An imaging system that makes no use of prior information isdiffraction-limited in its resolution to $\sim \lambda/2n$, where $\lambda$ is the wavelength of the light and $n$ is the refractive index of the immersion medium. The propagating far-field waves captured by the imaging lens do not contain the high-spatial-frequency information required for subdiffraction-limited imaging of an object. In the optical near-field, however, nonradiative evanescent waves do contain such information. The transfer of information from the near-field to the far-field thus has a limited bandwidth, requiring that high-resolution images be captured extremely close to the sample. This is typically performed with a near-field scanning optical microscope (NSOM) in which a sharp metallic probe, subwavelength metallic aperture, or metallic nanoparticle is introduced into the vicinity of an object of interest. Metallic structures are able to confine the electromagnetic energy—typically as surface plasmons—into subdiffraction-limited volumes. This can thus serve both to illuminate the sample locally with an extremely small light source and to collect or scatter near-field electromagnetic energy into the far-field, where it can be captured by a conventional optical system. The collected signal contains a wealth of information about the electronic, vibrational, structural, and dynamical properties of the sample.

When a small aperture (aperture NSOM) is used for local illumination of a sample, the ultimate lateral imaging resolution depends largely on the size of the aperture and the skin depth of the metal coating. Apertureless NSOM, on the other hand, uses sharp metal tips or nanoparticles to concentrate or scatter light into and out of deeply subwavelength volumes. Here, the tip diameter determines the imaging resolution, which can be $\sim 10\times$ better than aperture NSOM because of the fabrication difficulties and optical throughput limitations of ultrasmall apertures. Tip-based NSOM has been used to study the local field distribution on a variety of samples, including plasmonic waveguides and antennas, biological membranes, and integrated circuits. The method has also been successfully

**KEYWORDS:** near-field scanning optical microscopy · plasmonics · template stripping · tip-enhanced Raman scattering · single-molecule fluorescence · super-resolution imaging · optical antenna
applied for single-molecule fluorescence imaging\textsuperscript{12,13} and for vibrational imaging via Raman scattering.\textsuperscript{14–16} Clearly, as a super-resolution imaging technique, NSOM provides important nanoscale information inaccessible by other methods.

Unfortunately, the use and widespread implementation of NSOM is severely limited by the low reproducibility of near-field probes. To date, there is no reliable fabrication method that would provide reproducible probes with high yield and throughput. The grain

structure and roughness of evaporated metals introduces a high degree of variability in the fabrication of both aperture and tip-based near-field probes.\textsuperscript{1} Sculpting and shaving of probes by focused ion beam (FIB) milling provides smoother surfaces, but it is slow and costly and leads to unwanted ion implantation.\textsuperscript{17} Better reproducibility is achieved with near-field probes made of colloidal metal nanoparticles,\textsuperscript{18,19} but the picking and attaching of single nanoparticles to the end of pulled fibers is painstaking and inefficient. Furthermore, because of quenching, these nanoparticle probes require particles larger than 60 nm, which limits the attainable optical resolution. Smaller nanoparticles can be used only in combination with low-quantum-yield emitters,\textsuperscript{20} high-index substrates,\textsuperscript{21} or more elaborate probe geometries.\textsuperscript{22}

In this paper, we demonstrate the mass fabrication of high-quality, uniform, ultrasharp (<10 nm) metallic probes suitable for single-molecule fluorescence, single-molecule tip-enhanced Raman spectroscopy (TERS), and other techniques where the local field enhancement must be large and lateral imaging resolution must be high. The tip geometry is shown to be optimized for such imaging experiments, with a usable tip yield of >95%.

RESULTS AND DISCUSSION

Our method is based on a template stripping technique\textsuperscript{17,23–25} that has been shown to produce a variety

Figure 1. Template-stripping fabrication technique. (a) A silicon nitride mask (green) is used to define regions for anisotropic etching of silicon with KOH. The apex angle is 70.52°. (b) Following a single metal deposition and (c) a lift-off procedure, the pyramidal tips are ready (d) for template stripping.

Figure 2. Scanning electron microscopy (SEM) images of template-stripped metallic tips. (a) \(\sim 1.5\) million nominally identical tips can be fabricated over an entire wafer. (b) Close-up image of a single 200 nm thick gold tip resting in the inverted silicon mold after lift-off. The tips can be stored over an extended period of time without degradation. (c) SEM image after template stripping with epoxy and a thin tungsten wire. (d) The radius of the tip is 10 nm, suitable for high-resolution near-field imaging.
of metallic nanostructures, including ultrasharp tips, with ultrasmooth patterned metallic surfaces. The tips are template-stripped from a high-quality silicon mold, as shown in Figure 1. Critical fabrication of the tips does not require the use of slow and expensive nanofabrication tools such as FIB or electron-beam lithography. Using a 100 nm thick Si₃N₄ mask on a silicon wafer, large openings (tens of μm s) are created in the silicon nitride with standard photolithography exposure, development, and etching. A subsequent chemical etch in potassium hydroxide (KOH) creates inverted pyramidal molds in the silicon (Figure 1a).²⁶ This anisotropic etching process exposes the {111} crystal facets of the silicon, which join at the apex with an open angle of 70.52°.¹⁷ Interestingly, this large apex angle is particularly well suited to scatter the near-field optical signal into the far-field, as shown below. It is important that the KOH etching recipe provides smooth {111} sidewalls and sharp tips. This is accomplished with prolonged overetching (>1 h) in a mixture of 30% KOH, 10% isopropyl alcohol, and water. Following a single metal deposition (Figure 1b) and a hydrofluoric acid lift-off bath to remove the Si₃N₄ (Figure 1c), the tips are ready to be template-stripped (Figure 1d). The use of standard photolithography allows parallel fabrication and with the density used here (50 μm between pyramids) can easily produce over 1.5 million tips over a single 4 in. wafer, each with uniform properties, as shown in Figure 2a. A single tip in the silicon mold is shown in Figure 2b. By placing a small droplet of quick-setting epoxy on a short piece of 15 μm diameter tungsten wire (Alfa Aesar), a single pyramidal tip can be plucked from the mold, shown in Figure 2c. The tips have a radius of ~10 nm. These are then mounted to the prong of a quartz tuning fork and implemented into an NSOM system.

In previous work we have used gold tips fabricated by electrochemical etching or FIB milling for near-field fluorescence imaging⁹ and for near-field Raman scattering.¹⁵,²⁷ However, the quality of the results depended strongly on the particular properties of the tip. Although all the tips looked very similar when imaged by scanning electron microscopy (SEM), most of them did not yield any measurable field enhancement. Thus, near-field imaging with gold tips so far has relied on a trial-and-error approach with a low yield (~5%) and low repeatability. On the other hand, we find that gold tips fabricated by template stripping have a high yield. In fact, over 95% of the tips that were tested were usable for near-field imaging and provided similar resolution, in both fluorescence and Raman scattering.

A sketch of the experimental setup is shown in Figure 3. The sample is placed onto an x–y piezo scan-stage on top of an inverted confocal optical microscope. Additionally, a home-built atomic force microscope scan head is placed on top of the microscope, allowing the pyramidal tip to be positioned in the center of the optical focus. A tightly focused radially polarized optical excitation is used, providing a strong longitudinal electric field at the optical focus and giving maximum electric field enhancement from the pyramidal tip.¹ The sample is raster-scanned below the pyramidal tip, allowing for simultaneous topographical and optical images. The tip–sample separation (~5 nm) is maintained by using either shear-force or dynamic normal mode feedback.²⁸ Photons emitted from the sample are collected by the objective and sent to either an avalanche photodiode (APD) or a spectrometer and liquid nitrogen-cooled charge coupled device.

Figure 3. Setup used for near-field imaging with template-stripped gold pyramidal tips. A radially polarized laser beam is focused on the sample surface. A pyramidal tip is centered into the diffraction-limited spot, generating an enhanced field concentrated at the tip–sample junction and generating a local optical response (scattering, fluorescence, Raman scattering, etc.). The optical response is collected with the same objective lens and directed to a spectrometer or a single-photon counting detector while the sample is being raster-scanned. Insert: SEM image of a pyramidal near-field probe. Scale bar: 10 μm.
hand, near-field fluorescence imaging not only resolves individual molecules but also identifies the orientation of the molecular transition dipole axis. Molecules oriented along the axis of the pyramid (z-axis) reveal an optical enhancement of around 200-fold and an optical resolution of 18 nm, both due to the pyramidal tip. In-plane molecules exhibit a characteristic double-lobe pattern. Figure 4c shows the fluorescence emission rate of a single z-oriented dye molecule as a function of the pyramid/sample distance. The fluorescence enhancement is determined by the ratio between the maximal and minimal measured fluorescence rate in the approach curve, after adjusting for the photoluminescence of the tip. A maximum fluorescence rate enhancement of approximately 200-fold is observed. The resolution and enhancement far exceed that of an 80 nm gold sphere imaging at a glass—air interface, which has been used in other work for near-field fluorescence imaging.

We have also tested pyramidal probes for near-field Raman imaging. Past efforts in TERS have focused on using chemically etched gold or silver tips. While some of these probes provide exceptionally large field enhancements, they are not reproducible due to the variations in etch parameters and crystal structure of the metal wire. However, the pyramidal tips presented here allow for higher measurement reproducibility as well as quantitative models because of the well-defined probe geometry. To demonstrate the feasibility of using these pyramids for TERS and near-field Raman imaging, we used a sample of carbon nanotubes (CNTs) produced by the arc-discharge method. Because the same tube bundle can be located and measured repeatedly with each tip, the CNTs provide an excellent sample to characterize the pyramidal tips in terms of field enhancement and reproducibility.

Figure 5a shows a near-field image of the Raman G band, excited with a 785 nm laser, for an arc-discharge CNT bundle. The cross section of the near-field optical signal (arrow in Figure 5a) yields a width of 40 nm (Figure 5a inset). Note that the 40 nm corresponds to the convolution of the optical field localization (the resolution) with the actual width of the nanotube bundle. The corresponding topographic image, Figure 5b, shows a nanotube bundle width of 6.2 nm. The spectra of the CNT bundle with the tip close to the surface and retracted are shown in Figure 5c. Taking the ratio of these two
spectra for a Raman band provides a measure of the enhancement factor, which in this case is approximately 10. Figure 5 clearly demonstrates that the pyramidal tips are well suited for TERS experiments.

The measured Raman enhancement is less than the fluorescence enhancement because a carbon nanotube is a one-dimensional (1-D) object, while a fluorescent molecule is zero-dimensional (0-D). Illuminating a 0-D object with a confocal excitation extending an area A and a near-field antenna enhancing the field by a factor $f_0$ in a subdiffraction area $a$ will yield a measured enhancement roughly (no quenching) equal to $f_0^2$. For the 1-D case, it will be scaled by $(a/A)^{1/2} \approx 0.2$ because of the large physical extent of the confocal excitation. Additionally, the nanotube bundles imaged further increase the background by having more than one nanotube in the confocal excitation simultaneously, lowering the measured enhancement. The Raman enhancement mechanism for one-dimensional objects has been analyzed and discussed in detail previously.27

To understand the factors that are responsible for the large enhancements, we have performed finite-difference time-domain (FDTD) calculations for both pyramidal tips and conical tips of variable tip angle $\alpha$ and for different wavelengths $\lambda$. In particular, the calculations were used to determine the radiative properties of a quantum emitter placed in front of a tip. In the experiments, the tip is irradiated from the front by a focused higher-order laser beam. The same objective lens that is used for focusing is also used to collect photons due to the tip–sample interaction. Thus, it is evident that the signal-to-noise depends on the fraction of power that is radiated in the backward direction, that is, away from the tip and toward the objective lens. The fraction of power that is radiated in the forward direction couples predominantly to surface plasmons propagating along the sides of the tip.30–32

The energy associated with these modes is ultimately dissipated to heat, although a structured tip shaft could be used to release some of this energy into the far-field.33

![Figure 6](image-url)

Figure 6. Backward radiation efficiency of a dipole in front of a pyramidal tip as a function of wavelength $\lambda$ and opening angle $\alpha$. (a) Illustration of the computational model. The backward radiation efficiency is equal to $1$ for a dipole in free space (absence of tip). The dipole is placed $3.75$ nm from the tip apex, and its dipole axis is parallel to the tip axis. (b) Computational results. The solid curves are FDTD results obtained from a computational window of size $2.43 \mu m \times 2.43 \mu m \times 1.63 \mu m$. The mesh size is $2.5$ nm in all cases. The results clearly indicate that backward radiation increases with opening angle $\alpha$.

To calculate the fraction of power radiated in a backward direction, we place an electric dipole at a distance of $3.75$ nm in front of a gold tip and evaluate the radiation patterns (Figure 6a). The dipole orientation is chosen to be parallel to the tip axis. We use perfectly matched layers at the boundaries to avoid spurious reflections and evaluate the backward radiation (BR) efficiency, defined as the power flux through the bottom half space ($z < 0$) normalized with the corresponding power radiated by an isolated dipole in free space. Accordingly, the BR efficiency in the absence of the tip is 1. Calculations are performed for both pyramidal and conical tips of variable opening angles $\alpha$, and the results are similar. We therefore only show the data for pyramidal tips. Note that a tip represents an infinitely extended structure and that terminating its length for computational reasons can generate severe artifacts. This is even the case if perfectly absorbing layers are used. It is thus necessary that the computational window is comparable to or larger than the surface plasmon propagation length. Because the latter increases with wavelength, memory and processing time constraints prevent us currently from accurately calculating the BR efficiency at near-infrared wavelengths.

Our theoretical results, shown in Figure 6b, show that the BR efficiency increases as the wavelength $\lambda$ and the cone angle $\alpha$ are increased. This trend is expected because plasmon propagation along the facets of the pyramid becomes strongly mode-mismatched for large $\alpha$. As pointed out by Stockman and co-workers, plasmons propagate most efficiently along the tip shaft in the so-called adiabatic limit, that is, for small $\alpha$.31 However, for our situation the adiabatic limit is detrimental since it leads to enhanced dissipation. We find that increasing $\alpha$ from 10$^\circ$ to 70$^\circ$ enhances the backward radiation by more than a factor of 10 at a wavelength of $\lambda = 650$ nm. Note that this enhancement is due not only to a redistribution of the radiation pattern but mostly to electromagnetic back-action; that is, the tip enhances the dipole’s ability to release energy. Thus, an enhanced BR
efficiency corresponds to an increased radiative decay rate. This increase prevents a quantum emitter from complete quenching (Figure 4c) and is the reason for our ability to perform high-quality near-field fluorescence imaging on samples with single molecules.

CONCLUSION

In conclusion, we have presented a highly reproducible and effective method for the fabrication and assembly of high-quality near-field probes. The method is based on template stripping and yields many tips fabricated in parallel on a standard silicon wafer. The optimized pyramidal geometry provides spatial resolutions of less than 20 nm and fluorescence enhancements of ~200x. The template-stripped near-field probes introduced here promise to make near-field microscopy and spectroscopy a routine technique. This will impact not only nanoscale characterization of material surfaces but also high-throughput sensing, failure analysis, and heat-assisted magnetic recording.

METHODS

Pyramid Fabrication. First, 100 nm of low-stress nitride was grown on new Si wafers. Then, MEGAPOST SPR-955 photostripe (Rohm and Haas) was spin-coated on the wafers, exposed with an i-line stepper (Canon 2500 I3) using a mask to produce 5, 10, 15, and 20 μm holes, and developed for 70 s in MF CD 26 (Rohm and Haas) using a CEE 200X (Brewer Science) spray developer. Next, using the resist as an etch mask the nitride was etched using RIE (STI model 320) with CF4. The resist was then removed with an oxygen plasma, and the wafers were put in a bath of 30% KOH, 10% isopropyl alcohol, and water for 90 min for the anisotropic etching. After etching, the wafers were rinsed for 30 min and cleaned with a 1:1 solution of sulfuric acid and hydrogen peroxide, removing any excess KOH salt crystals, and dried. Next, 200 nm of Au was evaporated on the wafers using an electron-beam evaporator (CHA, SEC600) followed by a lift-off process where the wafers were soaked in 49% hydrofluoric acid for 20 min to remove the nitride mask, giving isolated Au pyramids. More fabrication details (including the fabrication of Ag pyramids and structured pyramids) are given in refs 24 and 25.

Spin Coating of a PMMA Film. To generate ~2 nm thick films of PMMA, a 20 μL drop of 0.01% 5000MW PMMA in toluene was placed onto a cover glass spinning at 3000 rpm. The thickness of the PMMA film has been measured by using a razor blade to generate a scratch and using an atomic force microscope to subsequently measure the depth of the cut relative to the PMMA surface. Conflict of Interest: The authors declare no competing financial interest.

Acknowledgment. We thank P. Bharadwaj for valuable input and advice. This work has been supported by grants to LN, from the Department of Energy (DE-FG02-01ER15204), the National Science Foundation (CBET-0930074), and grants to S.-H.O. from the Office of Naval Research (ONR) Young Investigator Award (N00014-11-1-0645), the National Science Foundation (NSF) CAREER Award (DII 1054191), and the WCU Program #R31-10032 funded by the Ministry of Education, Science & Technology and the National Research Foundation of Korea. Device fabrication was performed at the University of Minnesota Nanofabrication Center, which receives partial support from the NSF through the National Nanotechnology Infrastructure Network.

REFERENCES AND NOTES


