.

Figure 12 shows topography (a) and near-field Raman images (b), (c), (d) and spectra from selected sample positions, of stained DNA strands on a specially prepared silver sphere substrate. Since the substrate has a rough topography, it is important to carefully distinguish between topographically induced contrast and intensity changes due to absorption of the substrate, and finally the Raman signal of the sample itself. Fortunately reflectivity and Raman intensity can be separated, at least to a first approximation.

The two images, Figs. 12(b) and 12(c), which were recorded sitting on different Raman bands, show quite different features. Figure 12(b) was recorded at a glass Raman band of the near-field probe, whereas Fig. 12(c) correspond to a Raman signal of the labeled DNA. The signals of 12(b) represent the reflectivity (absorption) of the silver sphere substrate. Of course, a higher reflectivity of the sample results in an increase in the detected Raman intensity as well; therefore, we use the reflectivity signal to correct for the sample topography[107] in all our spectra. The result is displayed in Fig. 12(d). It demonstrates that one of the prominent spots of Fig. 12(c) is actually due to a locally increased

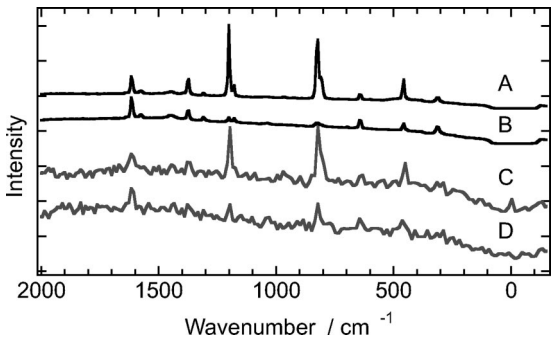

FIG. 13. Far-field (a), (b) and near-field Raman spectra (c), (d) of p-xylene. (a),(c) Parallel polarization; (b),(d) Vertical polarization geometry.

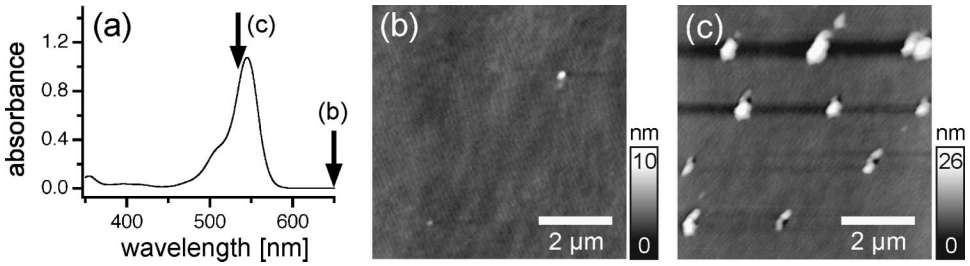

FIG. 14. Near-field optical ablation: (a) Absorption spectrum of rhodamine B at a concentration of 10^{-5} M in methanol solution. Arrows indicate the wavelengths used for the ablation experiments before recording the images on the right. (b),(c) Topography of a rhodamine B film after coupling 2.1 μJ into the SNOM tip at $\lambda = 650$ nm (b), and after coupling 1.4 μJ into the SNOM tip at $\lambda = 532$ nm (c), a wavelength near the maximum absorption of rhodamine B. The total topographic contrast in the *z* direction is 10 nm (b), and 26 nm (c), respectively.

reflectivity. The relation between reflectivity and Raman signal can be seen more clearly by comparing the selected uncorrected Raman spectra in Fig. 12(e). The arrows in the optical images correspond to the lateral positions of the chosen spectra. Obviously, a substantial change in the ratio of the BCB–DNA band (square) and the glass scattering (circle) occurs only from 1 to 2–3, whereas the absolute intensities of BCB–DNA and glass intensity vary in each spectrum.

In the case of surface enhanced Raman scattering the signal intensities can be very high. In fact, single-molecule sensitivity in experiments using conventional optical microscopy has already been claimed by several groups.[108,109] For NFO experiments the detection of a few hundred molecules was shown.[110] The method thus seems to have the potential for molecular recognition on a nm scale.

Near-field Raman spectroscopy is not limited only to solid surfaces, but can be used as well in liquids.[111] Figure 13 compares the near-field and far-field Raman spectra of xylene. The data demonstrate that the distinct polarization properties in the near field of an optical probe are useful to investigate, for instance, the conformation and the degree of crystallinity of a sample.

D. Pulsed laser ablation through SNOM tips

Optical probes produced by tube etching are stable enough to withstand several tenths of mJ of pulsed laser radiation coupled into them.[112] The intensity at the aperture can be in excess of 10^8 W/cm^2, which is enough to cause ablation of a sample close to the tip. Several mechanisms are possible for laser-induced ablation through SNOM tips.[113] The photochemical model assumes that the energy is deposited in the sample via optical absorption, and ablation follows as a consequence of chemical bond breaking. Similarly, in a photothermal scenario, the absorbed optical energy would be converted into heat, causing ablation from the locally heated spot. Third, a ballistic mechanism is conceivable, where material from the tip is sputtered onto the sample, causing ablation of material from its surface. A ballistic mechanism was the most likely explanation for an experiment carried out on an anthracene sample using a completely metallized SNOM tip.[114] In the absence of any photons reaching the sample, sputtering by either metal atoms from the tip, or by material adsorbed on its surface, was believed to be the reason for the ablation process. Fourth, one also has to take into account transient thermal expansion[69,115,70] as a possible mechanism for the creation of surface indentations, but it is likely to be negligible.[114]

We now discuss an experiment that gave very clear results about the ablation mechanism from a rhodamine B film. This sample was irradiated either on resonance (532 nm) or off resonance (650 nm) with respect to its optical absorption maximum [Fig. 14(a)]. SNOM tips with apertures of <100 nm diameter were used. No significant drop in optical output of the SNOM tip was measured for similar input energies at 400 and 650 nm input wavelength, although some decrease is expected due to the increasing difference between wavelength and aperture size. Figures 14(b) and 14(c) show shear-force topographic images recorded after several laser pulses had been fired onto the sample. No ablation was observed for the off-resonance wavelength [Fig. 14(b), $\lambda = 650$nm, 2.1 μJ pulse energy]. In contrast, well defined, ≈ 70 nm diameter (FWHM), 5 nm deep holes were created by a wavelength that is strongly absorbed by the sample [Fig. 14(c), $\lambda = 532$ nm, 1.4 μJ pulse energy]. This clearly points to an optical ablation mechanism, either photochemical or photothermal. Energy transfer by absorption of photons appears to be more efficient than by a ballistic process. The photon energy used here is not sufficient for inducing chemical bond dissociation in the rhodamine film, in contrast to ablation processes with ultraviolet laser radiation, where a photochemical ablation mechanism is often used to rationalize the observed phenomena.[116] We therefore suggest that a photothermal mechanism is responsible for the ablation of the rhodamine film: we believe that the absorbed photon energy is converted to vibrational excitation, causing a rapid temperature increase that leads to thermal desorption of the molecules.

Another interesting observation in Fig. 14(c) is that the ablated material is redeposited on the sample surface close to the ablation crater. The distribution of redeposited material was not symmetric in this experiment, but was always to the left side of the crater, independent of the scan direction. In other SNOM ablation experiments, the distribution was more symmetric. The origin of the directionality appears to be the tip itself: frequently, the apex of an SNOM tip exhibits some uncontrolled asymmetry, causing the ablated material to be blown off in a specific direction. This has interesting applications for the analytical nanosampling of material surfaces.

VI. PERSPECTIVES

With this review, by focusing on fundamentals and selected applications of aperture SNOM, we cover only a small part of the research currently pursued in the area of NFO. For a more complete picture the reader is referred to other NFO review articles in this issue and to recent reviews.[2–5,9]

The rapid development of NFO techniques is far from coming to an end. Many problems remain to be solved and many careful experiments are still to be done.

With respect to the technical problem of tip fabrication, there is still plenty of room for improvements, in particular regarding tip stability, damage threshold, and transmission coefficients. Microfabricated near-field optical probes[55,56] will help to solve these problems and improve reproducibility by providing better defined experimental conditions. Also, new tip concepts which deviated from the simple aperture scheme may be of advantage.

Another possibility to improve the performance of aperture probes is the use of ultraviolet (UV) light.[117–122] This will shift the cutoff region in the taper towards the apex and therefore strongly increase the transmission coefficient. This can be exploited by using smaller apertures. In addition, direct excitation of certain biomolecules might become possible instead of using fluorescent markers. Furthermore, in Raman spectroscopy it is advantageous to use UV excitation because of the ν^4 dependence of the scattering efficiency and because of the potential of resonant field enhancement at small silver particles in this spectral region (for a review, see Ref. 123).

Apertureless techniques are becoming more and more important.[26–34,124] The reported resolutions are exciting, but their utilization for relevant applications such as fluorescence imaging seems to be less straightforward than for aperture SNOM. This is mainly related to the difficulties of discriminating between effects generated in the enhanced field at the tip apex and the background related to the diffraction limited spot illuminating the tip. For fluorescence imaging, fluorescence resonant energy transfer (FRET) microscopy,[125–128] another apertureless contraption, is one of the most promising candidates for a significant improvement of resolution.

We also did not mention all of the numerous new contrast mechanisms that are constantly evolving. We did not mention, e.g., time-resolved SNOM, which combines nanometer scale optical resolution with femtosecond time resolution. A number of groups are already working in this area.[129–132] Pulsed laser techniques in principle provide the possibility to utilize nonlinear optical techniques like second harmonic generation, hyper-Raman, etc. Such techniques can help to avoid stray light problems in SNOM[34,124] and also can improve the signal-to-noise ratio.

As another example, the combination of local desorption using SNOM with mass spectrometric detection provides a promising new analytic method.[133]

From our point of view, SNOM is a field where the interaction between modeling and experiment can be extremely fruitful. This interaction, however, has not yet been as successful as one might wish. This is probably related to the intrinsic limitations of numerical simulations to account for a realistic experimental configuration (including tip, sample, and detection scheme) and conversely to experimental difficulties to precisely control all the relevant parameters. As a matter of fact, our simulations illustrated, for example, in Fig. 8, the importance of the illumination polarization for image interpretation; still, the exact polarization state of the field emerging from a real SNOM tip is difficult to control. One might wish, for the future, that experimentalists try to design their experiments such that they can be compared to theoretical models. On the other hand, theoreticians have to extend their numerical capabilities. Both sides would profit from such collaborations.

ACKNOWLEDGMENTS

The authors wish to thank A. Bernard for the kind gift of the contact printed sample and L. Novotny for helpful discussions and for providing the tip transmission data. We further acknowledge financial support from the NFP36 and the ETH-Zürich; O.J.F.M. gratefully acknowledges support from the Swiss National Science Foundation.

[1] D. Pohl, in *Scanning Near-field Optical Microscopy (SNOM)*, *Advances in Optical and Electron Microscopy Vol. 12* (Academic, New York, 1991), pp. 243–312.
[2] H. Heinzelmann and D. Pohl, Appl. Phys. A: Mater. Sci. Process. **59**, 89 (1994).
[3] C. Girard and A. Dereux, Rep. Prog. Phys. **59**, 657 (1996).
[4] R. Dunn, Chem. Rev. **99**, 2891 (1999).
[5] P. Barbara, D. Adams, and D. O'Connor, Annu. Rev. Mater. Sci. **29**, 433 (1999).
[6] *Near Field Optics*, *NATO ASI Series E: Applied Sciences*, Vol. 242, edited by D. Pohl and D. Courjon (Kluwer Academic, Dordrecht, 1993), Proceedings of NFO I: Besançon/Arc & Senans, France, 26–28 October 1992.
[7] *Near Field Optics*, *Ultramicroscopy*, Vol. 57, edited by P. Kruit and M. Isaacson (North-Holland, Elsevier, Amsterdam, 1995), Proceedings of NFO II, Raleigh, NC, 20–22 October 1993.
[8] *Near Field Optics and Related Techniques*, *Ultramicroscopy*, Vol. 61, edited by P. Kruit, M. Paesler, and N. van Hulst (North Holland, Elsevier, Amsterdam, 1995), Proceeding of NFO III, Brno, Czech Republic, 9–11 May 1995.
[9] *Proceedings NFO-5*, *J. Microscopy*, Vol. 194, edited by S. Kawata (Blackwell Science Ltd., Oxford, 1999), Proceeding of NFO 5, Shirahama, Japan, 6–10 December 1998.
[10] J. W. Goodman, *Introduction to Fourier Optics, Physical and Quantum Electronics Series* (McGraw-Hill, New York, 1968).
[11] M. Abbé, Archiv. Mikroscop. Anat. Entwicklungsmech. **9**, 413 (1873).
[12] T. Wilson and C. J. R. Sheppard, *Theory and Practice of Scanning Optical Microscopy* (Academic, London, 1984).
[13] *Handbook of Biological Confocal Microscopy*, edited by J. Pawley (Plenum, New York, 1995).
[14] E. Synge, Philos. Mag. **6**, 356 (1928).
[15] E. Synge, Philos. Mag. **11**, 65 (1931).
[16] G. Binnig and H. Rohrer, Helv. Phys. Acta **55**, 726 (1982).
[17] D. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984).
[18] U. Dürig, D. Pohl, and F. Rohner, J. Appl. Phys. **59**, 3318 (1986).
[19] D. Pohl, US Patent US4,604,520 (1986).
[20] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, Ultramicroscopy **13**, 227 (1984).
[21] A. Harootunian, E. Betzig, M. Isaacson, and A. Lewis, Appl. Phys. Lett. **49**, 674 (1986).
[22] E. Betzig, M. Isaacson, and A. Lewis, Appl. Phys. Lett. **51**, 2088 (1987).
[23] E. Betzig and J. Trautman, Science **257**, 189 (1992).
[24] H. Heinzelmann, B. Hecht, L. Novotny, and D. Pohl, J. Microsc. **177**, 115 (1994).
[25] B. Hecht, H. Heinzelmann, and D. Pohl, Ultramicroscopy **57**, 228 (1995).
[26] U. Fischer and D. Pohl, Phys. Rev. Lett. **62**, 458 (1989).
[27] F. Zenhausern, Y. Martin, and H. Wickramasinghe, Science **269**, 1083 (1995).
[28] Y. Inouye and S. Kawata, J. Microsc. **178**, 14 (1995).

[29] P. Gleyzes, A. Boccara, and R. Bachelot, Ultramicroscopy **57**, 318 (1995).
[30] C. Girard and O. M. A. Dereux, Phys. Rev. Lett. **75**, 3098 (1995).
[31] J. Koglin, U. Fischer, and H. Fuchs, J. Biomed. Opt. **1**, 75 (1996).
[32] T. Sugiura, S. Kawata, and T. Okada, J. Microsc. **194**, 291 (1999).
[33] B. Knoll and F. Keilmann, Nature (London) **399**, 134 (1999).
[34] E. Sanchez, L. Novotny, and X. Xie, Phys. Rev. Lett. **82**, 4014 (1999).
[35] D. Courjon, K. Sarayeddine, and M. Spajer, Opt. Commun. **71**, 23 (1989).
[36] R. Reddick, R. Warmack, and T. Ferrell, Phys. Rev. B **39**, 767 (1989).
[37] O. Marti *et al.*, Opt. Commun. **96**, 225 (1993).
[38] J. Krenn *et al.*, Phys. Rev. Lett. **82**, 2590 (1999).
[39] A. Choo *et al.*, Ultramicroscopy **57**, 124 (1995).
[40] N. van Hulst, M. Moers, and E. Borgonjen, *NATO Series E: ''Photons and Local Probes,''* edited by O. Marti and R. Möller (Kluwer Academic, Dordrecht, 1995), Vol. E 300, p. 165.
[41] R. Toledo-Crow, P. Yang, Y. Chen, and M. Vaez-Iravani, Appl. Phys. Lett. **60**, 2957 (1992).
[42] E. Betzig, P. Finn, and S. Weiner, Appl. Phys. Lett. **60**, 2484 (1992).
[43] K. Karrai and R. Grober, Appl. Phys. Lett. **66**, 1842 (1995).
[44] B. Hecht, *Forbidden Light Scanning Near-Field Optical Microscopy*, Ph.D. thesis, University of Basel (Hartung–Gorre, ISBN 3-89649-072-9, Konstanz, 1996).
[45] G. Tarrach, M. Bopp, D. Zeisel, and A. Meixner, Rev. Sci. Instrum. **66**, 3569 (1995).
[46] M. Pfeffer, P. Lambelet, and F. Marquis-Weible, Rev. Sci. Instrum. **68**, 4478 (1997).
[47] T. Okajima and S. Hirotsu, Appl. Phys. Lett. **71**, 545 (1997).
[48] S. Davy, M. Spajer, and D. Coujon, Appl. Phys. Lett. **73**, 2594 (1998).
[49] M. Gregor, P. Blome, J. Schöfer, and R. Ulbrich, Appl. Phys. Lett. **68**, 307 (1996).
[50] H. Bielefeldt, Ph.D. thesis, University of Konstanz, 1994.
[51] A. Ruiter, J. Veerman, K. van der Werf, and N. van Hulst, Appl. Phys. Lett. **71**, 28 (1997).
[52] B. Hecht, D. Pohl, H. Heinzelmann, and L. Novotny, in Ref. 40, Vol. 300, p. 93.
[53] T. Huser *et al.*, J. Opt. Soc. Am. A **16**, 141 (1999).
[54] B. Hecht *et al.*, J. Appl. Phys. **84**, 5873 (1998).
[55] S. Münster *et al.*, J. Microsc. **186**, 17 (1997).
[56] W. Noell *et al.*, Appl. Phys. Lett. **70**, 1236 (1997).
[57] E. Betzig *et al.*, Science **251**, 1468 (1991).
[58] G. Valaskovic, M. Holton, and G. Morrison, Appl. Opt. **34**, 1215 (1995).
[59] D. Turner, US Patent 4,469,554 (1984).
[60] P. Hoffmann, B. Dutoit, and R.-P. Salathé, Ultramicroscopy **61**, 165 (1995).
[61] D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, Appl. Phys. Lett. **68**, 2491 (1996).
[62] T. Yatsui, M. Kourogi, and M. Ohtsu, Appl. Phys. Lett. **73**, 2090 (1998).
[63] R. Stöckle *et al.*, Appl. Phys. Lett. **75**, 160 (1999).
[64] P. Lambelet *et al.*, Appl. Opt. **37**, 7289 (1998).
[65] M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Cambridge University Press, Cambridge, 1980).
[66] L. Novotny and C. Hafner, Phys. Rev. E **50**, 4094 (1994).
[67] A. L. Rosa, B. Yakobson, and H. Hallen, Appl. Phys. Lett. **67**, 2597 (1995).
[68] D. Kavaldjiev, R. Toledo-Crow, and M. Vaez-Iravani, Appl. Phys. Lett. **67**, 2771 (1995).
[69] M. Stähelin *et al.*, Appl. Phys. Lett. **68**, 2603 (1996).
[70] C. Lienau, A. Richter, and T. Elsaesser, Appl. Phys. Lett. **69**, 325 (1996).
[71] L. Novotny, D. Pohl, and B. Hecht, Opt. Lett. **20**, 970 (1995).
[72] H. Bethe, Phys. Rev. **66**, 163 (1944).
[73] C. J. Bouwkamp, Philips Res. Rep. **5**, 321 (1950).
[74] D. V. Labeke, D. Barchiesi, and F. Baida, J. Opt. Soc. Am. A **12**, 695 (1995).
[75] C. Hafner, *The Generalized Multiple Multipole Technique for Computational Electromagnetics* (Artech, Boston, 1990).
[76] J. A. Veerman, A. M. Otter, L. Kuipers, and N. F. van Hulst, Appl. Phys. Lett. **72**, 3115 (1998).
[77] H. Levine and J. Schwinger, Commun. Pure Appl. Math. **3**, 355 (1950).
[78] O. J. F. Martin, C. Girard, and A. Dereux, J. Opt. Soc. Am. A **13**, 1801 (1996).
[79] B. Hecht *et al.*, Phys. Rev. Lett. **77**, 1889 (1996).
[80] W. Moerner *et al.*, Phys. Rev. Lett. **73**, 2764 (1994).
[81] See Ref. 3.
[82] J.-J. Greffet and R. Carminati, Prog. Surf. Sci. **56**, 133 (1997).
[83] O. J. F. Martin, J. Microsc. **194**, 235 (1999).
[84] O. J. F. Martin, in *Near-field Optics: Physics, Devices, and Information Processing, Proceedings of SPIE* Vol. 3791 (SPIE, Bellingham, 1999), pp. 1–8.
[85] O. J. F. Martin and N. B. Piller, Phys. Rev. E **58**, 3909 (1998).
[86] N. B. Piller and O. J. F. Martin, IEEE Trans. Antennas Propag. **46**, 1126 (1998).
[87] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
[88] B. Hecht *et al.*, J. Appl. Phys. **81**, 2492 (1997).
[89] U. C. Fischer and H. P. Zingsheim, J. Vac. Sci. Technol. **19**, 881 (1981).
[90] H. W. Deckman and J. H. Dunsmuir, Appl. Phys. Lett. **41**, 377 (1982).
[91] L. Novotny and D. Pohl, in Ref. 40, Vol. 300, p. 21.
[92] L. Novotny, D. W. Pohl, and B. Hecht, Ultramicroscopy **61**, 1 (1995).
[93] M. Garcia-Parajo *et al.*, Bioimaging **6**, 43 (1998).
[94] N. van Hulst, J. Chem. Phys. **112**, 7799 (2000), this issue.
[95] A. Bernard *et al.*, Langmuir **14**, 2225 (1998).
[96] T. Schmidt, G. Schütz, H. Gruber, and H. Schindler, Anal. Chem. **68**, 4397 (1996).
[97] W. Trabesinger *et al.*, Anal. Chem. **71**, 279 (1999).
[98] D. A. Smith *et al.*, Ultramicroscopy **61**, 247 (1995).
[99] C. L. Jahncke, H. D. Hallen, and M. A. Paesler, J. Raman Spectrosc. **27**, 579 (1996).
[100] D. Zeisel *et al.*, Anal. Chem. **69**, 749 (1997).
[101] S. Webster, D. N. Batchelder, and D. A. Smith, Appl. Phys. Lett. **72**, 1478 (1998).
[102] M. K. Hong *et al.*, SPIE **2863**, 54 (1996).
[103] A. Piednoir, C. Licoppe, and F. Creuzet, Opt. Commun. **129**, 414 (1996).
[104] A. Lahrech, R. Bachelot, P. Gleyzes, and A. C. Boccara, Appl. Phys. Lett. **71**, 575 (1997).
[105] M. Moskovits, Rev. Mod. Phys. **57**, 783 (1985).
[106] A. Otto, I. Mrozek, H. Grabhorn, and W. Akemann, J. Phys.: Condens. Matter **4**, 1143 (1992).
[107] V. Deckert, D. Zeisel, and R. Zenobi, Anal. Chem. **70**, 2646 (1998).
[108] K. Kneipp *et al.*, Phys. Rev. Lett. **78**, 1667 (1997).
[109] S. Nie and S. R. Emory, Science **275**, 1102 (1997).
[110] D. Zeisel, V. Deckert, R. Zenobi, and T. Vo-Dinh, Chem. Phys. Lett. **283**, 381 (1998).
[111] J. Grausem, B. Humbert, and A. Burneau, Appl. Phys. Lett. **70**, 1671 (1997).
[112] See Ref. 63.
[113] B. Dutoit, D. Zeisel, V. Deckert, and R. Zenobi, J. Phys. Chem. B **101**, 6955 (1997).
[114] D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, Appl. Phys. Lett. **68**, 2491 (1996).
[115] H. J. Mamin, Appl. Phys. Lett. **69**, 433 (1996).
[116] D. Singleton, G. Paraskevopoulos, and R. S. Taylor, Chem. Phys. **144**, 415 (1990).
[117] I. Smolyaninov, D. Mazzoni, and C. Davis, Appl. Phys. Lett. **67**, 3859 (1995).
[118] S. Monobe *et al.*, Opt. Commun. **146**, 45 (1998).
[119] A. Kirsch *et al.*, Biophys. J. **75**, 1513 (1998).
[120] A. Jenei *et al.*, Biophys. J. **76**, 1092 (1999).
[121] S. Nishikawa and T. Isu, J. Microsc. **194**, 15 (1999).
[122] A. Kramer *et al.*, Biophys. J. **78**, 458 (2000).
[123] H. Metiu, Prog. Surf. Sci. **17**, 153 (1984).
[124] Y. Kawata, C. Xu, and W. Denk, J. Appl. Phys. **85**, 1294 (1999).
[125] V. L. S. K. Sekatskii, Appl. Phys. B: Lasers Opt. **63**, 525 (1996).
[126] V. L. S. K. Sekatskii, JETP Lett. **65**, 465 (1997).
[127] S. Vickery and R. Dunn, Biophys. J. **76**, 1812 (1999).
[128] G. Shubeita, S. Sekatskii, M. Chergui, and G. Dietler, Appl. Phys. Lett. **74**, 3453 (1999).
[129] J. Levy *et al.*, Phys. Rev. Lett. **76**, 1948 (1996).
[130] A. Richter *et al.*, Appl. Phys. Lett. **73**, 2176 (1998).
[131] B. A. Nechay *et al.*, Appl. Phys. Lett. **74**, 61 (1999).
[132] T. Guenther *et al.*, Appl. Phys. Lett. **75**, 3500 (1999).
[133] D. A. Kossakovski *et al.*, Ultramicroscopy **71**, 111 (1998).