ORIGINAL PAPER

Signal limitations in tip-enhanced Raman scattering: the challenge to become a routine analytical technique

Samuel Berweger · Markus B. Raschke

Abstract The fundamental parameters and limitations that determine the signal strength in tip-enhanced Raman scattering (TERS) are discussed. A semiquantitative analysis of the Raman signal expected in different experimental geometries and with different sample systems is presented, taking into account experimental parameters including Fresnel factor, numerical aperture of the illumination and collection optics, detection efficiency, and the Raman scattering cross section of the material. A top/side-on illumination geometry is essential for the study of non-transparent samples. It can yield the highest signal levels when strong tip–sample coupling using a metallic substrate provides large field enhancement. In contrast, axial/through-sample illumination is limited to transparent sample materials. Although conceptually simpler in experimental implementation and despite high numerical aperture signal collection efficiency, signals are generally weaker due to limited field enhancement. Crystalline solids with small Raman cross sections and dense molecular/biological systems with unavoidable far-field background provide the biggest challenge for TERS analysis yet at the same time hold the most exciting outstanding scientific questions TERS has the potential to answer.

Keywords TERS · Raman scattering cross section · Sensitivity · Signal levels

Tip-enhanced scanning near-field optical microscopy has emerged as a promising new route for ultrahigh spatial resolution imaging. Here, the nanoscopic apex of a plasmonic scanning probe tip is used to locally enhance and confine the electromagnetic field. In its Raman implementation as tip-enhanced Raman scattering (TERS), the nanometer spatially resolved vibrational spectroscopic surface chemical analysis of molecules [1–7], organic nanocomposites [8], and carbon nanotubes [9–11], as well as crystalline solids [12–17] has been demonstrated. The development of TERS has been significantly aided by the fundamental understanding of the possible role of the excitation of a surface plasmon polariton in the probe tip [18], the associated field enhancement [19–23], and improvements in tip fabrication [24–26]. This enabled the demonstration of spatial resolution down to the sub-10-nm range [10, 27, 28] and single-molecule sensitivity [3, 6, 7]. More recent experiments have explored TERS for the study of biological systems [29–33] and the extension of TERS for optical nanocrystallography [12, 13] using the symmetry selectivity of the phonon Raman response. However, despite almost 10 years of continuous effort [5, 34], the routine application of TERS for surface analysis has remained challenging. From a technical perspective, the reproducible fabrication of TERS-active tips with a defined plasmon resonance and sufficient local field enhancement has been difficult. Even for appropriate tips with apex radii of ~10 nm, the enhancement values for the local optical electric field are limited to order of ~10–20 and thus lower than early expectations (due to, e.g., damping and skin depth) [21, 22]. Therefore, a frequently overlooked critical fundamental question is how high a signal level can actually be expected within the physical limitations set by field enhancement, damage threshold, tip scattering efficiency, collection and detection efficiency, and the Raman scattering cross section of the sample material.

S. Berweger · M. B. Raschke (✉)
Department of Chemistry, University of Washington,
Seattle, WA 98195, USA
e-mail: raschke@chem.washington.edu

Here, we provide a semiquantitative analysis of the signal levels that can be expected in TERS from different sample materials and in different experimental geometries. The estimate is based on the range of values for the relevant physical parameters including illumination in terms of the orientation, polarization, and numerical aperture (NA) of the illumination and collection optics used, detection efficiency, tip enhancement, and Raman scattering cross section (including resonant excitation) of the analyte. The most commonly used top/side-on and axial illumination geometries are compared. We establish that the highest signal levels can be achieved in a top/side-on configuration when strong tip sample coupling via a metallic sample provides large local field enhancement. Axial illumination benefits from high NA signal collection efficiencies and can be the method of choice for transparent samples and in case a metal substrate is not beneficial or suitable.

Our results show that TERS in either geometry can be of readily detectable signal strength over a wide range of experimental parameters. However, the true challenge arises from the combination of far-field and fluorescence background, bleaching, reproducibility of fabrication of suitable TERS-active tips, and imaging artifacts due to topography and tip–sample forces. Therefore, many factors which lead to the successful demonstration of ultrahigh resolution (~10 nm) and single-molecule TERS may prove difficult to export to a routine application.

In the following, we first discuss the different TERS-relevant parameters which are calculated individually. Then, the resulting signal levels are discussed for various sample material and experimental scenarios, focusing specifically on molecular and crystalline systems.

Field enhancement

The field enhancement underlying TERS primarily arises from the evanescent near-field generated at the apex region of an ultrasharp metallic scanning probe tip as determined by its structural and material parameters [18] and further increased by possible excitation of plasmon polariton modes. A schematic of the multistep (conceptualized as a sequential time ordered process) tip enhancement and scattering process is shown in Fig. 1. The pump light is focused onto the tip–sample gap, where the incident far-field E_i generates a spatially localized near-field distribution in the immediate tip apex vicinity. Here, the tip serves as an optical antenna, spatially focusing part of the incident far-field radiation into the tip–sample gap. This near field induces an optical polarization in the material at the Raman-shifted frequency $\vec{P}_{NF}(\omega + \Delta) = \bar{\bar{\chi}}\vec{E}_{NF}(\omega)$, where $\bar{\bar{\chi}}$ is the Raman tensor and $\vec{E}_{NF} = L_i'\vec{E}_{inc}$ is the tip apex near-field, enhanced by a factor L_i'. The tip acts again as an

optical antenna for enhanced radiation coupling of the Raman polarization into the far field by scattering. The total electric field amplitude of the Raman scattered light is then given by $\vec{E}_s = L_s'\bar{\bar{\chi}}L_i'\vec{E}_{inc}$ with the scattered light enhanced by the factor L_s'.

As a consequence of this two-step excitation and emission tip enhancement process, the intensity I_s of the Raman scattered light is given by $I_s = (L_s'L_i')^2 \bar{\bar{\chi}}^2 I_i$. Due to, e.g., the plasmonic nature of the enhancement, L_s' and L_i' are frequency dependent as given by the spectral profile of the tip response in relation to the excitation and Raman scattered wavelengths [35, 36]. However, as plasmonic line widths [18] are generally broader than or at least comparable to the amount of Raman shift, we approximate $L_s' \approx L_i'$ in the following. This underlies the commonly used assumption of $I_{\text{Raman}} \propto |L'|^4$ in SERS and TERS [36, 37].

In addition to its spectral variation, the tip-field enhancement is also polarization dependent [17, 20]. This is due to the tip asymmetry, where a larger field enhancement occurs for excitation of the longitudinal plasmon mode when the light is polarized along the tip axis (axial direction, p-polarized), as opposed to the perpendicular direction (radial direction, s-polarized) [38]. Throughout this work but without loss in generality, we assume p-polarized excitation and emission.

Experimentally, Raman enhancement factors in the range of 10^4–10^5 have been found for free-standing tips (i.e., not placed near a metallic surface as will be discussed further below; [7, 10, 15, 17, 19, 23] and [39], and references therein). Assuming $I_{\text{Raman}} \propto |L'|^4$, this translates into field enhancement values of order ~10 in good agreement with theoretically predicted values in the range of 10–50 [19–23, 40] for typical apex radii of ~10–20 nm.

The effective enhancement can be increased by taking advantage of the near-field coupling between the tip and a metallic substrate [19, 23, 41]. The resulting Raman enhancement of up to 10^9 has been shown to be sufficient to observe a single-molecule TERS response in some cases [3, 4, 6, 7, 42]. This enhancement results from imposing the boundary conditions at the substrate surface onto the evanescent near field of the tip. The resulting surface charge density distribution screening the evanescent field can be expressed in the form of an image dipole of the tip located within the substrate. The field enhancement in the gap is then significantly increased over that of a free-standing tip with a typical field enhancement of ≥100 for typical tip–sample distances of several nanometers [19, 23].

The tip enhancement, however, may not be the only source of local field enhancement. In addition, the local incident far-field experienced by the tip apex itself is modified through the presence of a reflective substrate surface. This is expressed through the Fresnel factor F, as well as established in classical optics [43, 44], which will
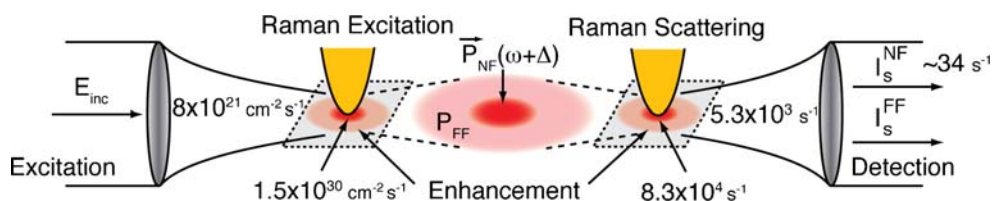
Fig. 1 Illustration of the TERS mechanism. Incident light is focused onto the tip–sample gap with a focus size significantly larger than the tip apex. The tip (plasmon) enhanced local near field interacts with the sample, resulting in a local polarization at the Raman-shifted frequency $\vec{P}_{NF}(\omega + \Delta)$. Subsequently, the tip-mediated and enhanced Raman scattered light is detected in the optical far field. Associated with the far-field illumination, the diffraction-limited excitation generates a far-field Raman background signal I_{FF}. Assuming an incident photon flux of $8 \times 10^{21} \, \mathrm{cm}^{-2}\mathrm{s}^{-1}$, a Raman enhancement $(L_i L_s)^2 = 1.5 \times 10^8$ gives rise to an effective flux of $1.5 \times 10^{30} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$ in the tip-enhanced near-field region. This results in an 8.3×10^4 $\mathrm{sr}^{-1}\mathrm{s}^{-1}$ near-field scattered photons assuming $d\sigma/d\Omega = 10^{-26} \, \mathrm{cm}^2 \, \mathrm{sr}^{-1}$. Of these near-field scattered photons, ~34 will be detected by the pixel of maximum intensity on the CCD array detector

affect both the incident and scattered light. Although frequency dependent, we assume no change in the value of F over the frequency range between the illumination and Raman scattered frequencies. Thus, in a backscattering geometry with the same illumination and detection angles, F will be the same for the incident and scattered light. In combination with the tip enhancement, the effective TERS enhancement becomes $L_i = L_i' F_i$ and $L_s = L_s' F_s$. With the maximum value of $F=2$, the Fresnel factor can enhance the Raman signal by up to a factor of 16.

In addition, the Fresnel factor affects the far-field Raman signal originating from the diffraction-limited focus of the incident light. The far-field tip scattering may further contribute to the detected background signal. The a priori separation from the near-field signal is difficult and a frequent source of signal artifacts. We further address this issue below, in particular the elimination of the far-field background in TERS from crystalline samples [17, 45].

We note that collectively $(L_i L_s)^2$ is the total effective TERS enhancement frequently referred to in the literature. However, field enhancement values are often calculated through comparison of near-field and far-field intensity of, e.g., the molecules on the substrate in the same geometry; thus, the Fresnel factors generally cancel, but they need to be considered when comparing TERS with, e.g., the molecular far-field Raman response in solution or gas phase.

Experimental geometry

TERS is generally implemented in one of two geometries: side-on illumination using a long working distance microscope objective or axial illumination in an inverted microscope configuration. The choice is a tradeoff between distinct advantages and disadvantages of one or the other geometry as discussed below.

A side-on illumination setup (as shown schematically in Fig. 2) allows for the use of optically nontransparent samples and substrates. In particular, this gives access to crystalline samples and allows for the easy use of metallic substrates, where plasmonic tip–sample coupling as well as the associated Fresnel factors of ~2 can significantly increase the field enhancement. Also, this simplifies the selection of s- and p-polarized excitation and detection, in particular the p-polarized illumination of the tip apex to drive the axial tip plasmon, necessary for large field enhancement and symmetry-selective TERS probing of particular Raman tensor elements as discussed above. As a consequence, the largest Raman enhancement values obtained to date in TERS have been achieved in such a geometry [3, 7]. However, this geometry restricts the use of high NA objectives, with values typically in the range of 0.28–0.5, with a maximum theoretical value of 0.7 given by the geometrical constraints imposed by a tip orientation normal with respect to surface. This geometry is illustrated in Fig. 2.

In contrast, axial illumination allows for the use of high NA, including oil-immersion objectives, providing for increased collection efficiency and reduced far-field background. However, the need for a longitudinal field
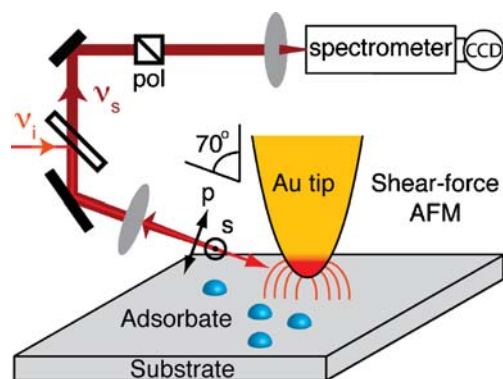

Fig. 2 Illustration of side illumination TERS. The incident light is focused onto the tip–sample gap by means of a long working distance objective (NA=0.2–0.5). The backscattered light is collected through the same objective; the laser line is spectrally filtered, and the signal is analyzed using a spectrograph. With independent polarization control of the incident and scattered light, the incident polarization is generally set to p polarization to drive the axial tip plasmon

in the focus for efficient plasmon excitation requires the use of specialized beam profiles, e.g., radial polarization [46] or a Hermite–Gaussian (1, 0) mode [47]. Yet, even then, the transverse electric field remains significantly stronger than the longitudinal component [47], thus reducing the excitation efficiency. While this geometry does allow for the use of metallic substrates if sufficiently thin (less than the optical skin depth δ), this gives rise to additional loss as Raman light has to pass through the metal, and the coupling enhancement is reduced for thicknesses below δ.

In general, the collection efficiency in TERS using oil-immersion objectives further benefits from the spatial emission characteristics of emitters placed on a dielectric substrate [48]. The emission profile of an emitter placed on a dielectric substrate will exhibit an anisotropic emission pattern (i.e., deviate from that of a perfect dipole). In particular, increased emission will be observed into the optically denser medium, with a significant portion of the total emission refracted into the substrate at an angle greater than the critical angle, allowing for up to 80% of the emitted light to be collected by an oil-immersion objective, with details depending on dipole orientation. While the effect of the dielectric substrate on the tip-enhanced near-field emission has not yet been explored, we expect the preferential emission into the denser medium to be beneficial for TERS detection efficiency.

Alternatively, parabolic mirrors with the tip positioned at their center have been used for top sample illumination using radially polarized pump light to achieve an axial polarization excitation at the tip [49]. This allows for values up to NA=1 [50, 51], while still maintaining the capability for probing nontransparent samples.

In all implementations, the tip–sample distance is controlled using either an atomic force microscope operated with shear force feedback [50, 52], contact mode [11, 31, 53], tapping mode [28, 29], or through a scanning tunneling microscope (STM) [1, 3, 4]. All of these except tapping mode maintain the tip–sample separation at very short length scales (shear force ~10 nm, contact mode <1 nm, STM ~1 nm) as required for maximum time-averaged enhancement.

Raman scattering cross section and signal strength

The strength of the Raman response is characterized by the Raman scattering cross section σ for a given Raman-active normal mode. The differential scattering cross section is then $d\sigma/d\Omega = N_s/N_i$, where N_i is the incident photon flux, and N_s is the number of photons scattered into a solid angle $d\Omega$. The number of detected photons S is then given by $S = \Gamma N_i d\sigma/d\Omega$, where Γ is the overall detection efficiency of

the experiment. In extending this description to the tip-enhanced near field, we can account for the field enhancement through the use of an effective photon flux $N_i' = N_i L_i^2 L_s^2$.

The near-field Raman signal can then be written as:

$$S = \Gamma N_i L_i^2 L_s^2 \cdot 4\pi d\sigma/d\Omega \tag{1}$$

where N_i can readily be calculated from the incident pump power, incidence angle ϕ, and the focus diameter d approximated by the diameter of the Airy disk $d = 1.22\lambda/$NA [43]. The factor of 4π accounts for the emission into all space. Γ will depend sensitively on the particular detection system used. We describe Γ via $\Gamma_{coll}\Gamma_{det}$, describing the collection and detection efficiencies, respectively. The collection efficiency Γ_{coll} expressed in terms of the solid angle over which light is collected depends on the NA used as:

$$\Gamma_{coll} = \frac{1}{2}\left(1 - \sqrt{1 - \frac{NA}{n_i}}\right) \tag{2}$$

which follows from $NA = n_i \sin(\theta)$, where n_i is the index of refraction surrounding the objective (air or immersion oil), and θ is the semiangle of the collection cone.

The spatial emission profile and therefore the collection efficiency depends on the orientation of the Raman emitter with respect to the surface, type of emitter (e.g., dipole or quadrupole), surface polarization density, and molecule–substrate interactions [54]. While these factors can readily be included for specific sample materials, for the general discussion in this work, we assume isotropic emission for both top/side-on illumination and axial illumination using nonimmersion objectives due to random orientation of the molecules or molecular motion.

The detection efficiency $\Gamma_{det} = \Gamma_{trans}\Gamma_{spec}$ depends on the optics used for directing the near-field scattered light to the detector Γ_{trans}, in addition to the detector efficiency itself Γ_{spec}. For the objective, we assume a transmission coefficient of 80% of the scattered light. We assume an additional combined 50% loss from mirrors, polarizing optics, dichroic beam splitter, and possible fiber coupling and transmission. For the spectrometer detection efficiency Γ_{spec}, we discuss the case of a grating spectrometer as is commonly used in TERS experiments and assume a typical 50% throughput and 50% quantum efficiency of the charge-coupled device (CCD) array. The spectral dispersion of the spectrometer will further reduce the signal obtained: the photons scattered by a given Raman-active mode will be distributed over different pixels of the CCD array both vertically and horizontally depending on the spectral resolution desired. For a Raman peak with a Lorentzian line shape with full-width half-maximum (FWHM)= 10 cm^{-1} and a resolution setting of 1 cm^{-1}, each pixel

detects a spectral range of 1 cm^{-1}, and the pixel corresponding to the peak intensity will detect only ~6.4% of the total signal. Taking into account losses and dispersion of the spectrum across multiple pixels, we obtain a detection efficiency order of only ~1% at the pixel of highest intensity for a given mode, i.e., with the specific values from above Γ_{det}=0.64%. We note that our estimates of losses are all chosen at the lower end, representing an optimized experimental setup.

The limiting factors in signal detection are then given by the signal-to-noise ratio arising from detector noise, background light, and far-field signal. For the case of a completely background-free signal with a typical detector noise on the order of two counts per second per pixel, five detected TERS photons/s will yield a sufficient signal-to-noise ratio for coarse signal detection and spectral assignment. In realistic cases, however, a far-field signal and some background stray light will accompany the TERS signal. It has been shown to be beneficial in some cases to use spectrally integrated acquisition [10, 12] with, e.g., an avalanche photodiode. This may result in improved detection efficiency and prevent dispersion of the signal across multiple pixels, increasing the signal-to-noise ratio. The use of longer acquisition times also allows for improving the signal-to-noise ratio but is limited by sample drift and molecular bleaching.

TERS signal: molecular systems

Since the first TERS experiments [5, 34], much of the interest has been centered on the spectroscopy of molecules using resonant or near-resonant excitation in order to increase the Raman signal which can be obtained. Here, we will specifically discuss the signal levels which can be achieved in molecular TERS, in particular, establishing detection limits for systems exhibiting smaller and nonresonant Raman scattering cross sections.

Raman cross sections range from $\leq 10^{-30}$ cm^2 sr^{-1} up to ~10^{-24} cm^2 sr^{-1} per molecule for resonant excitation [55–57]. We begin by examining signal levels of resonant Raman of molecules in a side-on illumination geometry that has been shown to lead to single-molecule sensitivity [3, 4, 7]. As experimental parameters, we consider an NA=0.35 objective, oriented 60° from the surface normal with an incident power of 0.2 mW and λ_{inc}=633 nm. This corresponds to an incident photon flux of 8.3×10^{21} s^{-1} cm^{-2} in a focus of $A \approx 7.6$ µm^2. We assume a Raman enhancement $\left(L_s' L_i'\right)^2 = \sim 10^7$ and $F_i = F_s = 2$ for a Au substrate.

For our analyte, we use a single resonant molecule with $d\sigma/d\Omega = 10^{-26}$ cm^2 sr^{-1} for a particular Raman-active vibra-

tional mode [55]. Based on Eq. 1, this results in 1.3×10^4 s^{-1} sr^{-1} near-field scattered photons. Taking into account $\Gamma = \Gamma_{coll} \Gamma_{trans} \Gamma_{spec}$ and assuming a Raman-active mode with FWHM=10 cm^{-1} and spectral resolution of 1 cm^{-1}, this will result in a readily detectable signal of ~34 counts per second at the pixel of maximum intensity in the CCD array for this vibrational mode.

In order to test the validity of our approximations, we compare our calculated values for the single-molecule case with experiments. In the work of Zhang et al., a signal of ~8 photons per second was detected in side-on illumination under similar experimental conditions [3], i.e., somewhat lower than our estimated values, with the discrepancy presumably due to differences in the Raman cross section of brilliant cresyl blue and details in experimental parameters. Likewise, in previous work by our group [7, 42], a signal of comparable strength was obtained, with 170–230 counts per second observed for a single mode (integrated over ~20 pixels), i.e., ~10 counts per second per pixel.

We now contrast the case of side-on single-molecule TERS with axial illumination using a dielectric substrate. In these experiments, we model common experimental conditions with an objective of NA=1.4, n_i=1.5, λ=488 nm, incident power of 0.2 mW, and a Ag tip with a Raman enhancement $\left(L_s' L_i'\right)^2 \approx 10^5$. In the focus area of $A \approx 0.15$ µm^2, this results in a physical photon flux of 4×10^{23} s^{-1} cm^{-2}. For the same resonant molecule as in the side-on illumination geometry with a cross section of $d\sigma/d\Omega = 10^{-26}$ cm^2 sr^{-1} and assuming Γ_{coll}=0.7 [48], of the 400 photons per second per steradian for a given vibrational mode, ~24 counts per second would be detected by the CCD pixel showing maximum intensity. Although weaker than the side-on illumination case discussed above, this shows that single-molecule sensitivity is possible in an axial illumination geometry.

We now turn to the effect of the background signal. For the case of single-molecule TERS, we require sufficiently dilute surface coverage to find on average only one molecule under the tip-enhanced signal region. With an average coverage of one molecule per ~100 nm^2, we have 7.6×10^4 molecules in our side-on illumination far-field focus of 7.6 µm^2. Therefore, with a resonant molecule with $d\sigma/d\Omega = 10^{-26}$ cm^2 sr^{-1}, we expect 0.25 counts per second of far-field signal at the pixel of maximum intensity, i.e., smaller than our near-field signal by a factor of >100 and well below the detection limit. In increasing the surface density, the ratio of molecules within the far-field focus and in the tip-enhanced near-field region will remain constant, thus maintaining this contrast level. As such, the background signal is not expected to be a limiting factor in TERS from molecular monolayer or submonolayer coverage on a metal substrate. This is in contrast to dense samples, where the molecular layer reaches thicknesses

greater than the spatial extent of the tip near field, and the background signal could become significant.

A comparison of signal levels in top/side-on illumination and axial illumination geometries is presented in Fig. 3, showing the calculated signal achieved for a single molecule at the detector pixel of maximum intensity as a function of NA assuming FWHM=10 cm^{-1} and a spectral resolution of 1 cm^{-1}. In all cases, we assume identical parameters as in the previously discussed cases, i.e., incident power of 0.2 mW, incident wavelengths of $\lambda=$ 488 nm and $\lambda=$633 nm for axial and top/side-on illumination, respectively. For the axial illumination objective, we have assumed, $n_{substrate}=n_{oil}=1.5$, resulting in an emission profile with 40% and 35% emitted into the substrate below and above the critical angle, respectively. The resulting collection efficiency Γ_{coll} under axial illumination based on [48] is shown in the inset, where the vertical dashed line indicates collection at the critical angle of the substrate–air interface.

For the case of top/side-on illumination using a metallic substrate allowing for plasmonic coupling, we see that a large signal can be obtained for molecules with large scattering cross sections of $d\sigma/d\Omega=10^{-26}$ cm^2 sr^{-1}. With a maximum of NA=0.7 up to ~610 counts/s would be expected. For an experimentally typical value of NA=0.5, this still allows for the detection of 150 counts per second.
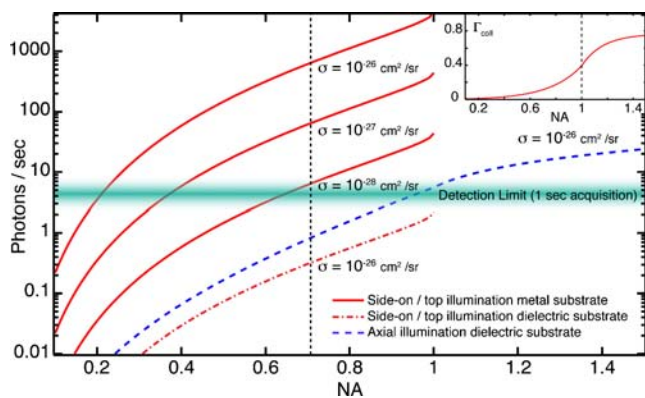

Fig. 3 Number of detected photons on a detector pixel as a function of the objective NA. *Red curves* are in a side-on/top illumination geometry with and without the use of a metallic substrate as indicated. With the dielectric substrate, we assume a Raman enhancement of $(L_s L_i)^2 \approx 10^5$. The *dashed line* indicates the maximum NA achievable in side-on illumination without the use of a parabolic mirror. The *blue curve* is under axial illumination ($n_{substrate}=n_{oil}=1.5$). The collection efficiency (*inset*) of the oil-immersion lens under axial illumination is based on results from [48]. The increased emission of light into the denser medium at angles greater than the critical angle is highly beneficial to TERS measurements. *Thick green line* represents the detection limit for 1 s acquisition time. It can be seen that side illumination benefits significantly from the increased field enhancement that arises due to plasmonic coupling to the metallic substrate. In contrast, the emission pattern of an emitter placed on a dielectric substrate significantly increases collection efficiency in an axial geometry when using oil-immersion objectives

Through the use of a top illumination parabolic mirror of NA=0.95, this allows for the detection of up to 2.7×10^3 counts per second. Therefore, with sufficiently high NA, this geometry even allows for the detection of single molecules with smaller scattering cross sections of $d\sigma/d\Omega=$ 10^{-27} cm^2 sr^{-1} with 15 counts per second expected for NA= 0.5. In the absence of plasmonic tip–sample coupling, however, we see that the detected signal, even for a molecule with a large scattering cross section, remains small, and single-molecule detection is likely not possible.

For the case of axial illumination, we see that the additional emission into the substrate at angles greater than the critical angle is beneficial for the TERS signal. In particular, the additional signal collected for NA>1 allows for single-molecule detection for the case of $d\sigma/d\Omega=10^{-26}$ cm^2 sr^{-1}. However, single molecules with smaller scattering cross sections will be difficult or impossible to detect.

Although top/side-on illumination can yield larger signal values, there are issues regarding its analytical implementation for the identification of molecules on a surface. The large field in the tip–sample gap has been associated with rapid photodecomposition of analytes [3, 4, 7, 42], and spectral diffusion may complicate molecular identification. As axial illumination experiences a reduced gap field, which is partially compensated for through the use of high NA, this may be desirable to reduce decomposition.

Furthermore, in probing larger structures (e.g., biological [30]), their size in the tens of nanometer range may prevent the metallic substrate-induced field enhancement through the tip–sample coupling from being effective. Thus, it may not always be possible to utilize a top/side-on illumination geometry to obtain intense TERS signals, and axial illumination may be preferable.

For comparison, we consider the case of a biological molecule, studied under axial illumination, using a value of $d\sigma/d\Omega=10^{-28}$ cm^2 sr^{-1} and otherwise identical parameters as above ($\lambda=$488 nm, $n_i=1.5$, NA=1.4). This would result in a very small signal of 0.2 photons per second at the detector pixel of highest signal intensity. While this signal level effectively renders a single-molecule undetectable, it needs to be considered that, e.g., for a single RNA chain [29], multiple nucleotide bases are present under the tip (on the order of 30), increasing the local oscillator density. Also, longer acquisition times can be used (up to 30 s [32]), allowing for an increased signal [29, 30, 32, 58]. Considering scattering from 15 nucleotide bases with 30-s acquisition, we expect a robust signal of ~110 photons in the presence of ~20 dark counts.

Therefore, although single-molecule sensitivity can be achieved in TERS, the application of the technique to molecular systems with smaller Raman scattering cross sections may not allow for the detection of single molecules. In particular, utilizing axial illumination to

probe, e.g., biological molecules requires the probing of small ensembles in combination with longer acquisition times. However, with ongoing improvements in tip fabrication [59, 60], larger tip enhancement and higher signal strength may be possible.

Thus, for systems which can be studied in a top/side-on illumination geometry on a metallic substrate (e.g., small molecules), the ultrahigh sensitivity of this implementation can be utilized to its full degree. In contrast, while lower signal levels are achievable in the axial geometry, the increased versatility requires its use in some situations. Despite the reduced signal levels, axial illumination is suitable for probing small ensembles of molecules, even those exhibiting reduced scattering cross sections [29–32].

Crystalline systems

Through the capability to interact directly with zone-center optical phonons [61], Raman scattering has long been a powerful tool for studying crystalline solids [62]. Accordingly, there have been initial studies toward applying TERS to the study of crystalline materials. The majority of studies have focused on Si [16, 17, 24, 53, 63], driven by the desire to image nanoscale variations in strain. However, studies of other crystalline materials have been limited [12, 13, 15, 24]. This can be attributed to the fact that many crystalline materials have small Raman scattering cross sections. For crystalline structures, the differential Raman scattering cross section is given as a volume-normalized value [61] with values ranging from $10^{-9} \mathrm{cm}^{-1} \mathrm{sr}^{-1}$ up to $10^{-5} \mathrm{cm}^{-1} \mathrm{sr}^{-1}$ [64, 65]. As phonon Raman is a volume-based effect, for a scattering volume of $1 \mathrm{nm}^3$, comparable to a single dye molecule, we obtain cross sections ranging from $10^{-30} \mathrm{cm}^2 \mathrm{sr}^{-1}$ up to $10^{-26} \mathrm{cm}^2 \mathrm{sr}^{-1}$, smaller than corresponding molecular values.

We begin our discussion of phonon TERS by considering the case of Si which has a comparably large scattering cross section of $\sim 5 \times 10^{-6} \mathrm{cm}^{-1} \mathrm{sr}^{-1}$ [65]. With the spatial extent of the near field given by the tip apex radius, we estimate an effective near-field probing volume of $10^3 \mathrm{nm}^3$. Thus, we obtain an effective near-field Raman scattering cross section of $5 \times 10^{-24} \mathrm{cm}^2 \mathrm{sr}^{-1}$. For our experimental geometry, we consider an identical top/side-on illumination geometry as discussed above, necessitated by the fact that Si is opaque. In keeping with experimental values, we assume a Raman enhancement factor of 10^4 [53] with a Fresnel factor of $F=1.1$ in the visible spectral range.

Unlike molecular systems, bleaching is not expected to be a concern, so we consider incident power of 1 mW, which results in 4.2×10^{22} photons per square centimeter in a focus of $A \approx 7.6 \mu m^2$. From Eq. 1, we then expect $\sim 3 \times 10^3 \mathrm{sr}^{-1} \mathrm{s}^{-1}$ near-field scattered photons, of which $\sim 8 \mathrm{s}^{-1}$ will be detected by the pixel of maximum

signal intensity. While this signal is weaker than the top/side-on illumination single-molecule case discussed above, the lack of photobleaching allows for longer acquisition times (limited by sample drift), with upwards of 30 s still practical [63]. As a consequence, an appreciable near-field signal can be obtained from Si crystals.

However, the far-field signal background is a significant concern [63]. With a light penetration depth of ~5 μm at a wavelength of $\lambda=633$ nm [66], a bulk volume of $7.6 \mu m^2 \times 5 \mu m$ would contribute to the far-field Raman signal. Therefore, $\sim 2 \times 10^4$ photons per second would be detected at the center pixel of the peak. Such a high far-field signal would make a near-field signal effectively indiscernible. Fortunately, experimentally observed near- to far-field contrast is significantly higher [17, 63], owing in part to the absorption of far-field Raman light within the crystal and losses at the Si–air interface. Furthermore, the near-field probe volume is likely larger than what has been assumed with the tip apex as a lower limit [3, 25]. The use of shorter wavelengths (e.g., $\lambda=532$ nm), resulting in a reduced absorption depth [66], has also been beneficial [17, 53]. Thus, experimentally obtained near- to far-field contrast values are typically in the range of 1:10.

Further reducing or eliminating the background has been a significant challenge, and of particular interest has been the use of tip depolarization [45]. It has been shown that the tip-enhanced near field may alter the polarization state of the TERS signal, that is, a Raman mode which is expected to be observed under a given far-field polarization condition will also have a polarization component in a forbidden configuration. As a consequence, setting a polarizer to an angle which significantly suppresses the far-field background will still allow for an appreciable near-field signal [16, 17, 24, 67]. While this effect has been applied successfully for imaging strain in Si, in crystal classes more complex than cubic Si, there remains a desire for the capability to utilize polarization control for isolating specific phonon modes rather than just its use for background suppression [68].

Despite these challenges posed by the far-field background, phonon TERS is well suited for studying crystalline nanostructures or thin films, as the background signal would be significantly reduced. In recent work by Matsui et al. [15], TERS on GaN was demonstrated with a Raman enhancement of 2.8×10^4. With a Raman scattering cross section of $\sim 3 \times 10^{-7} \mathrm{cm}^{-1} \mathrm{sr}^{-1}$ [69], this is 20 times smaller than the corresponding Si value. Accordingly, the near-field signal was observed to be weak, requiring acquisition times of 10 min. However, through the use of a thin film, the effective far-field scattering volume is reduced, significantly improving the near- to far-field contrast. Likewise, results on CdS, which has a comparable cross section to GaN, have been reported using thin films [24].

The study of crystalline nanostructures also allows for the use of metallic substrates to obtain enhancement through the Fresnel factor. Our group has demonstrated this recently, performing phonon TERS measurements from $BaTiO_3$ and $LiNbO_3$ nanocrystals [12, 13]. In particular, we have utilized the symmetry selectivity of the Raman selection rules in combination with the tip enhancement polarization anisotropy to perform optical nanocrystallography.

Conclusion

In recent years, TERS has been proposed to study a variety of systems, and work is underway to implement TERS as an analytical technique capable of detecting small amounts of analyte with ultrahigh spatial resolution. Here, we have addressed fundamental questions of the signal levels achievable in TERS through a semiquantitative analysis of the expected signal levels in different experimental geometries with different sample materials. Given constraints on the achievable TERS signal as imposed by the scattering cross section, degree of field enhancement, and NA of the illumination and detection optics, we find that even single-molecule sensitivity may be routinely achieved in a top/side-on illumination geometry through the additional field enhancement obtained from plasmonic tip–sample coupling. Even for cases where the additional field enhancement is unachievable, the use of an axial illumination geometry with high NA optics can yield sufficient signal for many analytical implementations. However, the overall signal strength obtainable may preclude rapid scanning or spectrographic imaging with high spectral resolution.

In particular, we considered the effect of the NA of the illumination and detection objective and detection efficiency including the spectrometer, field enhancement, and Raman scattering cross section of the analyte. We find that the high field enhancement achievable in plasmonic tip–sample coupling in a top/side-on illumination geometry outweighs the benefits of the high NA which can be used for axial illumination with the achievable signal higher by a factor of >5. However, in many cases, the extra enhancement in top/side-on illumination may not be accessible or pose difficulties in chemical identification, thus requiring the use of axial illumination. In such cases, we find axial illumination suitable for the study of molecules with large scattering cross sections or ensembles of molecules with smaller scattering cross sections.

For crystalline systems, we find the far-field background and small Raman scattering cross sections to be a significant impediment to the study of bulk crystal systems. However, with a reduced background, TERS has large potential for the study of crystalline thin films and nanocrystals. Our results show that, while restrictions exist in the systems which can be studied with TERS, the technique carries strong potential for the study of a variety of solids.

References

1. Pettinger B, Ren B, Picardi G, Schuster R, Ertl G (2004) Phys Rev Lett 92:096101
2. Anderson MS (2000) Appl Phys Lett 76:3130
3. Zhang W, Yeo BS, Schmid T, Zenobi R (2007) J Phys Chem C 111:1733
4. Domke KF, Zhang D, Pettinger B (2006) J Am Chem Soc 128:14721
5. Hayazawa N, Inouye Y, Sekkat Z, Kawata S (2000) Opt Commun 183:333
6. Steidtner J, Pettinger B (2008) Phys Rev Lett 100:236101
7. Neacsu CC, Dreyer J, Behr N, Raschke MB (2006) Phys Rev B 73:193406
8. Yeo BS, Amstad E, Schmid T, Stadler J, Zenobi R (2009) Small 8:952
9. Hartschuh A, Qian H, Georgi C, Böhmer M, Novotny L (2009) Anal Bioanal Chem 394:1787
10. Hartschuh A, Sanchez EJ, Xie XS, Novotny L (2003) Phys Rev Lett 90:095503
11. Yano T-A, Inouye Y, Kawata S (2006) Nano Lett 6:1269
12. Berweger S, Neacsu CC, Mao Y, Zhou H, Wong SS, Raschke MB (2009) Nat Nanotechnol 4:496
13. Berweger S, Raschke MB (2009) J Raman Spectrosc (in press)
14. Lee N, Hartschuh RD, Mehtani D, Kisliuk A, Maguire JF, Green M, Foster MD, Sokolov AP (2007) J Raman Spectrosc 38:789
15. Matsui R, Verma P, Ichimura T, Inouye Y, Kawata S (2007) Appl Phys Lett 90:061906
16. Motahashi M, Hayazawa N, Tarun A, Kawata S (2008) J Appl Phys 103:034309
17. Ossikovski R, Nguyen Q, Picardi G (2007) Phys Rev B 75:045412
18. Neacsu CC, Steudle GA, Raschke MB (2005) Appl Phys B 80:295
19. Behr N, Raschke MB (2008) J Phys Chem C 112:3766
20. Demming AL, Festy F, Richards D (2005) J Chem Phys 122:184716
21. Denk W, Pohl DW (1990) J Vac Sci Tech B 9:510
22. Novotny L, Bian RX, Xie XS (1997) Phys Rev Lett 79:645
23. Roth RM, Panoiu NC, Adams MM, Osgood RM, Neacsu CC, Raschke MB (2006) Opt Express 14:2921
24. Mehtani D, Lee N, Hartschuh RD, Kisliuk A, Foster MD, Sokolov AP, Maguire JF (2005) J Raman Spectrosc 36:1068
25. Ren B, Picardi G, Pettinger B (2004) Rev Sci Instr 75:837
26. Saito Y, Murakami T, Inouye Y, Kawata S (2005) Chem Lett 34:920
27. Hartschuh A (2008) Angew Chem Int Ed 47:8178
28. Ichimura T, Fujii S, Verma P, Yano T, Inouye Y, Kawata S (2009) Phys Rev Lett 102:186101

29. Bailo E, Deckert V (2008) Angew Chem Int Ed 47:1658
30. Cialla D, Deckert-Gaudig T, Budich C, Laue M, Möller R, Naumann D, Deckert V, Popp J (2009) J Raman Spectrosc 40:240
31. Schmid T, Messmer A, Yeo BS, Zhang W, Zenobi R (2008) Anal Bioanal Chem 391:1907
32. Yeo BS, Mädler S, Schmid T, Zhang W, Zenobi R (2008) J Phys Chem C 112:4867
33. Yeo BS, Stadler J, Schmid T, Zenobi R, Zhang W (2009) Chem Phys Lett 472:1
34. Stockle RM, Suh YD, Deckert V, Zenobi R (2000) Chem Phys Lett 318:131
35. Chen CK, Heinz TF, Ricard D, Shen YR (1983) Phys Rev B 27:1965
36. Le Ru EC, Grand J, Felidj N, Aubard J, Levi G, Hohenau A, Krenn JR, Blackie E, Etchegoin PG (2008) J Phys Chem C 112:8117
37. Moskovits M (1985) Rev Mod Phys 57:783
38. Aigouy L, Lahrech A, Gressilon S, Cory H, Boccara AC, Rivoal JC (1999) Opt Lett 24:187
39. Roy D, Wang J, Williams C (2009) J Appl Phys 105:013530
40. Micic M, Klymyshyn N, Suh YD, Lu HP (2003) J Phys Chem B 107:1574
41. Pettinger B, Domke KF, Zhang D, Schuster R, Ertl G (2007) Phys Rev B 76:113409
42. Neacsu CC, Dreyer J, Behr N, Raschke MB (2007) Phys Rev B 75:236402
43. Born M, Wolf E (1999) Principles of optics. Cambridge University Press, Cambridge
44. Boyd RW (2003) Nonlinear optics. Elsevier, Amsterdam
45. Poborchii V, Tada T, Kanayama T (2005) Jap J Appl Phys 44:202
46. Hayazawa N, Saito Y, Kawata S (2004) Appl Phys Lett 85:6239
47. Novotny L, Sanchez EJ, Xie XS (1998) Ultramicroscopy 71:21
48. Enderlein J, Ruckstuhl T, Seeger S (1999) Appl Opt 38:724
49. Debus C, Lieb MA, Drechsler A, Meixner AJ (2002) J Microsc 210:203
50. Sackrow M, Stanciu C, Lieb MA, Meixner AJ (2008) Chem Phys Chem 9:316
51. Steidtner J, Pettinger B (2007) Rev Sci Instr 78:103104
52. Neacsu CC, Berweger S, Raschke MB (2007) NanoBiotechnol 3:172
53. Hayazawa N, Motohashi M, Saito Y, Ishitobi H, Ono A, Ichimura T, Verma P, Kawata S (2007) J Raman Spectrosc 38:684
54. Enderlein J, Böhmer M (2003) Opt Lett 28:941
55. Le Ru EC, Blackie E, Meyer M, Etchegoin PG (2007) J Phys Chem C 111:13794
56. Shim S, Stuart CM, Mathies RA (2008) Chem Phys Chem 9:697
57. Skinner JG, Nilsen WG (1968) J Opt Soc Am 58:113
58. Rasmussen A, Deckert V (2006) J Raman Spectrosc 37:311
59. Barrios CA, Malkovskiy AV, Kisliuk AM, Sokolov AP, Foster MD (2009) J Phys Chem C 113:8158
60. Taguchi A, Hayazawa N, Saito Y, Ishitobi H, Tarun A, Kawata S (2009) Opt Express 17:6509
61. Kuzmany H (1998) Solid-state spectroscopy. Springer, Berlin
62. Cardona M (ed) (1983) Light scattering in solids I. Springer, Berlin
63. Saito Y, Motohashi M, Hayazawa N, Iyoki M, Kawata SS (2006) Appl Phys Lett 88:143109
64. Johnston WD, Kaminow IP (1968) Phys Rev 168:1045
65. Ralston JM, Chang RK (1970) Phys Rev B 2:1858
66. Palik ED (ed) (2000) Handbook of optical constants of solids. Springer, Berlin
67. Gucciardi PG, Lopes M, Deturche R, Julien C, Barchiesi D, de la Chapelle ML (2008) Nanotechnol 19:215702
68. Hayes W, Loudon R (1978) Scattering of light by crystals. Dover, New York
69. Loa I, Gronemeyer S, Thomsen C, Ambacher O, Schikora D, As DJ (1998) J Raman Spectrosc 29:219