SERSbot: Revealing the details of SERS multi-analyte sensing using full automation

Journal:	*ACS Sensors*
Manuscript ID	se-2021-02116f.R1
Manuscript Type:	Article
Date Submitted by the Author:	10-Nov-2021
Complete List of Authors:	Grys, David-Benjamin; University of Cambridge, Cavendish Laboratory de Nijs, Bart; University of Cambridge, Department of Physics Cavendish Laboratory, Huang, Junyang; University of Cambridge, Department of Physics Scherman, Oren; University of Cambridge, Melville Laboratory for Polymer Synthesis, Department of Chemistry Baumberg, Jeremy; University of Cambridge, Department of Physics

SERSbot: Revealing the details of SERS multi-analyte sensing using full automation

David-Benjamin Grys[1], Bart de Nijs[1], Junyang Huang[1], Oren A Scherman,[2] Jeremy J. Baumberg[1*]

*E-mail: jjb12@cam.ac.uk

[1] NanoPhotonics Centre, Cavendish Laboratory, Department of Physics, JJ Thompson Avenue, University of Cambridge, Cambridge, CB3 0HE, United Kingdom

[2] Melville Laboratory for Polymer Synthesis, Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

ABSTRACT

Surface-enhanced Raman spectroscopy (SERS) is considered an attractive candidate for quantitative and multiplexed molecular sensing of analytes whose chemical composition is not fully known. In principle, molecules can be identified through their fingerprint spectrum when binding inside plasmonic hotspots. However, competitive binding experiments between methyl viologen (MV^{2+}) and its perdeuterated isotopomer (d_8-MV^{2+}) show here that determining individual concentrations by extracting peak intensities from spectra is not possible. This is because analytes bind to different binding sites inside and outside of hotspots with different capacities. Only by knowing all binding constants and geometry-related factors, can a model revealing accurate concentrations be constructed. To collect sufficiently reproducible data for such a sensitive experiment, we fully automate measurements using a high-throughput SERS optical system integrated with a liquid handling robot (SERSbot). This now allows us to accurately decompose analyte mixtures through independent component analysis (ICA) and to quantitatively map out competitive binding of analytes in nanogaps. Its success demonstrates the feasibility of automated SERS in a wide variety of experiments and applications.

KEYWORDS competitive binding, lab automation, lab robot, surface-enhanced Raman, liquid handling, quantitative SERS, multiplexed sensing, nanogap sequestration, Langmuir isotherm

Surface-enhanced Raman spectroscopy (SERS) is a powerful technique for molecular sensing. Its inherent specificity is what distinguishes SERS most from other techniques and makes it a desirable platform for multi-analyte sensing applications without the need for chemical recognition, e.g. antibodies.[1,2] The basic principle of SERS sensing is to employ the local field enhancements of optically-excited collective electron oscillations (surface plasmons) that arise in nano-patterned metals to enhance the Raman scattering process of analytes. Typically, desirable nanoscale features required for such field enhancements are achieved through either forming nano-sized cavities, vertices, or sharp edges from noble metals.

Analytes bound and trapped inside SERS hotspots provide significantly lower (many orders of magnitude) detection limits compared to Raman sensing. Unlike Raman, which allows for relating peak intensities directly to the probed chemical composition and concentrations (linear system), decomposing SERS spectra in a multi-analyte system, however, is not straightforward. This is because signal intensities, in addition to their cross-sections and individual concentrations, now depend on analytes competing for various binding sites inside and outside SERS-active hotspots.

In this study, we demonstrate this dependence by systematically analysing and quantifying the SERS response of a bi-analyte system of methyl viologen (MV^{2+}) and its perdeuterated d_8-ring-MV^{2+} (d_8-MV^{2+}). We find that the peak intensities are highly non-linear as the result of competitive binding for several limited binding sites. Only by comparing the SERS response to a complex ligand/receptor-type model (nested Hill-Langmuir equations), can the correct concentrations in mixtures be extracted. This result has far-reaching implications for many SERS sensors that target real analytes. If the chemical compositions are not entirely known, the concentrations cannot be determined.

To make this study possible, a high degree of reproducibility for the SERS measurements are crucial. This is achieved by (1) using a simple colloidal gold SERS substrate, (2) employing more sophisticated data analysis tools such as independent component analysis (ICA)[3,4], and (3) fully automating substrate and sample preparation through combining a fully automated custom-built liquid handler and a SERS optical setup into a SERS robot or 'SERSbot'. This SERSbot autonomously prepares the SERS substrates, mixes the analytes, controls aggregation and incubation times, and records the SERS spectra.

In previous studies[5–8] we have characterised a simple yet robust SERS substrate formed by mixing gold nanoparticles (AuNPs) with an off-the-shelf molecular linker (cucurbit[n]uril = CB[n]). This straightforward self-assembly protocol produces AuNP clusters with precise nanogaps, yielding highly repeatable SERS. Analytes mixed into the suspension are sequestered by the nanogaps resulting in strong SERS signals. With such facile chemistry, reproducibility is only limited by extrinsic factors such as accurate pipetting of the AuNP, CB[n], and analyte solutions as well as timing of aggregation and analyte incubation, which is all taken care of by the SERSbot.[9]

EXPERIMENTAL SECTION

SERS robot system overview. The aim of the SERSbot is to fully automate sample preparation of our nano-assemblies as well as the acquisition of SERS spectra. It is therefore comprised of a custom-built liquid handling robot and a Raman microscope. The liquid handler is designed to automate all steps required to form the SERS substrate and deliver analytes (Figure 1a). This involves preparing concentration series and on-demand mixing of arbitrary analyte mixtures. To achieve this, the robot is equipped with two single channel micropipettes, which operate on a 30 x 30 cm platform. Up to six different modules can be fixed to the platform. In the standard configuration it contains two 96 multiwell plates, two pipette tip containers and two additional modules for up to 32 glass vials (2 mL each) and six large (50 mL each) centrifuge tubes. The platform can be moved independently of the pipettes in the x_p and y_p directions, allowing it to precisely position containers under the microscope objective for SERS measurements.

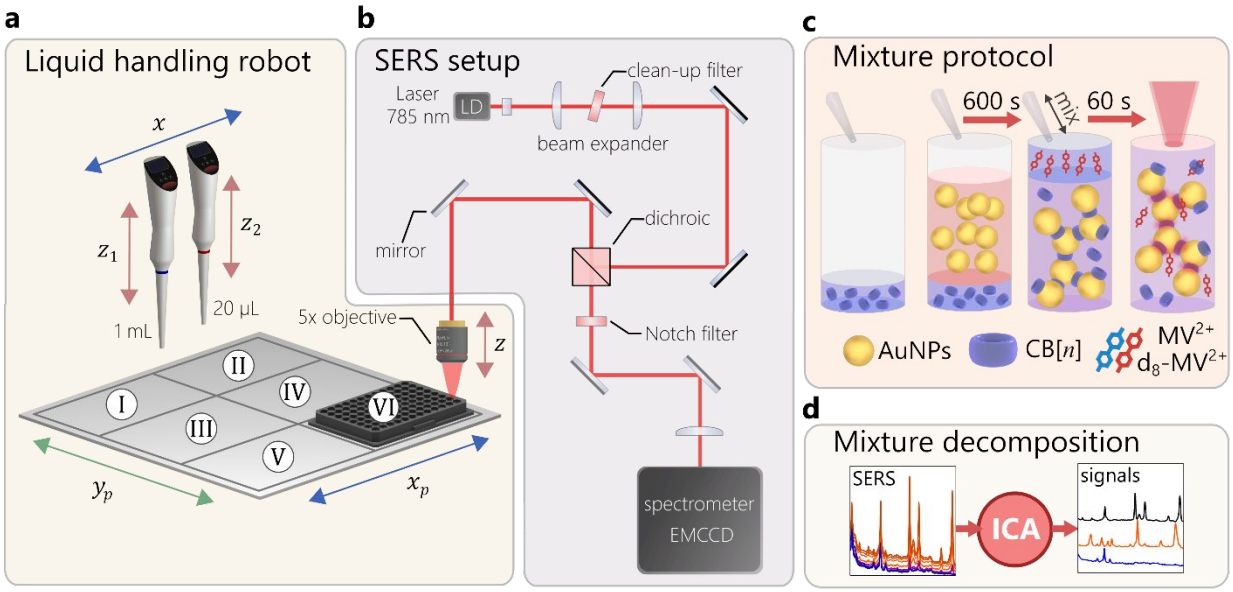


Figure 1. **SERSbot.** (a) Liquid handler robot with two micropipettes operating on platform divided into six regions designed to hold *e.g.* multiwell plates, pipette tips and vials with stock solutions. (b) Integrated SERS setup. (c) Protocol that the robot follows to combine AuNPs and CB[n] to make substrate and addition of analyte(s). (d) Scheme for ICA processing of data into component signals.

SERS setup. The SERS setup operates in the near-infrared (NIR) at 785 nm pumped by a narrow frequency volume Bragg grating filtered diode laser (Integrated Optics: 785 nm MatchBox) with up to 500 mW output power (Figure 1b). A cylindrical lens at the laser output shapes the beam profile to correct for astigmatism. After a beam expander (~3x) and a laser line clean-up filter, the beam reflects from a dichroic beam splitter sending it into the back aperture of the microscope objective. The NA=0.25 5x objective

(Zeiss) is optimised for NIR applications. Focussing of the beam onto SERS samples mounted on the liquid handler platform is optimised once at the start of each full data run, to give the largest signals. The collected SERS emission is transmitted through the dichroic beam splitter, laser scatter is removed by a 33 nm FWHM 785 nm notch filter, and then focussed onto the entrance slit of a monochromator (Shamrock, 1200l/mm grating) paired with cooled EMCCD (Andor Newton 970FI).

Liquid handler robot design. The liquid robot handler is built entirely using off-the-shelf components and 3D-printed parts. The 30 x 30 cm main platform (x_p,y_p) and the two micropipettes (x, z_1, z_2) are attached to motorised linear stages driven by steppers and belts/pulleys allowing them to move along five axes with a resolution of < 100 µm. The two linear stages (z_1, z_2) that move the micropipettes (StarLab) up and down against gravity are counterbalanced by springs to prevent the tips crashing into the main platform. The total footprint of the optics plus liquid handling measures 1×1×0.5 m (width × length × height), but could be readily compacted by 2-3 fold.

To make the robot fully autonomous, it is crucial to load fresh pipette tips whilst releasing and discarding used ones. This is normally done manually by triggering the spring-loaded ejector mechanism of the micropipette. To release tips automatically, servo motors press the release buttons (Figure S1, Video S1 showing pipette tip release), with 100% reliability.

All mechanical components are controlled by an 8-bit microcontroller (Microchip AVR Atmega256), which receives G-code-like instructions from a PC via USB. To ensure correct and safe execution of every instruction, polling in conjunction with a three-way handshake and checksums are used. The stepper motors are driven by an integrated stepper-driver, each equipped with two full H-bridges and overcurrent protection (Allergo A4988). End-stop switches at both ends of the linear stages prevent the platform and pipettes from overruns and also set the home position for each axis. The electronic pipette buttons (up, down, left, enter, dispense/aspirate) are contacted by wires connected through MOSFET-drivers to the microcontroller. The firmware is written in C and AVR assembler, and the high-level software in Python.

Sensing protocol. The protocol (Figure 1c) for the sensing experiments starts with pipetting 7 µL of 32.5 µM CB[n] followed by 313 µL of 50 nm gold nanoparticles. To allow for the formation of CB[n] AuNP aggregates, the system then waits for an optimal 600±0.1 s. While manual pipetting has ±5 s accuracy, the SERSbot electronics delivers tolerances of ±0.1s. Subsequently, the analyte or analyte mixture is added and stirred into the well plate (using the pipette tip to 'suck and dispense' three times). It is left to infuse and equilibrate for exactly 60 s (±0.1 s), and then the well plate is moved by the SERSbot under the microscope objective so that a SERS spectrum is immediately taken (or placed manually under the Renishaw inVia).

Independent component analysis. High throughput from automating SERS measurements presents both challenges and opportunities for analysis and interpretation. The ultimate goal in SERS sensing is to decompose a measured spectrum such that its source spectra and scaling factors (the mixing scores) can be extracted. In a spectrum that represents a mixture of analytes, the scores should reflect the individual analyte concentrations. The problem of extracting the source spectra without *a-priori* knowledge of either the source spectra themselves or their scores, is termed blind source separation (BSS).

This is similar to the widely-employed principal component analysis (PCA) technique whose eigen-spectra and -scores represent an orthogonal coordinate basis that maximises the variance in the data.[10] PCA works well for classification of features in SERS spectra, but fails to extract the true source spectra. For the spectral analysis of analyte mixtures (see section 3), independent component analysis (ICA) is preferable to reliably retrieve the source spectra and the scores.

This assumes that the observed spectra $\boldsymbol{x}$ are a linear combination of source spectra $\boldsymbol{s}$, mixed according to the mixing matrix $\mathbf{A}$ as $\boldsymbol{x} = \mathbf{A}\boldsymbol{s}$, with the vectors $\boldsymbol{x} = \{x_0,\ldots,x_n\}$ and $\boldsymbol{s} = \{s_0,\ldots,s_m\}$ representing the n observed and m source spectra. The $n \times m$ mixing matrix $\mathbf{A}$ is composed of the mixing scores a_{nm}. The solution $\boldsymbol{s} = \mathbf{W}\boldsymbol{x}$, requires $\mathbf{W} = \mathbf{A}^{-1}$ for the un-mixing matrix.

A solution to the BSS problem is ICA, which is robust for determining $\mathbf{W}$ and $\boldsymbol{s}$ given only the observed spectra $\boldsymbol{x}$. The presented algorithm here is based on Fast ICA.[11,12] Briefly, the key idea of Fast ICA is based on the central limit theorem stating that the distribution of a mixture of uncorrelated random variables becomes more 'Gaussian' than the original variables. Thus, an independent component can be found by maximising the non-Gaussianity of the projection $y = \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}$, so that if $y = \pm s_m$, $\boldsymbol{w}$ is one row of the un-mixing matrix $\mathbf{W}$. For SERS signals, a simple measure of Gaussianity such as kurtosis[13] proves to be sufficiently robust to recover the source signals.

Before the ICA algorithm can be executed, the spectral data $\boldsymbol{x}$ requires some pre-processing. In the first step PCA dimensionality-reduction is performed which reduces noise and removes spectral lines due to cosmic rays. ICA is prone to mistakenly identify these lines as independent components. The next step is to remove the sample mean and de-correlate (whitening) the spectral data such that $\mathrm{cov}(\boldsymbol{x},\boldsymbol{x}) = \boldsymbol{I}$**.** The resulting data vector after pre-processing $\hat{\boldsymbol{x}}$ is then fed into the ICA algorithm.

This implementation of ICA is based on a simple gradient descent (starting with a random guess for $\boldsymbol{w}$) which in every round k updates

$$\boldsymbol{w}_{k+1} = \mathrm{norm}\left[\left(\boldsymbol{E}\{\hat{\boldsymbol{x}}y^3\} - 3\boldsymbol{E}\{y^2\}\right)\boldsymbol{w}_k\right]$$

with $\boldsymbol{E}$ referring to the expected value, norm representing vector normalisation, and $\hat{\boldsymbol{x}}$ the whitened and zero-mean SERS spectra. The gradient descent algorithm for determining

the source spectra and mixing coefficients is written in Matlab (see SI), and also available as a free Python implementation in the machine learning package 'scikit-learn'.

SERS measurements. SERS spectra taken on the robot and the commercial Renishaw inVia Raman system are each averaged over three acquisitions of 10 s integration time. The laser power after the microscope objective of both systems are set to 145 mW. The Renishaw system uses a similar 5x objective (Renishaw). The reported counts are normalised to the laser power and total acquisition time (cps/mW).

DFT/thermochemistry calculations. The extracted spectra are compared with DFT calculations. This uses B3LYP at the 6-31G*/GD3 level of theory, SMD implicit water model, pre optimisation in the gas phase, as well as counterpoise correction (see results in Figure S3) The test analyte molecules used later in the work here are methyl viologen (MV^{2+}) and its deuterated version (d_8-MV^{2+}), allowing us to then evaluate the CB[7]:MV and CB[7]:d_8-MV^{2+} complexation enthalpy and Gibbs free energy.

RESULTS AND DISCUSSION

SERSbot characterisation. To identify how well the SERSbot compares to manual pipetting/high-end Raman (Renishaw inVia), the assay protocol depicted in Figure 1c is used (firstly without any analyte present). This straightforward SERS substrate (Figure 2a) is used throughout the paper and is comprised of colloidally-suspended AuNP aggregates providing plasmonically-active nanogaps delivering the SERS enhancement. Each nanogap is precisely controlled by the molecular linker cucurbit[n]uril (CB[n], n=5,7) exhibiting a fixed gap width of 0.9 nm.[14,15] These CB[n=5,7] are used because of their water solubility.[16]

For the SERSbot, a total of eight fresh CB[5]:AuNP samples were measured on different days (2-4 d) using the same AuNP stock suspension (Figure 2b, left). The spectra show typical CB[5] signals with a ring-breathing signature mode at 830 cm^{-1}.[5] All eight spectra are almost perfectly congruent, exhibiting nearly identical background shifts and peak intensities. As SERS spectra are usually known to exhibit background fluctuations, this emphasises the robustness of the CB[5]:AuNP substrate.[17] To make each spectrum more visible, they are also plotted with vertical offsets (Figure 2b, bottom).

For the commercial Raman system with careful manual pipetting, five fresh samples taken on consecutive days also show good reproducibility despite the less precise manual timing for the aggregation (±5 s). As both sets of spectra show little variances, similarities between spectra are quantified using a Pearson correlation coefficient (PCC) $r(\boldsymbol{x},\boldsymbol{y})$ which is an accepted and useful Figure of Merit for quantifying reproducibility and repeatability.[9] For two spectra $\boldsymbol{x}$ and $\boldsymbol{y}$ the PCC is defined as the ratio of the covariance $cov(\boldsymbol{x},\boldsymbol{y})$ to the product of their standard deviations $\sigma_x\sigma_y$. This is estimated by calculating the empirical covariance and standard deviation between the first spectrum of each set and the

subsequent samples (Figure 2c). As expected, the correlation coefficients (here reported as 1-r) obtained from both the SERSbot and the manual setup approaches zero, meaning that the spectra are nearly identical and highly reproducible. Average PCCs for both the best manual procedure and for the SERSbot are comparable.

With the high reproducibility of the CB[5]:AuNP aggregates the noise performance of the SERS setup and commercial Raman system are compared. This is done by normalising the sample variance of the CB[5] series to the peak intensity of its strongest vibration at 830 cm^{-1} (Figure 2d). This normalisation step is important to remove system-dependent efficiencies: the Renishaw system generates higher counts for the same CB[5]:AuNP samples, but its noise level is comparable (24.6 dB compared to 22.5 dB for the SERSbot setup). The only measurable contribution of the CB[5]:AuNP system on top of this noise floor is the variation arising from the ring-breathing mode (830 cm^{-1}).

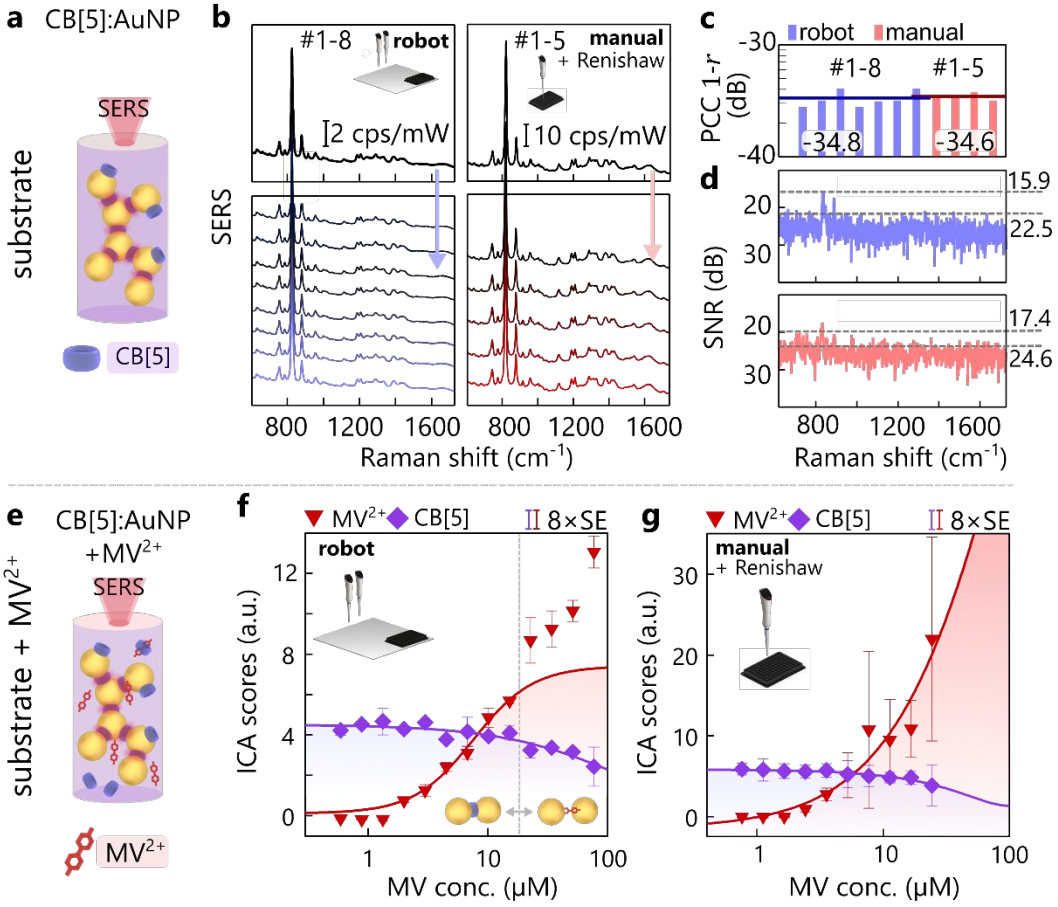


Figure 2 | SERS spectra of SERSbot vs manual pipetting. (a) Schematic CB[5]:AuNP assembly. (b) Series of SERS spectra collected from eight CB[5]:AuNP samples over 2 days using the robot liquid handler (left) and manual pipetting (right). Spectra are displayed (top) without any background correction or scaling, (bottom) spectra are stacked. (c) Pearson correlation coefficient (PCC) shown as (1-r) between first (reference) and subsequent samples, on logarithmic scale. (d) Signal-to-noise comparison of the SERS robot (top) and manual setup (bottom), normalised to the CB[5] signature peak at 830 cm^{-1}. (e) Schematic

CB[5]:AuNP substrate for MV^{2+} measurements. (f-g) ICA scores for CB[5] and MV^{2+} comparing data collected manually (g) with the SERSbot(f). Error bars span eight times the standard error to make them visible.

The overall performance of the fully automated SERSbot is thus comparable to the best manual pipetting with a high-end Raman system, for this very simple protocol. However, for more complex protocols it is evident that the SERSbot will outperform manual pipetting and spectroscopy, particularly when consistent mixing of analytes is required.

This improved performance is found when introducing even a single analyte to the substrate (Figure 2e). We compare the robot setup to manual pipetting measuring the CB[5] mixed with an analyte of methyl viologen (MV^{2+}). To do so, the protocol is extended. The first two steps, pipetting of CB[5] followed by AuNPs for 600 s aggregation time, remain the same. After this, the analyte or analyte mixture is added and stirred into the well plate (using the pipette tip to 'suck and dispense' three times). It is left to infuse and equilibrate for exactly 60 s (±0.1 s) and eventually a SERS spectrum is taken. For every concentration a total of three repeat samples are taken. To make comparison easier, the series dilution of MV^{2+} is performed by the SERSbot, which are then reused for the manual pipetting experiment. This ensures that there is no relative concentration uncertainty between the two experiments on the two different systems.

From the spectra (Figure 2f,g), the differences between the robotic system and manual pipetting/acquisition are highly evident. The manual data exhibits several times higher uncertainty (random error) for most data points (less apparent at low concentrations of MV^{2+}). This proves that for liquid SERS measurements not only precise timing and volumes are necessary but also reproducible mixing of the analyte into the suspension, which are very difficult to control for manual pipetting. This clearly shows the advantages of such measurements with the SERSbot.

For high MV^{2+} concentrations, the robotic system also shows an increase of random error, and the concentration series begins to deviate from the Langmuir-Hill fit (see SI). The reason for this deviation and increased error is likely caused by MV^{2+} molecules contributing to the aggregation of AuNPs (Figure S2). This means that a substantial number of the probed nanogaps are no longer defined by the precise gap-spacing of CB[5], thus lowering the reproducibility of the SERS measurement. This is confirmed by the CB[5] ICA score which decays for higher MV^{2+} concentrations.

Quantitative multi-analyte SERS. To show the robustness and strength of the SERSbot in combination with ICA, we now demonstrate the system's performance for a double analyte system with CB[7]:AuNP constructs. CB[7] is employed because it has been shown capable of sequestering small molecule analytes, therefore adding additional binding sites to the system.[6,15,18] Besides the CB[7] nanogap spacer (at fixed concentration), the analytes used (Figure 3a) are methyl viologen (MV^{2+}) and the perdeuterated d_8-ring-MV^{2+} (d_8-MV^{2+}) with the hydrogen atoms of the benzene rings

replaced by deuterium. Such chemical identical bi-analyte systems have proven very useful in SERS to investigate performance of nanogap systems.[19] Four sets of measurements with a total of 60 spectra were taken, preparing each sample afresh. The first two sets of measurements (I and II) are concentration series of MV^{2+} and d_8-MV^{2+} for calibrating and training the ICA. In III and IV, a competitive binding assay with MV^{2+} concentration series is performed whilst keeping the d_8-MV^{2+} fixed (at 4.6 µM and 0.9 µM, respectively), thus combining three source spectra simultaneously (Figure 3b).

The employed ICA algorithm (see methods) runs through the entire data set and returns three independent components (Figure 3c). These components clearly resemble the individual spectra of CB[7], MV^{2+} and d_8-MV^{2+} matching the measured SERS (Figure S2). This shows that ICA is indeed able to retrieve the source spectra without any *a-priori* information from complex mixture data.

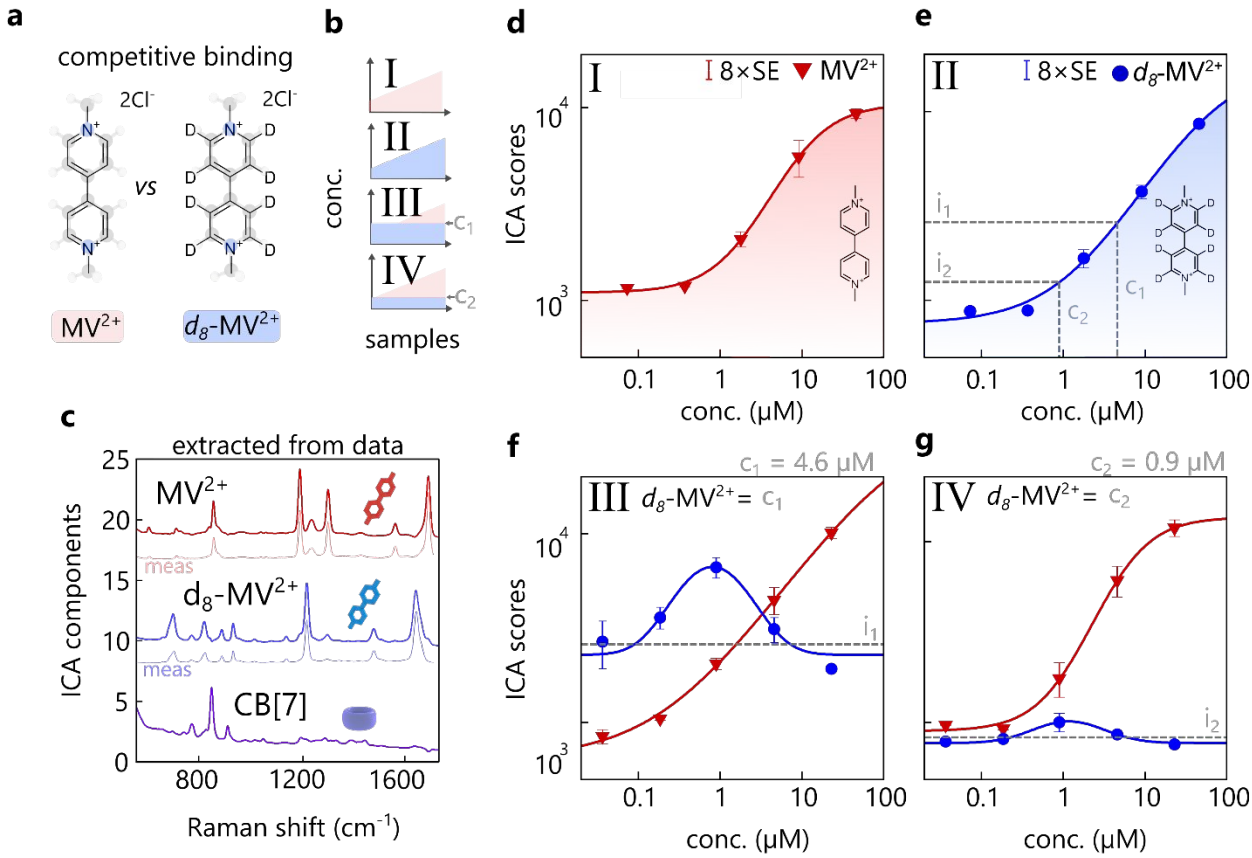


Figure 3 | SERS competitive binding assay. (a) Chemical structure of MV^{2+} and deuterated d_8-MV^{2+}. (b) Experimental protocol for CB[7]:AuNP sensing illustrating the sets of different concentration series analysed. (c) Extracted source spectra from ICA, matching measured MV^{2+}, d_8-MV^{2+} and CB[7] SERS spectra . (d-g) ICA scores for (d) MV^{2+} and (e) d_8-MV^{2+} concentration series including Hill-Langmuir fits, as well as for MV^{2+} concentration series with d_8-MV^{2+} concentration fixed at (f) 4.6 µM and (g) 0.9 µM.

Plotting the extracted ICA scores against the MV^{2+} and d_8-MV^{2+} concentrations (Figure 3d,e) reveals the expected sensing response in the nanogaps. Both can be fitted

with the Hill–Langmuir equation[20–22] (see section SI5) to retrieve the disassociation constants of MV^{2+} and d_8-MV^{2+} binding into the nanogaps, $K_d^{MV} = 20 \pm 5$ μM , $K_d^{d8-MV} = 32 \pm 5$ μM. These represent the analyte concentration at half-occupation of the nanogaps. As expected from their chemical similarity (see DFT calculations in Figure S3), the binding for both molecules is nearly identical. These micromolar values evidence the strong binding affinity of the viologen derivatives to the hydrophobic CB-filled nanogaps.

Competitive binding assay. The two molecules MV^{2+} and d_8-$MV^{2+}$are structurally identical and possess similar dissociation constants but have very different SERS spectra (Figure S2c) as the vibrational energies are inversely proportional to the square-root of the reduced mass. They are thus ideal candidates to explore nanogap sensing chemistries in conjunction with our high throughput SERSbot.

In a competitive binding assay (Figure 3 f,g) mixtures of MV^{2+} and d_8-MV^{2+} are prepared and added to the CB[7]:AuNPs according to the same protocol as for (I,II). Each sample contains a fixed concentration of d_8-MV^{2+} (4.2 μM or 0.9 μM), while the MV^{2+} concentration is varied from 73 nM to 23 μM. As we show below, despite their chemical similarity, the SERSbot assay is clearly able to show how these analytes compete with each other due to the different binding sites available.[23]

Plotting the ICA scores from the SERS spectra of the MV^{2+} (red) against the MV concentration yields another Langmuir-Hill isotherm which slightly deviates from the MV-only concentration series (Figure 3f,g). Surprisingly, the d_8-MV^{2+} scores (blue) differ significantly from the expected ICA score value, despite the fixed concentration of d_8-MV^{2+} for every data point (grey dashed lines). For MV^{2+} concentrations >10 μM, the d_8-MV^{2+} scores are well below the expected values from the d_8-MV^{2+} concentration series (Figure 3e). At such high MV^{2+} concentrations, the majority of SERS-probed nanogaps are occupied by MV^{2+}, which therefore leads to d_8-MV^{2+} scores below the expected values from the site competition. This evidences the limited number of sites available in the nanogap.

As the MV^{2+} concentration decreases, ever more d_8-MV^{2+} molecules are sequestered by the nanogaps seen in the rising d_8-MV^{2+} scores. Counterintuitively, these scores rise above the expected value, to a maximum at ~1 μM MV^{2+} concentration. For further decreases in MV^{2+} concentration, the d_8-MV^{2+} scores decay back to the expected values $c_{1,2}$ from Figure 3e.

This peculiar behaviour of the d_8-MV^{2+} response is attributed to the presence of spare CB[7] molecules outside the plasmonically-active nanogaps which form strong inclusion complexes [$\log(K_d^{MV:CB7})$ = -7] with MV^{2+}/d_8-MV^{2+}.[24,25] The probed nanogaps thus compete with CB[7] in sequestrating d_8-MV^{2+}/MV^{2+} which 200 to 300-fold prefer CB[7] ($K_d^{MV:G}$~20-30 μM *vs* $K_d^{MV:CB7}$~0.1 μM). With this knowledge it is evident how an increasing MV^{2+} concentration complexes preferentially with CB[7] thus promoting ever more d_8-MV^{2+} into the nanogaps. Once CB[7] is saturated with MV^{2+}, the d_8-MV^{2+} response reaches

its maximum and a further increase of MV^{2+} begins to displace d_8-MV^{2+} inside the nanogaps. This leads to a drop of the d_8-MV^{2+} signal for high MV^{2+} concentrations.

Nanogap sensing model. As shown in the previous section, the fixed concentration of d_8-MV^{2+} produces different signal intensities (ICA scores) depending on the MV^{2+} concentration. Evidentially, it is not possible to extract the analyte concentration simply by comparing peaks or peak ratios. Here, we introduce a quantitative model that incorporates all relevant sensing mechanisms to replicate and fit the measured data.

We assume that the plasmonic gaps act as receptors with a total nanogap binding site concentration $[G_0]$ (Figure 4a). The nanogaps sequester MV^{2+} and d_8-MV^{2+} to form the complexes $[G \cdot MV]$ and $[G \cdot dMV]$. In the same fashion CB[7] is assumed to have a total concentration of $[CB_0]$ and form the complexes $[CB \cdot MV]$ and $[CB \cdot dMV]$.

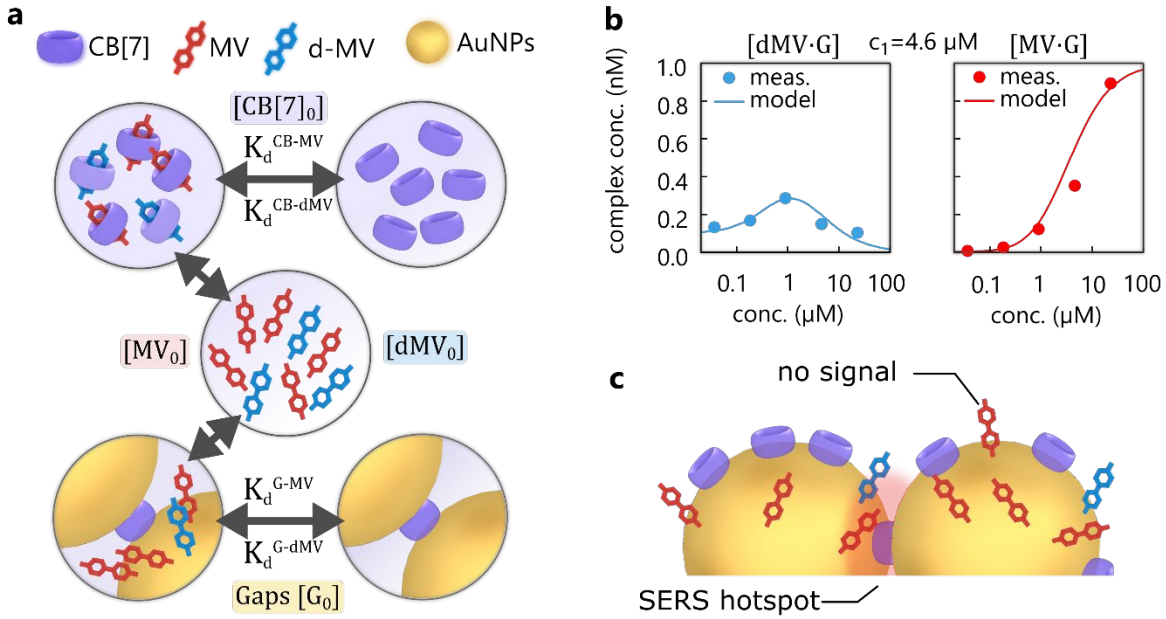


Figure 4 | Nanogap sensing model. (a) Interactions between the nanogaps G, MV, d_8-MV^{2+} and CB[7] mapped by the nanogap model. (b) Concentrations for the complexes $[G \cdot MV]$ and $[G \cdot dMV]$ which are proportional to the SERS signal replicating the SERS response measured in Figure 3g,f. (c) Illustration of nanogap hotpot showing that only a fraction of the AuNP surface contributes to the SERS signal.

The dissociation constants are defined as

$$K_d^{G:MV} = \frac{[G][MV]}{[G \cdot MV]}, \; K_d^{G:d8\text{-}MV} = \frac{[G][dMV]}{[G \cdot dMV]}$$

$$K_d^{CB:MV} = \frac{[CB][MV]}{[CB \cdot MV]}, \; K_d^{CB:d8\text{-}MV} = \frac{[CB][dMV]}{[CB \cdot dMV]}.$$

Together with the mass conservation equations

$$[\mathrm{G}_0] = [\mathrm{G}] + [\mathrm{G}\cdot\mathrm{MV}] + [\mathrm{G}\cdot\mathrm{dMV}],$$
$$[\mathrm{MV}_0] = [\mathrm{MV}] + [\mathrm{G}\cdot\mathrm{MV}] + [\mathrm{CB}\cdot\mathrm{MV}],$$
$$[\mathrm{dMV}_0] = [\mathrm{dMV}] + [\mathrm{G}\cdot\mathrm{dMV}] + [\mathrm{CB}\cdot\mathrm{dMV}],$$
$$[\mathrm{CB}_0] = [\mathrm{CB}] + [\mathrm{CB}\cdot\mathrm{MV}] + [\mathrm{CB}\cdot\mathrm{dMV}],$$

a system of eight equations is obtained.[26] These can be solved numerically for the nanogap complexes $[\mathrm{G}\cdot\mathrm{MV}]$ and $[\mathrm{G}\cdot\mathrm{dMV}]$, which are directly proportional to the SERS intensity

$$I_{SERS}^{MV} \propto [\mathrm{G}\cdot\mathrm{MV}] \text{ and } I_{SERS}^{dMV} \propto [\mathrm{G}\cdot\mathrm{dMV}].$$

Directly using this model to replicate the experimental competitive binding assay as a function of the [CB[7]]:[d_8-MV^{2+}] ratio using the concentrations and extracted disassociation constants from our data, yields a response which does not fully reproduce the d_8-MV^{2+} peaking at ~1 μM (Figure S6a). The reason for this is that the model does not account for MV^{2+} and d_8-MV^{2+} binding to the gold surface *outside* the SERS-active hotspots (Figure 4c). This means that the real disassociation constants are considerably lower (stronger affinity) and the effective concentration available for binding into the nanogaps is lower.

To compensate for this 'unspecific' binding in our model, we first estimate the relative fractions of MV^{2+}/d_8-MV^{2+} bound inside and outside the nanogaps. Electron microscopy of the fractal aggregates[11] shows that every AuNP connects to ~2.5 adjacent AuNPs. Approximating the AuNP shape as icosahedral (with 20-faced (111) facets), the effective MV^{2+}/d_8-MV^{2+} concentrations available for nanogap-binding are then ~$(2.5/20)^{-1}$=8-fold lower.

Including this geometry-specific factor into the model, a good fit of the measured data is now achieved (Figure 4b). Most convincingly, the surprising peak at ~1 μM seen in Figure 3f,g is reproduced, supporting the validity of our model. The extracted dissociation constants are then $K_d^{G:MV}$~0.63 μM and $K_d^{G:dMV}$~1.5 μM, showing that the deuterated molecule again finds it harder to bind into the nanogap, likely due to changes in its solvation in the confined environment of the gap. The extraction of 10-fold lower K_d values in this competitive binding assay than in the single-component assays (Fig. 3d,e) is due to the non-specific analyte 'theft' outside the nanogaps, and shows that understanding molecular binding in such real nano-construct substrates is important. The nanogaps possess much higher fundamental binding efficiencies for analytes than previously measured, emphasizing the need to remove surface sequestration outside nanogaps in order to maximize sensing detection limits.

Further increasing the CB[7] concentration in this model calculation (Figure S6d) shifts the d_8-MV^{2+} detection peak below ~1 μM, and sharpens it. Conversely, decreasing

the CB[7] concentration flattens the d_8-MV^{2+} signal, confirming that the presence of CB[7] is essential to form this peak. The agreement between experiment and theory also confirms that analyte binding into nanogaps is reversible, as previously suggested.[27]

From this model fit to the data, it is possible to extract the upper bound of the nanogap binding site concentration. The detection peak solely arises from the competitive binding of MV^{2+} and d_8-MV^{2+} into the nanogaps. With increasing gap concentration, this competition disappears as sufficient binding sites are available for both compounds, while lower gap concentrations also do not shift the peak in the response. Sweeping the gap concentration (Figure S6c) shows the peak is found where the gap concentration matches the CB[7] concentration (~10 μM or below). This approach thus provides a new way to independently estimate the number of nanogap binding sites per unit volume, which is required for quantitative SERS, without having further knowledge of the experimental parameters such as enhancement factor of the substrates. Without systematic data from the SERSbot, all such effects would be difficult to ascertain.

CONCLUSION

The full automation of vibrational molecular analysis by combining SERS measurements with a liquid handler into a SERS robot proves to be a viable option for providing and maintaining consistent high repeatability across an arbitrary number of samples. For the CB[n]:AuNP aggregates this is achieved by accurately dispensing solutions of CB, gold nanoparticles and analytes and carefully controlling aggregation and incubation times. The large spectral data sets produced are ideal for sophisticated data analysis, which enables quantitative and multiplexed characterisation of systematically-controlled sample sets. Using independent component analysis to characterise mixture of two molecules (MV^{2+} and d_8-MV^{2+}) we demonstrate the competition for various binding sites inside and outside the nanogap. Comparing the results to a ligand/receptor binding model confirms that the normally assumed Hill-Langmuir concentration dependence is altered. From this competitive binding assay, we also extract dissociation constants for ligand/nanogap binding, show their reversibility, and quantify competitive binding. Our nanogap sensing model confirms the subtle interactions in binding mechanisms involved, even in a seemingly simple setting. Indeed, for future work we will extend the sensing capability to mixtures of even more analytes, that the SERSbot will tackle autonomously.

SUPPORTING INFORMATION

The Supporting Information is available free of charge at [link].

- Photos of SERSbot, DFT and thermochemistry, raw spectra, methyl viologen in the nanogap, model comparison, Langmuir-Hill fit calculations, and Matlab ICA code
- SERSbot pipette tip change (MP4)

ACKNOWLEDGEMENTS

We acknowledge financial support from EPSRC Grants (EP/L027151/1, EP/R020965/1, EP/P029426/1) and ERC PICOFORCE (883703). DBG is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L015889/1 for the EPSRC Centre for Doctoral Training in Sensor Technologies and Applications. B.d.N acknowledges support from the Leverhulme Trust and the Isaac Newton Trust in the form of an ECF.

REFERENCES

(1) Yuan, Y.; Panwar, N.; Yap, S. H. K.; Wu, Q.; Zeng, S.; Xu, J.; Tjin, S. C.; Song, J.; Qu, J.; Yong, K. T. SERS-Based Ultrasensitive Sensing Platform: An Insight into Design and Practical Applications. *Coord. Chem. Rev.* **2017**, *337*, 1–33.

(2) Gao, Z.; Burrows, N. D.; Valley, N. A.; Schatz, G. C.; Murphy, C. J.; Haynes, C. L. In Solution SERS Sensing Using Mesoporous Silica-Coated Gold Nanorods. *Analyst* **2016**, *141* (17), 5088–5095.

(3) Hyvärinen, A.; Oja, E. Independent Component Analysis: Algorithms and Applications. *Neural Networks* **2000**, *13* (4–5), 411–430.

(4) Abell, J. L.; Lee, J.; Zhao, Q.; Szu, H.; Zhao, Y. Differentiating Intrinsic SERS Spectra from a Mixture by Sampling Induced Composition Gradient and Independent Component Analysis. *Analyst* **2012**, *137* (1), 73–76.

(5) Mahajan, S.; Lee, T.-C.; Biedermann, F.; Hugall, J. T.; Baumberg, J. J.; Scherman, O. A. Raman and SERS Spectroscopy of Cucurbit[n]Urils. *Phys. Chem. Chem. Phys.* **2010**, *12* (35), 10429.

(6) de Nijs, B.; Kamp, M.; Szabo, I.; Barrow, S. J.; Benz, F.; Wu, G.; Carnegie, C.; Chikkaraddy, R.; Wang, W.; Deacon, W. M.; Rosta, E.; Baumberg, J. J.; Scherman, O. A. Smart Supramolecular Sensing with Cucurbit[n]Urils: Probing Hydrogen Bonding with SERS. *Faraday Discuss.* **2017**, *205*, 505–515.

(7) Kasera, S.; Hermann, L. O.; Del Barrio, J.; Baumberg, J. J.; Scherman, O. A. Quantitative Multiplexing with Nano-Self-Assemblies in SERS. *Sci. Rep.* **2014**, *4* (6785), 1–6.

(8) Taylor, R. W.; Lee, T. C.; Scherman, O. A.; Esteban, R.; Aizpurua, J.; Huang, F. M.; Baumberg, J. J.; Mahajan, S. Precise Subnanometer Plasmonic Junctions for SERS within Gold Nanoparticle Assemblies Using Cucurbit[n]Uril "Glue." *ACS Nano* **2011**, *5* (5), 3878–3887.

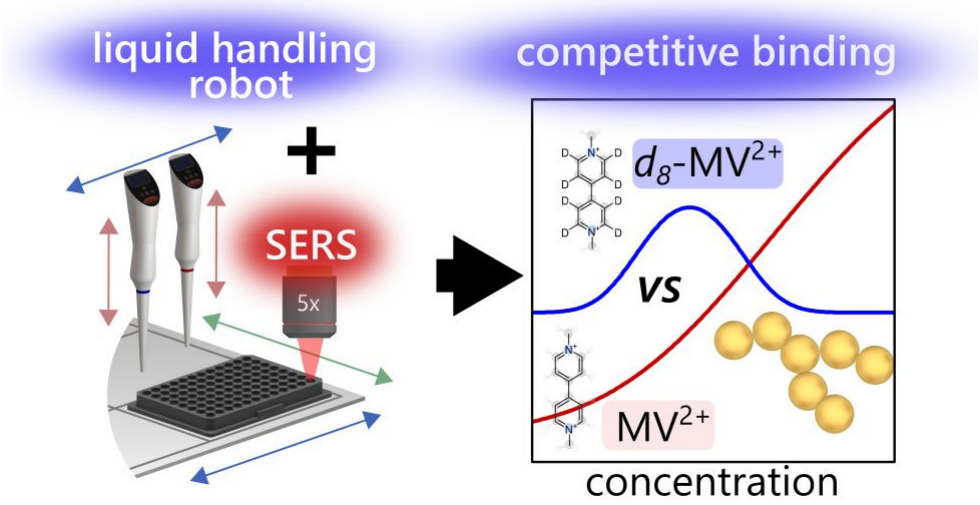
(9) Grys, D.-B.; Chikkaraddy, R.; Kamp, M.; Scherman, O. A.; Baumberg, J. J.; de Nijs, B. Eliminating Irreproducibility in SERS Substrates. *J. Raman Spectrosc.* **2020**.

(10) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2* (1–3), 37–52.

(11) Hyvärinen, A. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Trans. Neural Networks* **1999**, *10* (3), 626–634.

(12) Comon, P. Independent Component Analysis, A New Concept? *Signal Processing* **1994**, *36* (3), 287–314.
(13) Hyvarinen, A. Fast ICA for Noisy Data Using Gaussian Moments. *Proc. - IEEE Int. Symp. Circuits Syst.* **1999**, *5*.
(14) Assaf, K. I.; Nau, W. M. Cucurbiturils: From Synthesis to High-Affinity Binding and Catalysis. *Chem. Soc. Rev.* **2015**, *44* (2), 394–418.
(15) Barrow, S. J.; Kasera, S.; Rowland, M. J.; Del Barrio, J.; Scherman, O. A. Cucurbituril-Based Molecular Recognition. *Chem. Rev.* **2015**, *115* (22), 12320–12406.
(16) Murray, J.; Kim, K.; Ogoshi, T.; Yao, W.; Gibb, B. C. The Aqueous Supramolecular Chemistry of Cucurbit[: N] Urils, Pillar [n] Arenes and Deep-Cavity Cavitands. *Chemical Society Reviews*. Royal Society of Chemistry May 7, 2017, pp 2479–2496.
(17) Mahajan, S.; Cole, R. M.; Speed, J. D.; Pelfrey, S. H.; Russell, A. E.; Bartlett, P. N.; Barnett, S. M.; Baumberg, J. J. Understanding the Surface-Enhanced Raman Spectroscopy " Background." *J. Phys. Chem. C* **2010**, *114* (16), 7242–7250.
(18) Sigle, D. O.; Kasera, S.; Herrmann, L. O.; Palma, A.; De Nijs, B.; Benz, F.; Mahajan, S.; Baumberg, J. J.; Scherman, O. A. Observing Single Molecules Complexing with Cucurbit[7]Uril through Nanogap Surface-Enhanced Raman Spectroscopy. *J. Phys. Chem. Lett.* **2016**, *7* (4), 704–710.
(19) Ru, E. C. Le; Meyer, M.; Etchegoin, P. G. Bi-Analyte Surface Enhanced Raman Scattering for Unambiguous Evidence of Single Molecule Detection. **2005**.
(20) Gesztelyi, R.; Zsuga, J.; Kemeny-Beke, A.; Varga, B.; Juhasz, B.; Tosaki, A. The Hill Equation and the Origin of Quantitative Pharmacology. *Arch. Hist. Exact Sci.* **2012**, *66*, 427–438.
(21) Weiss, J. N. The Hill Equation Revisited: Uses and Misuses. *FASEB J.* **1997**, *11* (11), 835–841.
(22) Emmons, E. D.; Guicheteau, J. A.; Fountain, A. W.; Tripathi, A. Effect of Substituents on Surface Equilibria of Thiophenols and Isoquinolines on Gold Substrates Studied Using Surface-Enhanced Raman Spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22* (28), 15953–15965.
(23) Tripathi, A.; Emmons, E. D.; Kline, N. D.; Christesen, S. D.; Fountain, A. W.; Guicheteau, J. A. Molecular Structure and Solvent Factors Influencing SERS on Planar Gold Substrates. *J. Phys. Chem. C* **2018**, *122* (18), 10205–10216.
(24) Ong, W.; Kaifer, A. E. Salt Effects on the Apparent Stability of the Cucurbit[7]Uril-Methyl Viologen Inclusion Complex. *J. Org. Chem.* **2004**, *69* (4), 1383–1385.
(25) Ma, D.; Zavalij, P. Y.; Isaacs, L. Acyclic Cucurbit[n]Uril Congeners Are High Affinity Hosts. *J. Org. Chem* **2010**, *75*, 4786–4795.
(26) Wang, Z. X. An Exact Mathematical Expression for Describing Competitive Binding of Two Different Ligands to a Protein Molecule. *FEBS Lett.* **1995**, *360* (2), 111–114.
(27) De Nijs, B.; Carnegie, C.; Szabó, I.; Grys, D.-B.; Chikkaraddy, R.; Kamp, M.; Barrow, S. J.; Readman, C. A.; Kleemann, M.-E.; Scherman, O. A.; Rosta, E.; Baumberg, J. J.

Inhibiting Analyte Theft in Surface-Enhanced Raman Spectroscopy Substrates: Subnanomolar Quantitative Drug Detection. *ACS Sensors* **2019**, *4* (11).


liquid handling robot
competitive binding
SERS
5x
d_8-MV^{2+}
vs
MV^{2+}
concentration

For TOC only