Research article

Journal of RAMAN SPECTROSCOPY

Received: 18 August 2013 Revised: 1 November 2013 Accepted: 6 November 2013 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/jrs.4423

Tip-enhanced Raman spectroscopy – an interlaboratory reproducibility and comparison study

Carolin Blum,[a] Lothar Opilik,[a] Joanna M. Atkin,[b,c,d] Kai Braun,[e] Stefan B. Kämmer,[f†] Vasily Kravtsov,[b,c,d] Naresh Kumar,[g] Sergey Lemeshko,[h] Jian-Feng Li,[a] Karol Luszcz,[f‡] Teimour Maleki,[f] Alfred J. Meixner,[e] Steve Minne,[f] Markus B. Raschke,[b,c,d] Bin Ren,[i,j] Jan Rogalski,[e] Debdulal Roy,[g] Bruno Stephanidis,[a] Xiang Wang,[i,j] Dai Zhang,[e] Jin-Hui Zhong[i,j] and Renato Zenobi[a]*

Since its first experimental realization, tip-enhanced Raman spectroscopy (TERS) has emerged as a potentially powerful nanochemical analysis tool. However, questions about the comparability and reproducibility of TERS data have emerged. This interlaboratory comparison study addresses these issues by bringing together different TERS groups to perform TERS measurements on nominally identical samples. Based on the spectra obtained, the absolute and relative peak positions, number of bands, peak intensity ratios, and comparability to reference Raman and surface-enhanced Raman spectroscopy (SERS) data are discussed. Our general findings are that all research groups obtained similar spectral patterns, irrespective of the setup or tip that was used. The TERS (and SERS) spectra consistently showed fewer bands than the conventional Raman spectrum. When comparing these three methods, the spectral pattern match and substance identification is readily possible. Absolute and relative peak positions of the three major signals of thiophenol scattered by 19 and 9 cm^{-1}, respectively, which can probably be attributed to different spectrometer calibrations. However, within the same group (but between different tips), the signals only scattered by 3 cm^{-1} on average. This study demonstrated the suitability of TERS as an analytical tool and brings TERS a big step forward to becoming a routine technique. Copyright © 2014 John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: tip-enhanced Raman spectroscopy; interlaboratory comparison study; thiophenol self-assembled monolayer; spectra interpretation; metrology

Introduction

Since its experimental realization,[1–4] tip-enhanced Raman spectroscopy (TERS) has evolved remarkably and has become more and more mature. Starting with point measurements in the early days, e.g. on dyes,[4] TERS imaging experiments have become possible in recent years on single molecules,[5] graphene,[6] carbon nanotubes,[7] and even biological samples.[8] With TERS gaining increasing importance as a chemical analysis technique and being commercialized by an increasing number of manufacturers, questions about the comparability and reproducibility of TERS data arise. This calls to clarify whether deviations in TERS spectra obtained by different groups are due to differences in

* *Correspondence to: Renato Zenobi, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland. E-mail: zenobi@org.chem.ethz.ch*

† *Present address: Multiprobe, 93103 Santa Barbara, CA, USA*

‡ *Present address: Department of Information Technology and Electrical Engineering, Photonics Laboratory, ETH Zurich, 8093 Zurich, Switzerland*

a *Department of Chemistry and Applied Biosciences, Laboratory of Organic Chemistry, ETH Zurich, 8093 Zurich, Switzerland*

b *Department of Physics, University of Colorado, 80309 Boulder, CO, USA*

c *Department of Chemistry, University of Colorado, 80309 Boulder, CO, USA*

d *JILA, University of Colorado, 80309 Boulder, CO, USA*

e *Institute of Physics and Theoretical Chemistry, University Tübingen, 72076 Tübingen, Germany*

f *Bruker Nano Surfaces Division, 93117 Santa Barbara, CA, USA*

g *National Physical Laboratory, TW11 OLW Teddington, UK*

h *NT-MDT Europe B.V. Application Laboratory, 5656 Eindhoven, The Netherlands*

i *State Key Laboratory of Physical Chemistry of Solid Surfaces, College of Chemistry and Chemical Engineering, Xiamen University, 361005 Xiamen, China*

j *Key Laboratory of Analytical Sciences, College of Chemistry and Chemical Engineering, Xiamen University, 361005 Xiamen, China*

Copyright © 2014 John Wiley & Sons, Ltd.

samples, differences in TERS instruments, or differences in measurement procedures applied.

A TERS experiment is characterized by the substrate, the tip (full metal or metallized), and the illumination/detection geometry. Typical substrates include the following: glass slides,[2,9] mica,[9,10] template-stripped (TS) Au substrates,[5,11] Au platelets,[6,12] or single crystalline Au surfaces.[7,13] TERS tips are produced in many different ways, e.g. by electrochemical etching from a full metal wire,[8,14,15] by metal coating of an atomic force microscopy (AFM) cantilever,[2,9] or by more complicated procedures such as focused ion beam (FIB) milling (e.g. for grating coupled tips).[16] Ag and Au are most commonly used as tip materials in combination with excitation wavelengths in the green or red part of the visible spectrum. Setups based on AFM or scanning tunneling microscopy (STM) feedback mechanisms are in use. Another big difference in the TERS setups is the illumination/detection geometry applied: transmission illumination/detection,[17] top illumination/detection,[18] side illumination/detection (all using a microscope objective), or illumination/detection with a parabolic mirror[19] are most commonly used (Fig. 1). With such a diversity of different experimental parameters, it is not surprising that the results obtained are sometimes different. Examples include discrepancies in TERS spectra (in terms of band pattern and signal positions) recorded in different laboratories from malachite green,[20–22] phenylalanine,[23,24] and 4-nitrobenzenethiol.[17,25,26] It is therefore important to clarify what spectral differences are to be expected because of the experimental differences mentioned earlier.

Round robin or interlaboratory comparison studies have been carried out since analytical sciences exist to achieve reproducibility and comparability between different laboratories and research groups. They are very powerful means to validate a method and to show its potential as an analytical tool. Various examples can be found in the literature, e.g. a round robin study about X-ray photoelectron spectroscopy and Auger-electron microscopy,[27,28] the international evaluation program (IMEP) with several rounds, e.g. IMEP-9 on trace elements in water,[29] an intercomparison study on accurate mass measurements of small molecules,[30] or an interlaboratory study on protein glycosylation by mass spectrometry.[31] After more than 10 years of TERS development and research, one might ask whether it will soon reach the metrological level. At present, TERS still has some way to go until it is possible to directly compare TERS spectra obtained in different laboratories on different setups and to determine the efficiency and accuracy of a TERS setup by means of a standardized procedure (e.g. with a standard sample). In this study, we carry out an interlaboratory study to compare TERS spectra obtained by different groups using different TERS setups on equivalent samples.

As samples, self-assembled monolayers of thiophenol on two kinds of gold substrates – a transparent one for transmission illumination/detection and an opaque one for top/side illumination/detection – were sent to the participating groups to be studied using their TERS setups. They performed TERS measurements on the sample with different tips and sent the raw data back to the organizing group (ETH Zürich). If the measurements on the thiophenol sample were successful, a second sample of undisclosed nature followed. Only the organizing group knew the composition of the unknown sample (a tripeptide, CysPhePhe).

The aim of this interlaboratory comparison was to address the following questions: Is the same spectral pattern (relative peak

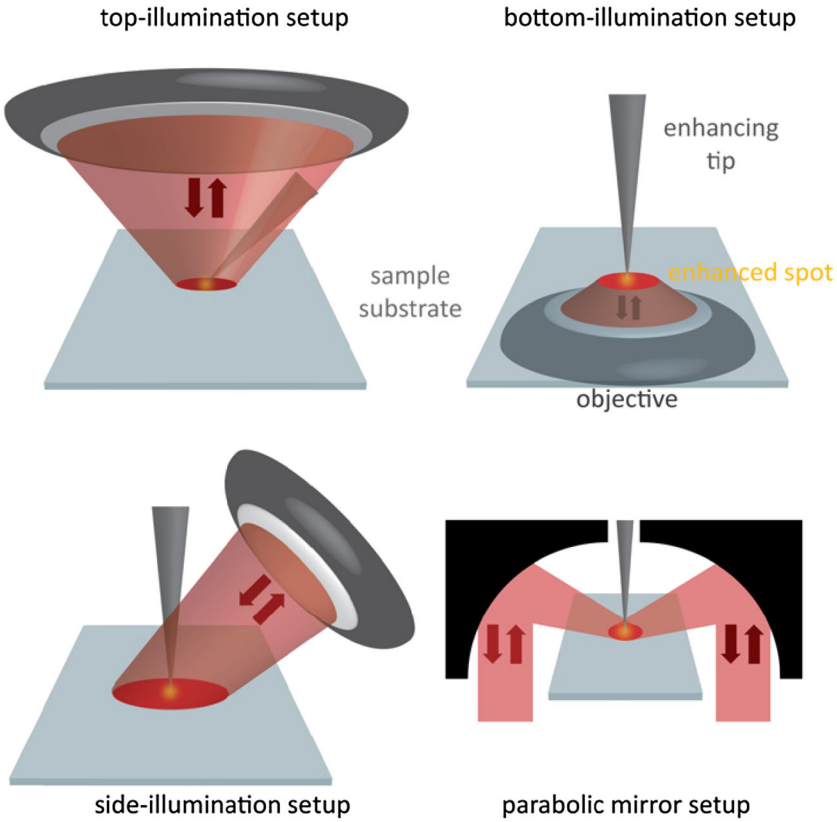

Figure 1. Schematics of different types of setups used in this study. Illumination and collection of scattered light occurs via the same optics in all four cases.

Copyright © 2014 John Wiley & Sons, Ltd.

positions) observed when measuring equivalent samples? Do the absolute peak positions vary and if yes by how much? Do the peak intensity ratios differ? How many bands are observed compared with conventional Raman spectroscopy and SERS? Are there unique TERS selection rules? What has to be considered when comparing TERS spectra obtained by different groups/setups? Is there a systematic difference in the TERS spectra that can be traced back to differences in the illumination/detection geometry, tip type, or tip-sample distance control feedback mechanism? Can the differences/similarities that were observed for the thiophenol sample also be observed for a more fragile tripeptide (used as the 'unknown' sample)? And finally, can TERS be used as a reliable analytical tool to identify unknown substances? This study shows that all groups observed a very similar spectral pattern although using various kinds of setups and tips. The absolute and relative signal positions were within 3 cm^{-1} for TERS spectra recorded by same group. However, absolute and relative peak positions scattered by 19 and 9 cm^{-1} respectively, which can probably be attributed to different spectrometer calibrations. In addition, the same signals are visible in the TERS and shell-isolated nanoparticle-enhanced Raman spectroscopy (SHINERS) spectra of thiophenol.[24] The differences between the SERS/TERS and conventional Raman spectra can be rationalized with density functional theory (DFT) calculations performed by Feugmo *et al.*[32]

Experimental

Because of the different requirements for the involved TERS setups, substrates had to be designed accordingly. Some setups could only accommodate certain sample sizes, and some required transparent substrates. The used substrates can be divided in two main groups: opaque TS gold substrates and semi-transparent gold-coated glass slides. Self-assembled monolayers were chosen as samples, because they can be prepared in a reproducible manner. For the known sample, thiophenol was chosen, because it has a strong Raman response and it readily self-assembles on gold.[33,34] For the unknown sample, the tripeptide CysPhePhe was chosen because of several reasons: (1) cysteine is known to self-assemble on gold,[35] (2) while cysteine itself has a small Raman cross section, the two additional phenylalanine residues of CysPhePhe lead to a much larger Raman cross section,[36,37] and (3) there is an ongoing discussion about TERS spectra of biomolecules, especially whether the amide I mode (which is related to the secondary structure of proteins) can be observed.[38,39] The choice of a peptide for this work could potentially address this question.

TERS setups

In Fig. 1, schemes of the different types of TERS setups used in this study are displayed. A detailed description of the different configurations of individual groups can be found in the Supporting Information pages 2–4. The most important experimental parameters are summarized in Table 1. The spectral resolution was calculated from the signal positions of two adjacent pixels (at 997 cm^{-1}). The results of this comparison study were anonymized, and all participating groups were labeled A to G.

TERS tips

A detailed description can be found in the Supporting Information pages 5–6. The most important details are summarized in Table 2.

Samples

All samples were prepared by the organizing group (Blum *et al.*, ETH Zürich) according to the procedures described in the following. The samples were then packaged and shipped to the participating groups for measurement. Because the participating laboratories are located in different parts of the world and mailing times

Table 1. Overview of the instrument characteristics of all participating groups

Group	A	B	C	D	E	F	G	
Feedback mechanism	STM	AFM (contact)	STM	STM	AFM (shear-force)	STM	STM	AFM (contact)
Tip material	Ag, Au	Ag	Au	Ag	Au	Au	Ag	Ag
Excitation wavelength (nm)	632.8	532.0	636.6	632.8	632.8	632.8	632.8	532.0
Numerical aperture	0.45	1.49	0.998	0.7	0.5	0.42	0.7	1.4
Illumination, angle of incidence	Side, 60°	Transmis.	Top, 15-86°	Top, 0°	Side, 60°	Side, 60°	Top, 0°	Transmis.
Tip angle with respect to the surface	90°	90°	90°	30–50°	90°	90°	30-50°	90°
Spectral resolution (cm^{-1}/pixel)	2.2	2.5	2.2	1.1	5.7	2.2	1.1	2.7
Polarization	Linear, parallel to tip axis	Radial	Radial	Linear, along tip axis	Linear, parallel to tip axis	Linear, parallel to tip axis	Mixed	Radial
Spectrometer calibration (cm^{-1})	Ne lamp, Si, HOPG	Ne lamp, Si	Ne lamp, Si	Si	H_2	Si	Si, diamond, Ne lamp	Ne lamp

Side, side/epi-illumination/detection; transmis., transmission illumination/detection; top, top illumination/detection; HOPG, highly ordered pyrolytic graphite; STM, scanning tunneling microscopy; AFM, atomic force microscopy. The surface normal is the reference for defining the angle of incidence.

Copyright © 2014 John Wiley & Sons, Ltd.

Table 2. Overview of the different tip preparation recipes

Group	A	B	C	D	E	F	G
Feedback mechanism	STM	AFM (contact)	STM	STM	AFM (shear-force)	STM	STM	AFM (contact)
Wire material	Ag, Au	—	Au	Ag	Au	Au	Ag	—
Wire diameter (mm)	0.25	—	0.25	0.25	0.125	0.125	0.25	—
Etching solution	HCl : EtOH	—	HCl	$HClO_4$: MeOH	HCl : EtOH	RX : ROH	$HClO_4$: MeOH	—
coating (nominal thickness)	—	42 nm Ag	—	—	—	—	—	50 nm Ag
Underlying material	—	300 nm SiO_2 (oxidation)	—	—	—	—	—	SiN

STM, scanning tunneling microscopy; AFM, atomic force microscopy.

varied, the samples were measured after different delay times (up to approximately 2 weeks). It had been tested by the organizing group that the same signals were still observed after this time.

TS Au substrates for top/side illumination/detection based setups

The opaque substrates were prepared according to a procedure adapted from Weiss *et al.*[40] that was described in detail before.[24]

Thiophenol samples on TS Au

The glass pieces were lifted off the silicon wafer, exposing a freshly cleaved TS gold surface that was immediately immersed into a 5 mM ethanolic thiophenol (Acros, USA) solution for approximately 18 h in order to allow chemisorption of thiophenol molecules on the gold surface.[33,34] The sample was then rinsed with ethanol and dried in a stream of nitrogen. The organizing group performed TERS measurements (top illumination/detection, Ag tips) with three tips on each sample to check for the presence of the thiophenol monolayer before it was shipped to one of the participating groups in a sealed container.

CysPhePhe samples on TS Au

A freshly cleaved TS Au substrate was immersed in a 1 mg/ml ethanolic CysPhePhe solution (custom synthesized, 99.68% purity; CanPeptide) in a 1.5 ml Eppendorf tube. The tubes with the submersed substrates were then shipped to the participating groups, with instructions to take the sample out of the solution, rinse it with ethanol, and dry it under a stream of nitrogen or argon. The participating groups then used the samples for the TERS experiments. The organizing group did not perform measurements on these samples before shipment.

Substrates for transmission illumination

Glass cover slides (borosilicate D263^{TM}M, 100 μm thickness; Marienfeld, Germany) were cleaned in piranha solution as described earlier. After rinsing them with water (NANOpure) and ethanol, they were vapor coated with an adhesion layer of `titanium (99.7% purity; ABCR, Germany) of 2 nm nominal thickness before coating them with 10 nm of Au (nominal thickness).

Thiophenol samples on semi-transparent Au

The substrates were immersed in a 5 mM thiophenol solution and shipped in a sealed container to the participating groups.

CysPhePhe samples on semi-transparent Au

The substrates were immersed in a 1 mg/ml ethanolic CysPhePhe solution and shipped immersed in the solution in a sealed container to the participating group.

SHINERS particles

For a detailed description of the SHINERS particles[41] preparation, please see the Supporting Information page 7.

Data processing

For each of the nine spectra from each group, the three characteristic thiophenol peaks at ≈998, ≈ 1022, and ≈1074 cm^{-1} (the peak at 1004 cm^{-1} for CysPhePhe) were fitted in order to obtain the peak positions and intensities (peak height; see the Supporting Information pages 9–17 and 35–40). A total of three spectra were selected from each group, with one spectrum for each tip (Supporting Information pages 8 and 36). Finally, in order to compare the spectra from different groups, selected spectra – one per group – were plotted in the same graph (Fig. 2a). The spectra were carefully chosen using the following criteria: high signal-to-noise ratio and absence of carbonaceous contamination signals (exception: the AFM spectrum of group G shows a broad carbonaceous signal background at around 1600 cm^{-1}. It is still displayed, because no other spectrum without that background was available.). All spectra are displayed as received. No additional averaging, background correction etc. were performed. The spectra were offset vertically for better visibility and scaled in their intensity as necessary to display the results from different laboratories in one graph. If a spectrum was scaled in the *y*-direction, it is clearly stated in the figure caption. None of the spectra were shifted or scaled along the *x*-axis.

Results and discussion

The thiophenol sample was initially sent to 12 different groups, operating 6 setups using an STM feedback mechanism and 7 setups using an AFM feedback mechanism (13 setups in total). Reproducible spectra were obtained on 8 among the 13 setups (62%). The CysPhePhe sample was then measured by seven groups who successfully obtained thiophenol TERS spectra (coauthors) in the first round (with five setups with STM feedback mechanism and three setups with AFM feedback mechanism, eight setups in total). Reproducible spectra were obtained on four of these eight setups (50%). Which group was successful on which sample is summarized in Table 3.

Band assignment: Raman *versus* SERS/TERS

Figure 2a shows a comparison of TERS thiophenol spectra from the different groups (see data evaluation). Figure 2b shows a conventional Raman spectrum of thiophenol adapted from Zayak *et al.*[43] In Fig. 2c, a thiophenol TERS spectrum is compared with a thiophenol Raman spectrum enhanced by SHINERS particles. The TERS spectrum of group G was chosen for this comparison because it had the

Copyright © 2014 John Wiley & Sons, Ltd.

Journal of RAMAN SPECTROSCOPY

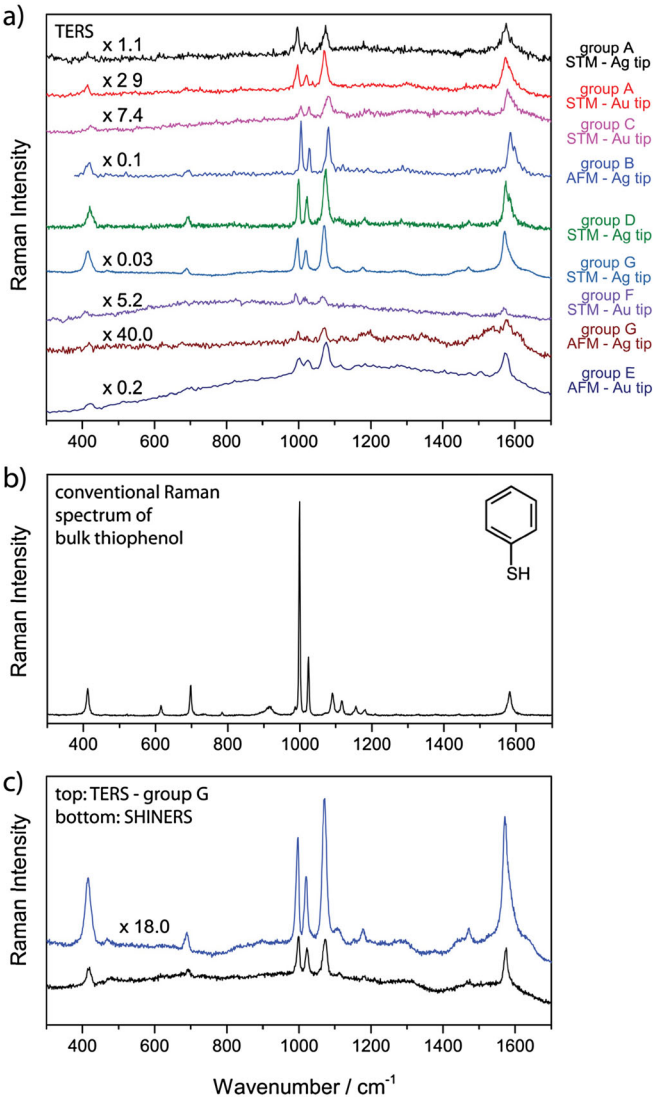

Figure 2. (a) Tip-enhanced Raman spectroscopy (TERS) thiophenol spectra from the different groups. The spectra were multiplied and offset along the *y*-axis in order to show them in one graph. (b) Conventional Raman spectrum of thiophenol (neat solution, excitation wavelength 632.8 nm; the spectrum was measured and kindly provided by Alexey Zayak, Jim Schuck, and Jeaffrey Neaton). (c) Comparison of a TERS spectrum of the thiophenol of group G with a shell-isolated nanoparticle-enhanced Raman spectroscopy (SHINERS) thiophenol spectrum. The TERS spectrum was multiplied and offset along the *y*-axis for better visibility. STM, scanning tunneling microscopy.

highest signal-to-noise ratio and all spectral features are clearly visible. Au SHINERS particles with a silica shell were chosen as enhancing substrates because their underlying SERS effect is similar to TERS and their shell prevents direct binding of the analyte (as well as contaminants) to the enhancing particle.

As can be seen, the absolute and relative peak positions in the thiophenol TERS spectrum match the ones in the thiophenol SHINERS spectrum. Signals at 417, 692, 999, 1023, 1070, 1178, 1473, and 1581 cm^{-1} are all due to aromatic ring vibrational modes.[44] An assignment of the vibrational modes can be found in the works by Scott *et al.*[44] and Feugmo *et al.*[32] These TERS/SERS modes match the signals in the reference Raman spectrum, although some modes show stronger or weaker relative intensities in the SERS/TERS spectra. Feugmo *et al.* claim that the sulfur atom of the thiophenol molecule most likely binds to the gold in a twofold-coordinated binding site, leading to the observed differences between the conventional Raman and the SERS (and TERS) spectra: The two most abundant peaks in the reference Raman spectrum are at 999 (ring in-plane deformation mode r-i-d + CC stretching mode ν_{CC}) and 1022 cm^{-1} (ν_{CC} + CH deformation mode δ_{CH}), whereas the most abundant signals in the TERS spectrum are the triplet of peaks at 998, 1022, and 1074 cm^{-1} ($\nu_{CC} + \delta_{CH}$). Additionally, the signals at 417 ($\nu_{CS} + \nu_{AuS}$) and 1581 cm^{-1} (ν_{CC}) have higher relative intensity, and the signal at 608 cm^{-1} (r-i-d) disappears. Additional weak signals at 470 (CH wagging mode ω_{CH}) and 1473 cm^{-1} ($\delta_{CH} + \nu_{CC}$) appear. The signal at 692 cm^{-1} (δ_{CCC} + r-i-d) remains unchanged. Because of the adsorption of the thiophenol molecules to the Au surface during the self-assembly process, the S–H bond is cleaved, and therefore, the S–H bending mode ν_{SH} at 914 cm^{-1} is also missing in the SERS and TERS spectra. Note that the use of different excitation wavelengths can also lead to different relative peak intensities due to the different coupling of the modes to the electronic states of the molecule (vibronic coupling). However, the spectra shown here for comparison (SERS, TERS, and conventional Raman) were all obtained with the same excitation wavelength of 632.8 nm.

The main conclusions from this comparison are as follows: (1) relative peak positions in the TERS spectra of the thiophenol monolayer coincide with the relative positions in the reference Raman spectrum of thiophenol. (2) The TERS (and SERS/SHINERS) spectra exhibit different relative band intensities compared with the reference Raman spectrum. This can be rationalized by DFT calculations taking the binding of the sulfur atom to the Au surface into account.[32] (3) All participating groups see the same spectral pattern, the same bands seem to be suppressed or enhanced compared with the conventional Raman spectrum, irrespective of the tip metal and TERS configuration used.

Peak positions

One major question addressed by this comparison study is do the absolute peak positions vary and if yes by how much? Are the same spectral patterns observed when measuring the same substance? The answer to these questions can be found in Fig. 2a and in the corresponding data evaluation that is summarized in Table 4 and is described in detail in the Supporting Information pages 18–32. The absolute positions of the three bands at 998, 1022, and 1074 cm^{-1} were evaluated. These three bands were chosen as they are clearly above the noise level in all the TERS spectra and can be fitted reliably. The signals at 417 and 1581 cm^{-1} were not chosen because the first one is not strong enough in most cases for a reliable peak fit and the latter one has an asymmetric peak shape. It was found that the absolute peak positions scatter by as much as 19 cm^{-1} (for all measurements). This has to be considered when using values from the literature for peak assignments and can especially lead to erroneous assignments if components of a complex mixture are identified by only relying on values from the literature for one major peak. This large variation in absolute peak position is most likely caused by differences in spectrometer calibrations. This argument is supported by the fact that absolute peak positions vary much less when comparing spectra from the same group. The absolute peak position of the thiophenol signal at 998 cm^{-1} for the spectra of a *single* group (Supporting Information page 32) lies within 3 cm^{-1}. For group E, the values scattered the most (8.8 cm^{-1}; Supporting Information page 32), which can be explained by the low spectral resolution used (5.7 cm^{-1}/pixel, compared with

Copyright © 2014 John Wiley & Sons, Ltd.

Table 3. Overview of the measured samples

Group	A	B	C	D	E	F	G	
Feedback mechanism	STM	AFM (contact)	STM	STM	AFM (shear-force)	STM	STM	AFM (contact)
TPhe (ts Au substrate)	✔	-	✔	✔	✔	✔	✔	-
TPhe (thin Au substrate)	-	✔	-	-	-	-	✔	✔
CysPhePhe (ts Au substrate)	✖	-	✔	✔	✖	✔	✔	-
CysPhePhe (thin Au substrate)	-	✖	-	-	-	-	-	✖

STM, scanning tunneling microscopy; AFM, atomic force microscopy; ✔, reproducible spectra were obtained; ✖, no reproducible spectra were obtained; -, not measured.

Table 4. Summarized results of the peak evaluation for the thiophenol sample

Group		A		B	C	D	E	F	G (STM)	G (AFM)
Tip material		Au	Ag	Ag	Au	Ag	Au	Au	Ag	Ag
Averaged peak position (cm^{-1})	Peak 1	996.8	996.7	1006.4	1006.4	999.8	996.1	992.0	997.0	999.1
Standard deviation σ (cm^{-1})		0.9	0.2	0.3	1.4	2.3	3.0	0.4	1.1	0.1
Averaged normalized signal to noise		21.3	159.5	143.0	0.4	403.9	29.1	50.9	64.3	44.4
Standard deviation σ		4.3	228.8	76.6	0.1	522.4	31.4	10.8	67.1	7.8
Averaged peak position (cm^{-1})	Peak 2	1022.0	1021.0	1029.8	1031.3	1023.4	1019.6	1016.7	1020.5	1023.4
Standard deviation σ (cm^{-1})		0.8	1.2	0.2	2.2	0.8	4.3	0.5	1.0	0.1
Averaged normalized signal to noise		12.0	56.3	81.7	0.4	312.3	20.1	28.7	40.7	29.4
Standard deviation σ		2.6	78.0	45.0	0.2	341.4	18.5	8.6	43.1	0.0
Averaged peak position (cm^{-1})	Peak 3	1071.8	1073.4	1082.9	1084.2	1076.9	1072.4	1067.9	1071.5	1071.4
Standard deviation σ (cm^{-1})		0.5	1.6	0.6	1.3	1.6	3.2	0.7	1.3	1.8
Averaged normalized signal to noise		34.8	144.0	128.6	0.6	703.6	43.9	41.3	88.1	52.3
Standard deviation σ		10.5	186.5	69.3	0.4	656.1	54.3	9.8	94.5	1.8

STM, scanning tunneling microscopy; AFM, atomic force microscopy.

1–2 cm^{-1}/pixel for the other groups; Table 1). The same is true for the peak positions from the measurements performed with the same tip: 84% of the peak positions vary by less than 1 cm^{-1} and 94% by less than 2 cm^{-1}.

In conclusion, absolute peak positions scatter only very little (3 cm^{-1}) within a group but quite significantly (19 cm^{-1}) from group to group. In the future, a unified calibration method is thus desirable that is applicable to different spectrometers in order to ensure that the signal positions are comparable between different laboratories. Additionally, at the moment, reliable substance identification requires reference spectra from the same instrument or from another spectrometer calibrated with the same (unified) method (which is often not the case for literature references).

Relative peak positions are of significant interest as well, when comparing band patterns. For the thiophenol spectra, the distances between the three peaks at 998, 1022, and 1074 cm^{-1} scatter by 7 (peaks 1 to 2), 9 (peaks 2 to 3), and 10 cm^{-1} (peaks 1 to 3). On average, the relative peak positions within a group scatter by only 3 cm^{-1}. As in the case of the absolute peak positions, the scatter in the relative peak positions is most probably due to the very different spectrometer calibration procedures that were used (Table 4).

Therefore, when calibrating a spectrometer, a reference material/lamp should be used that has several spectral lines that are distributed over the relevant spectral range in order to account for spectral aberration of the (imaging) spectrometer, e.g. due to astigmatism, coma, or spherical aberration. An accurate spectrometer calibration is also crucial to differentiate between an induced peak shift and a shifted signal position due to different instrument calibration.

Peak intensity ratios

It is already known from SERS and TERS that the signal intensity ratios can be different from conventional Raman

Copyright © 2014 John Wiley & Sons, Ltd.

measurements.[45] Here, we want to address the question about how the signal intensity ratios vary within different TERS measurements. The ratio of the signal intensities of the peaks at 998 (peak 1) and 1074 cm^{-1} (peak 3) of thiophenol and their standard deviation σ was calculated for each group (Supporting Information pages 28–30 and 34). These two peaks were chosen because they are strong in all the spectra, and other possible peaks were either asymmetric or too weak. The results are displayed in Table 5.

The intensity ratio (peak height) of peak 3/peak 1 varies from 0.83 to 1.66 (average value: 1.35 ± 0.31) meaning that for some groups, the peak at 998 cm^{-1} was dominating the spectrum (1.6 times stronger) and for other groups the peak at 1074 cm^{-1} (1.2 times stronger). This can partly be explained by different wavelength dependent spectrometer performances but is also greatly influenced by the wavelength dependence of the tip enhancement.[46] This explains the large variation among different tips from the same group (see, e.g., the spectra from group A measured with Au tips that exhibit a relative standard deviation σ in the peak 3/peak 1 intensity ratio of 38%). Apart from that, there are also encouraging examples of remarkably reproducible peak intensity ratios within a group, e.g. the spectra of group B have a much higher reproducibility in the peak 3/peak 1 intensity ratio with a standard deviation σ of only 9%. However, small variations in the peak 3/peak 1 intensity ratio are always observed, even for spectra collected with the same tip. In this case, the peak intensity ratio varies on average by about 0.16 (see the Supporting Information pages 28 to 30). Possible reasons for the variation include changes in tip shape during the experiment (e.g. due to heating effects[47]), changes of tip position with respect to the laser focus causing changes in the plasmon resonance, changes in the distance between tip and substrate, and differences in the orientation of the molecules underneath the tip (e.g. reorientation because of heating effects).

Comparing the TERS enhancement

The thiophenol and CysPhePhe samples used in this study do not yield a far field Raman spectrum (without tip) with the measurement times and laser powers applied here. Therefore, a direct comparison of the TERS spectra based on a contrast factor[48] obtained from a 'tip-in–tip-out' experiment is not possible. Nevertheless, in order to compare the spectra in terms of TERS enhancement, the following procedure was employed: The peak height was divided by the noise level (obtained by calculating the standard deviation σ of an approximately 100 cm^{-1} segment without Raman bands). Because every group used different measurement times, laser powers, and illumination/detection geometries, this signal-to-noise ratio was normalized by dividing it by the laser power per illuminated area (for details, see the Supporting Information pages 27–30), the measurement time, and the square root of the number of accumulations (as the signal-to-noise is proportional to the square root of the number of accumulations). The results were then multiplied by 10 and can be found in Table 4.

At a first glance, the *averaged* normalized signal-to-noise ratio of the groups seems to vary a lot (Table 4): The lowest value is 0.37 and the highest is 704. Taking a closer look at the *individual* values of the different measurements (see box charts and tables in the Supporting Information pages 33–34) reveals that 90% of the normalized signal-to-noise ratios are within 20–100, with two exceptions: one Ag tip from group A reaches values of 544 and a Ag tip of group D reaches a value of 1334 (for the peak at 998 cm^{-1}). These two tips apparently exhibit an exceptionally large signal enhancement in individual measurements, leading to the large variation in the averaged normalized signal-to-noise ratios. Interestingly, the normalized signal-to-noise ratio varies not only between groups, which is understandable because every group has its own tip preparation procedure, but also within a series of experiments with the same tip and the same group (e.g. group E, tip 1, and the Ag tips from group A; see the Supporting Information pages 27–30).

CysPhePhe sample

In Fig. 3, a comparison of the CysPhePhe TERS spectra obtained by the different groups is shown. For all four successful groups (C, D, F, and G), the most abundant modes are the aromatic ring breathing mode at 1004 cm^{-1} and the ring mode at 1600 cm^{-1}.[36,49] Also visible are the modes at 1200 and 1450 cm^{-1}, as well as the broad band between 620 and 634 cm^{-1}.[36,49] The lowest spectrum in Fig. 3

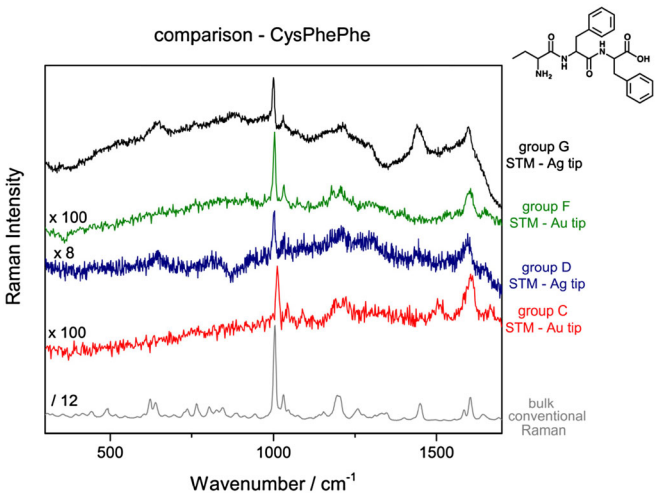

Figure 3. Representative tip-enhanced Raman spectroscopy spectra of CysPhePhe from the different groups (green, blue, red, and black). The bulk conventional Raman spectrum of CysPhePhe is displayed in gray at the bottom. The spectra were multiplied and offset in their intensity in order to show them in one graph. On the top right, the chemical structure of CysPhePhe is displayed. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

Table 5. Signal intensity ratios of the peaks at 998 and 1074 cm^{-1} and their standard deviation σ

Group	A		B	C	D	E	F	G (STM)	G (AFM)
Tip material	Au	Ag	Ag	Au	Ag	Au	Au	Ag	Ag
Peak 3/1 ratio	1.6	1.7	0.9	1.5	1.5	1.6	0.8	1.3	1.2
Standard deviation σ	0.3	0.6	0.1	0.4	0.4	0.4	0.2	0.1	0.3

STM, scanning tunneling microscopy; AFM, atomic force microscopy.

Copyright © 2014 John Wiley & Sons, Ltd.

Table 6. Summarized results of the peak evaluation for the CysPhePhe sample

Group	C	D	F	G (STM)
Tip material	Au	Ag	Au	Ag
Averaged peak position (cm^{-1})	1011.6	1004.6	1003.1	1002.2
Standard deviation σ (cm^{-1})	0.9	1.6	0.4	1.9
Averaged normalized signal-to-noise ratio	1.3	173.9	65.7	117.2
Standard deviation σ	0.6	125.7	43.4	89.5

STM, scanning tunneling microscopy.

shows the normal Raman spectrum of bulk CysPhePhe. In accordance with the thiophenol TERS spectra, the CysPhePhe TERS spectra match the reference Raman spectrum, although fewer bands are present: Only the bands for the two aromatic modes are clearly visible in all the spectra (Table 6). Note that to obtain the displayed spectra, measurement times between 10 and 30 s were necessary (varying laser powers were applied). In general, participants reported difficulties with this particular sample, from which carbonaceous contamination was much more easily produced. Accordingly, lower laser powers had to be chosen to obtain these spectra, generally around 200–400 μW (see the Supporting Information page 45).

The signal at 1004 cm^{-1} scatters by 13 cm^{-1}. This value is lower than the one for thiophenol, essentially because results from only four groups were considered for its calculation. Within three of the four groups, the absolute peak position scattered by less than 2 cm^{-1}, and for one group, it scattered by 5 cm^{-1}. The peak intensity ratios were not calculated for this sample, because the peak at 1600 cm^{-1} corresponds to two signals in the normal Raman spectrum that coalesce into one broad signal in the TERS spectrum, and there was no other strong signal present to which a distance could have been determined.

As mentioned earlier, there is an ongoing discussion in the TERS community about how well biomolecules can be detected and whether the observed spectral features are consistent. When looking at the spectra obtained from the CysPhePhe sample, the answer seems to be rather clear: Yes, biomolecules can be detected, and yes, the spectral pattern is not significantly different. Only the C–H modes at 1450 cm^{-1} exhibit significant differences in intensity. When inspecting the region above 1630 cm^{-1} in the spectra of Fig. 3, the question of the presence or absence of the amide I mode in TERS spectra can also be addressed. In some of the spectra, there seems to be a small bump in the baseline in the area where the amide I band is expected (around 1650 cm^{-1}), although no clearer statement can be made because of the low signal-to-noise level. There is definitely no strong amide I band present; it is not fully understood why this is the case. Note that CysPhePhe is only a model tripeptide, and things might be different for more complex polypeptides or proteins.

Conclusions and outlook

When performing a comparison study, the question arises why some groups obtained reproducible spectra and others did not. This question is difficult to answer, mainly because of the fact that there is no standard method to evaluate the instrument performance. Often, success or failure of a TERS experiment can be traced back to the experience of the user, the feedback mechanism, the illumination/detection geometry, and the tip type. Several parameters generally differ from one TERS experiment to another, which renders it virtually impossible to pinpoint a single aspect that is responsible for different observations.

Measurements on the opaque TS Au substrates were more successful (46% for thiophenol and 50% for CysPhePhe) than on the semi-transparent Au substrates (23% for thiophenol and 0% for CysPhePhe). This could be due to the fact that the semi-transparent substrates reduce the transmission (to 59% at 532 nm; see the Supporting Information page 35) of incoming and backscattered photons. Compared with the opaque TS Au substrates, which can only be used with top/side illumination/detection, the detection of Raman photons is therefore more hampered through the semi-transparent gold substrates.

Groups using an STM feedback were more successful (83% for the thiophenol sample and 80% for the CysPhePhe sample) than the groups using an AFM feedback mechanism (43% for the thiophenol sample and 0% for the CysPhePhe sample). One has to consider that most groups using AFM feedback work with transmission setups and therefore received semi-transparent substrates (with the restriction mentioned earlier). Other aspects that are expected to play a role are the following: (1) the tunnel current/bias voltage between the tip and substrate when using STM feedback; (2) the difference in the induced mirror dipole for different Au layer thicknesses; (3) the complexity and diversity of AFM feedback mechanisms, e.g., shear-force feedback, contact, or tapping mode is used; and (4) most groups (exception: group E) that use an AFM feedback mechanism employ metallized tips, whereas groups using an STM feedback used full metal tips.

Besides the differences mentioned earlier, it was found that

- All participating laboratories observed the same spectral pattern irrespective of the setup or tip used.
- The relative band positions in the TERS spectra match the relative peak positions in the SERS spectrum and in the reference Raman spectrum, and substance identification is reliably possible. Compared with the latter, fewer bands are present in the TERS spectrum. This can be rationalized with DFT calculations (Feugmo *et al.*).
- The relative peak positions of the major bands scattered by 3 cm^{-1} on average within a group. However, absolute signal positions scattered by 19 cm^{-1}, and the relative peak positions by 9 cm^{-1} when taking results from all groups into

Copyright © 2014 John Wiley & Sons, Ltd.

account. As a consequence, care should be taken when comparing different peaks to miscellaneous literature references. The lack of a standardized calibration procedure is suspected as the main reason for this observation.
- The peak intensity ratios vary between different groups (1.35 ± 0.31) and also within the same group. However, there are also examples where the peak intensity ratios between the tips vary by only 9%. High variations in the peak intensity ratio were also observed for measurements with the same tip.
- For the CysPhePhe samples, the two aromatic modes at 1004 and 1600 cm^{-1} dominate the spectra. The amide I mode is hardly visible even though long measurement times were used.

This study shows that using TERS, different research groups obtain the same spectral pattern on equivalent samples although using different TERS setups and tips. This brings TERS – as a reliable chemical nanoanalytical tool, e.g. for substance identification – a big step forward. In the future, TERS would greatly benefit from the availability of a standardized sample, a standardized measurement procedure, and a standardized calibration method that can be performed with all kinds of setups and tips. A research initiative funded by the European Union through the European Metrology Research Program (EMRP) on Raman metrology will certainly facilitate the development of reference samples for Raman spectroscopy. Additionally, further investigations will be necessary in order to understand the differences between AFM-TERS and STM-TERS.

Acknowledgements

The authors thank all the other groups (five in total) that put effort into measuring samples for this study but did not obtain any signals. C. B. thanks the Stiftung Stipendien-Fonds des Verbandes der chemischen Industrie e.V. (funding) and the German National Academic Foundation (ideational promotion) for scholarships. J. F. L. was supported by Bruker. Alexey Zayak, Jim Schuck, and Jeffery Neaton from the Lawrence Berkeley National Laboratory kindly provided the data of the conventional Raman spectrum of thiophenol bulk. Thomas Schmid is thanked for his advice. V. K., J. A., and M. B. R. acknowledge funding by the National Science Foundation (NSF CAREER grant CHE 0748226). B. R., X. W., and J. H. Z. acknowledge funding from MOST of China (2011YQ03012406) and NSFC (21021120456). D. R. and N. K. would like to acknowledge funding from EMRP to develop standards for Raman metrology.

Competing financial interest

T. M., S. M., S. K., K. L., and S. L. disclose competing financial interests. Their participating companies NT-MDT and Bruker actively develop and sell TERS systems and probes. Their products were used for the data collected by the NT-MDT respectively Bruker coauthors. Other authors may have NT-MDT/Bruker setups, or other NT-MDT/Bruker products in their facilities, but there were no vendor requirements or obligations for any of the no-company affiliated coauthors participating in this study.

References

[1] R. M. Stöckle, Y. D. Suh, V. Deckert, R. Zenobi, *Chem. Phys. Lett.* **2000**, *318*, 131.
[2] N. Hayazawa, Y. Inouye, Z. Sekkat, S. Kawata, *Opt. Commun.* **2000**, *183*, 333.
[3] M. S. Anderson, *Appl. Phys. Lett.* **2000**, *76*, 3130.
[4] B. Pettinger, G. Picardi, R. Schuster, G. Ertl, *Electrochemistry* **2000**, *68*, 942.
[5] R. Zhang, Y. Zhang, Z. C. Dong, S. Jiang, C. Zhang, L. G. Chen, L. Zhang, Y. Liao, J. Aizpurua, Y. Luo, J. L. Yang, J. G. Hou, *Nature* **2013**, *498*, 82.
[6] J. Stadler, T. Schmid, R. Zenobi, *ACS Nano* **2011**, *5*, 8442.
[7] A. Hartschuh, E. J. Sánchez, X. S. Xie, L. Novotny, *Phys. Rev. Lett.* **2003**, *90*, 095503.
[8] M. Paulite, C. Blum, T. Schmid, L. Opilik, K. Eyer, G. C. Walker, R. Zenobi, *ACS Nano* **2013**, *7*, 911.
[9] D. Kurouski, T. Deckert-Gaudig, V. Deckert, *J. Am. Chem. Soc.* **2012**, *134*, 13323.
[10] B. Pettinger, G. Picardi, R. Schuster, G. Ertl, *Single Mol.* **2002**, *3*, 285.
[11] J. Stadler, T. Schmid, R. Zenobi, *Nano Lett.* **2010**, *10*, 4514.
[12] T. Deckert-Gaudig, V. Deckert, *Small* **2009**, *5*, 432.
[13] Z. Liu, X. Wang, K. Dai, S. Jin, Z.-C. Zeng, M.-D. Zhuang, Z.-L. Yang, D.-Y. Wu, B. Ren, Z.-Q. Tian, *J. Raman Spectrosc.* **2009**, *40*, 1400.
[14] B. Ren, G. Picardi, B. Pettinger, *Rev. Sci. Instrum.* **2004**, *75*, 837.
[15] N. N. Jiang, E. T. E. Foley, J. M. J. Klingsporn, M. D. D. Sonntag, N. A. N. Valley, J. A. J. Dieringer, T. T. Seideman, G. C. G. Schatz, M. C. M. Hersam, R. P. R. Van Duyne, *Nano Lett.* **2012**, *12*, 5061.
[16] S. Berweger, J. M. Atkin, R. L. Olmon, M. B. Raschke, *J. Phys. Chem. Lett.* **2010**, *1*, 3427.
[17] E. M. van Schrojenstein Lantman, T. Deckert-Gaudig, A. J. G. Mank, V. Deckert, B. M. Weckhuysen, *Nat. Nanotechnol.* **2012**, *7*, 583.
[18] L. Opilik, T. Bauer, T. Schmid, J. Stadler, R. Zenobi, *Phys. Chem. Chem. Phys.* **2011**, *13*, 9978.
[19] D. Zhang, X. Wang, K. Braun, H.-J. Egelhaaf, M. Fleischer, L. Hennemann, H. Hintz, C. Stanciu, C. J. Brabec, D. P. Kern, A. J. Meixner, *J. Raman Spectrosc.* **2009**, *40*, 1371.
[20] K. F. K. Domke, D. D. Zhang, B. B. Pettinger, *J. Am. Chem. Soc.* **2006**, *128*, 14721.
[21] C. Neacsu, J. Dreyer, N. Behr, M. Raschke, *Phys. Rev. B* **2006**, *73*, 193406.
[22] K. F. Domke, D. Zhang, B. Pettinger, *J. Phys. Chem. C* **2007**, *111*, 8611.
[23] T. Deckert-Gaudig, E. Rauls, V. Deckert, *J. Phys. Chem. C* **2010**, *114*, 7412.
[24] C. Blum, T. Schmid, L. Opilik, S. Weidmann, S. R. Fagerer, R. Zenobi, *J. Raman Spectrosc.* **2012**, *43*, 1895.
[25] Z. Zhang, M. Sun, P. Ruan, H. Zheng, H. Xu, *Nanoscale* **2013**, *5*, 4151.
[26] M. Sun, Z. Zhang, H. Zheng, H. Xu, *Sci. Rep.* **2012**, *2*, 647.
[27] C. J. Powell, N. E. Erickson, T. E. Madey, *J. Electron Spectrosc.* **1979**, *17*, 361.
[28] C. J. Powell, N. E. Erickson, T. E. Madey, *J. Electron Spectrosc.* **1982**, *25*, 87.
[29] I. I. Papadakis, J. V. J. Nørgaard, E. E. Vendelbo, L. L. Van Nevel, P. D. P. Taylor, *Analyst* **2001**, *126*, 228.
[30] Bristow, Webb, *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 13.
[31] N. Leymarie, P. J. Griffin, K. Jonscher, D. Kolarich, R. Orlando, M. McComb, J. Zaia, J. Aguilan, W. R. Alley, F. Altmann, L. E. Ball, L. Basumallick, C. R. Bazemore-Walker, H. Behnken, M. A. Blank, K. J. Brown, S.-C. Bunz, C. W. Cairo, J. F. Cipollo, R. Daneshfar, H. Desaire, R. R. Drake, E. P. Go, R. Goldman, C. Gruber, A. Halim, Y. Hathout, P. J. Hensbergen, D. M. Horn, D. Hurum, W. Jabs, G. Larson, M. Ly, B. F. Mann, K. Marx, Y. Mechref, B. Meyer, U. Möginger, C. Neusüss, J. Nilsson, M. V. Novotny, J. O. Nyalwidhe, N. H. Packer, P. Pompach, B. Reiz, A. Resemann, J. S. Rohrer, A. Ruthenbeck, M. Sanda, J. M. Schulz, U. Schweiger-Hufnagel, C. Sihlbom, E. Song, G. O. Staples, D. Suckau, H. Tang, M. Thaysen-Andersen, R. I. Viner, Y. An, L. Valmu, Y. Wada, M. Watson, M. Windwarder, R. Whittal, M. Wuhrer, Y. Zhu, C. Zou, *Mol. Cell. Proteom.* **2013**, *12*, 2935–2951.
[32] C. G. Tetsassi Feugmo, V. Liégeois, *Chem. Eur. J. of Chem. Phys.* **2013**, *14*, 1633.
[33] B. Ren, G. Picardi, B. Pettinger, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 139.
[34] H. Kang, T. Park, I. Choi, Y. Lee, E. Ito, M. Hara, J. Noh, *Ultramicroscopy* **2009**, *109*, 4.
[35] I. Sirés, M. Delucchi, M. Panizza, R. Ricotti, G. Cerisola, *J. Appl. Electrochem.* **2009**, *39*, 2275.
[36] J. Guicheteau, L. Argue, A. Hyre, M. Jacobson, S. D. Christesen, *Proc. SPIE* **2006**, *6218*, 621801–6218011.
[37] S. Stewart, P. Fredericks, *Spectrochim. Acta A* **1999**, *55*, 1641.
[38] C. Blum, T. Schmid, L. Opilik, N. Metanis, S. Weidmann, R. Zenobi, *J. Phys. Chem. C* **2012**, *116*, 23061.
[39] D. Kurouski, T. Postiglione, T. Deckert-Gaudig, V. Deckert, I. K. Lednev, *Analyst* **2013**, *138*, 1665.
[40] E. A. Weiss, G. K. Kaufman, J. K. Kriebel, Z. Li, R. Schalek, G. M. Whitesides, *Langmuir* **2007**, *23*, 9686.

Copyright © 2014 John Wiley & Sons, Ltd.

[41] J.-F. Li, Y. F. Huang, Y. Ding, Z.-L. Yang, S. B. Li, X.-S. Zhou, F. R. Fan, W. Zhang, Z. Y. Zhou, D.-Y. Wu, B. Ren, Z. L. Wang, Z.-Q. Tian, *Nature* **2010**, *464*, 392.
[42] A. T. Zayak, Y. S. Hu, H. Choo, J. Bokor, S. Cabrini, P. J. Schuck, J. B. Neaton, *Phys. Rev. Lett.* **2011**, *106*, 083003.
[43] D. Scott, J. McCullough, W. Hubbard, J. Messerly, I. Hossenlopp, F. Frow, G. Weddington, *J. Am. Chem. Soc.* **1956**, *78*, 5463.
[44] D. H. Dressler, Y. Mastai, M. Rosenbluh, Y. Fleger, *J. Mol. Struct.* **2009**, *935*, 92.
[45] D. Mehtani, N. Lee, R. Hartschuh, A. Kisliuk, M. Foster, A. Sokolov, F. Cajko, I. Tsukerman, *J. Opt. A* **2006**, *8*, 183.
[46] A. Downes, D. Salter, A. Elfick, *Opt. Express* **2006**, *14*, 5216.
[47] B. Pettinger, P. Schambach, C. J. Villagómez, N. Scott, *Annu. Rev. Phys. Chem.* **2012**, *63*, 17.1.
[48] J. De Gelder, K. De Gussem, P. Vandenabeele, L. Moens, J. Raman *Spectrosc.* **2007**, *38*, 1133.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

Copyright © 2014 John Wiley & Sons, Ltd.