Inhibiting Analyte Theft in SERS Substrates: sub-Nanomolar Quantitative Drug Detection

Bart de Nijs[1], Cloudy Carnegie[1], István Szabó[2], David-Benjamin Grys[1], Rohit Chikkaraddy[1], Marlous Kamp[3], Steven J. Barrow[3], Charlie A. Readman[3], Marie-Elena Kleemann[1], Oren A. Scherman[3], Edina Rosta[2], and Jeremy J. Baumberg[1*]

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

[2] Department of Chemistry, King's College London, 7 Trinity Street, London SE1 1DB, United Kingdom

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

Abstract:

Quantitative applications of surface enhanced Raman spectroscopy (SERS) often rely on surface partition layers grafted to SERS substrates to collect and trap solvated analytes that would not otherwise adsorb onto metals. Such binding layers drastically broaden the scope of analytes that can be probed. However, excess binding sites introduced by this partition layer also trap analytes outside the plasmonic 'hot-spots'. We show that by eliminating these binding sites, limits of detection (LODs) can effectively be lowered by more than an order of magnitude. We highlight the effectiveness of this approach by demonstrating quantitative detection of controlled drugs down to sub-nanomolar concentrations in aqueous media. Such LODs are low enough to screen, for example, urine at clinically relevant levels. These findings provide unique insights into the binding behavior of analytes, which are essential when designing high performance SERS substrates.

Keywords: SERS, tetrahydrocannabinol, THC, drug detection, self-assembly, nanoparticles, synthetic cannabinoids, spice.

Tremendous efforts have been made in the development of SERS substrates, often utilizing colloidal self-assembly or complex patterning of metal surfaces, with many variants that showcase million-fold SERS enhancements factors (EFs).[1–5] However, since EFs scale as $|E|^4$, spatial inhomogeneities in field enhancement $|E(x,y)|$ result in highly varying Raman intensities across such high performance substrates.[6] As a consequence the majority of measured SERS spectra are generated by only a small fraction of the molecules, situated in highly localized optically-active sites (hot-spots)[1,7–9] (Figure S1). This means that the adsorption location of molecules on SERS substrates greatly affects the strength of their SERS signals. However since SERS is capable of single molecule sensing[7] then, as proposed by Le Ru *et al.*, a highly optimized SERS substrate should be able to

detect every single molecule at low analyte concentrations.[9] Local variations can be effectively mitigated by collecting signals over a large number of hot-spots, thus averaging SERS intensities for a given analyte concentration.[10–12] Averaging however, results in a large fraction of analyte molecules not contributing significantly to the collected SERS spectra. This effect becomes increasingly important at low analyte concentrations when the total number of analyte molecules approaches the (large) number of binding sites available outside the hot-spot, resulting in fewer analyte molecules reaching the high-performance hot-spots.[13,14] This is here termed 'analyte theft'.

These issues are often ignored when testing novel SERS substrates. Typically an 'optimized' sample is created by coating the substrates with a dense layer of molecules with strong (typically thiol) binding groups with the sole purpose of determining an idealized enhancement factor (EF). However, in practice, analytes do not have such strong metal-binding groups, for instance biomarkers,[15] controlled substances[16] or other polycyclic aromatic hydrocarbons of interest.[17,18] Therefore, in addition to reproducible high field enhancements, an ideal SERS substrate should have at least two more features. Firstly, the SERS substrate should have either a specific or ubiquitous affinity to the analyte. A number of SERS substrates have already been presented that employ supra-molecular chemistry to capture conventionally non-binding analytes. Such substrates typically employ biofunctionalization,[19,20] amphiphilic[15] or hydrophobic[21] partition layers or amphiphilic cage constructions such as cyclodextrins[22–24] or cucurbit[n]urils (CB[n]s).[12,25–27] Secondly, a SERS substrate should preferably only bind analytes near the hot-spot to minimize analyte theft. The majority of proposed substrates however are fully coated by these receptive partition layers resulting in the number of binding sites approaching or exceeding the total number of analyte molecules available in the system when sensing at sub-micromolar concentrations (see SI section 2 for example calculation). While several techniques have been introduced to achieve hot-spot selective adsorption,[9,28–30] no study has looked at how this affects the quantitative sensing of real analytes.

Here, we present a highly reproducible self-assembled SERS substrate consisting of gold nanoparticles and cucurbit[n]uril (CB[n]) as rigid molecular linkers, with a general amphiphilic affinity to analytes. We study quantitatively the effect of eliminating indiscriminant binding on the detection of analytes at sub-micromolar concentrations. The rigid CB molecular spacer provides precise control over the inter-particle spacing in AuNPs aggregates,[26] and their hydrophobic nature combined with surface-bound charged citrate molecules provides an environment rich in both hydrophilic and hydrophobic sites. In addition, locally replacing the bounding aqueous phase with a neighboring metal nanoparticle surface renders the local chemical environment

significantly different from that of a ligand-coated nanoparticle surface. We show that these properties combine to allow for interstitial incorporation of analytes (*i.e.* outside the CB molecular cavity but within the plasmonic hot-spot). We quantitatively demonstrate that by eliminating the indiscriminant binding (analyte theft) this localized interstitial incorporation allows detection of analytes down to sub-nanomolar concentrations in water. Our results show that this interstitial binding principle can be employed to detect a wide range of analytes as the binding does not depend on the analyte's affinity to metal, but rather on its preference for the amphiphilic interactions presented within the hot-spot.

Results and Discussion

SERS substrate formation. To demonstrate interstitial incorporation of analytes and show the benefits of preventing indiscriminate binding, plasmonic substrates consisting of self-assembled AuNPs with a range of molecular spacers of cucurbit[n]uril (CB[n]), were compared, where n is 5, 6, 7 or 8.[25] Adding CB[n] to a dispersion of citrate-stabilized AuNPs induces self-assembly, forming aggregates as the particles stick together via the CB[n] which act as rigid 0.9 nm molecular spacers (Figure 1**a**).[26,31] This aggregation takes about 10 minutes during which a gradual colour change from red to blue-grey is observed (inset Figures 1**a**,**b**). The resulting aggregates consist of a collection of plasmonic hot-spots with reproducible localized field enhancements as a result of the rigid sub-nm separations.[26,32]

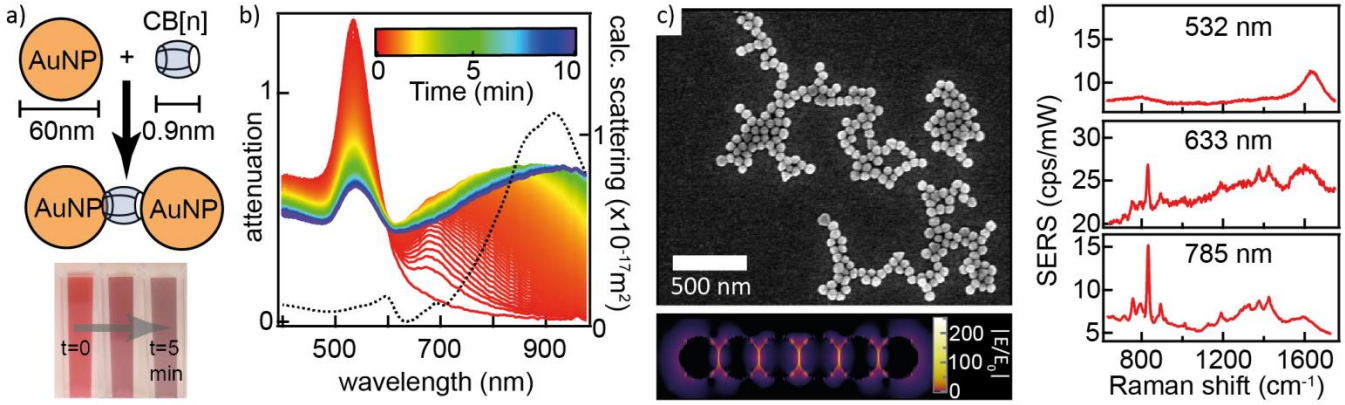

Figure 1: Surface enhanced Raman spectroscopy (SERS) substrate formation and properties. **a**) Adding CB[n] to a solution of AuNPs (diameter 60 nm) induces aggregation observed as a colour change from red to grey, with inter-particle spacing of 0.9 nm (the height of the spacer). **b**) Extinction spectra of the self-assembly process showing the formation of chain modes in solution over time. Dashed line: Finite difference time domain (FDTD) simulated far field scattering spectrum for a six membered AuNP chain. **c**) Scanning electron microscopy (SEM) image of AuNP aggregates formed by CB[n] self-assembly showing fractal-like structures. Inset: Modelled AuNP chain showing localised hot-spots between the nanoparticles with field enhancements $|E/E_0|$ up to 250. **d**) SERS spectra from AuNP aggregates under illumination at 532 nm (top), 633 nm (middle) and 785 nm (bottom) in counts per second per milliWatt (cts·s^{-1}·mW^{-1}).

SERS substrate characterization. Absorbance spectra during aggregation (Figure 1**b**) show a drop in the single nanoparticle mode (at 534 nm) combined with a rise of the dimer mode (at 690 nm) and chain modes (up to 1000 nm). The red-shifted chain modes visible at the culmination of aggregation result from the coupling together of the individual hotspot modes, feasible only due to the exact reproducibility of the gap spacing. After 10 minutes the aggregation is seen to terminate with a predominant scattering mode around 900 nm.[26,31,33] Scanning electron micrographs (SEM) of the aggregates (Figure 1**c**) show a fractal-like structure, with chain lengths between two and seven nanoparticles characteristic for this self-assembly process.[26,32] Modelling a chain of six 60 nm AuNPs with 0.9 nm spacings using finite difference time domain (FDTD) simulations matches the dominant scattering mode at 900 nm observed in the absorbance experiments (black dashed line, Figure 1b). Plotting the field enhancements of the modelled structure clearly shows that the highest enhancements are localized within the gaps between the nanoparticles.[34,35] This simplified linear chain is expected to have comparable field enhancements to our aggregates since bends in the chains are found to have limited effects on the resonant localization properties.[32] The simulated structure shows field enhancements up to $|E/E_0|$=250 at 900 nm, which implies enhancement factors around 10^9 when exciting and collecting at resonance.[6] The broad absorbance spectra in Figure 1**b** suggest relatively high enhancement factors are expected over a wide range of wavelengths, from 700 nm to 1000 nm, though a local maximum is also observed at 534 nm for transverse modes of the chains.

Comparing three different excitation wavelengths (Figure 1**d**) shows that at 532 nm (transverse mode) no clear SERS signals are observed. For both 633 nm and 785 nm excitations, the clear peaks seen around 830 cm^{-1} are characteristic for CB[n]. The highest emission (in counts per second per milliWatt: $cts \cdot s^{-1} \cdot mW^{-1}$) is observed for 785 nm excitation, as expected from the absorbance spectra in Figure 1**b**. Using a 5x microscope objective ensures a large volume of ~10^7 hot-spots are simultaneously probed in solution, providing the averaging as noted above, which is required for reproducible and quantitative SERS spectra.

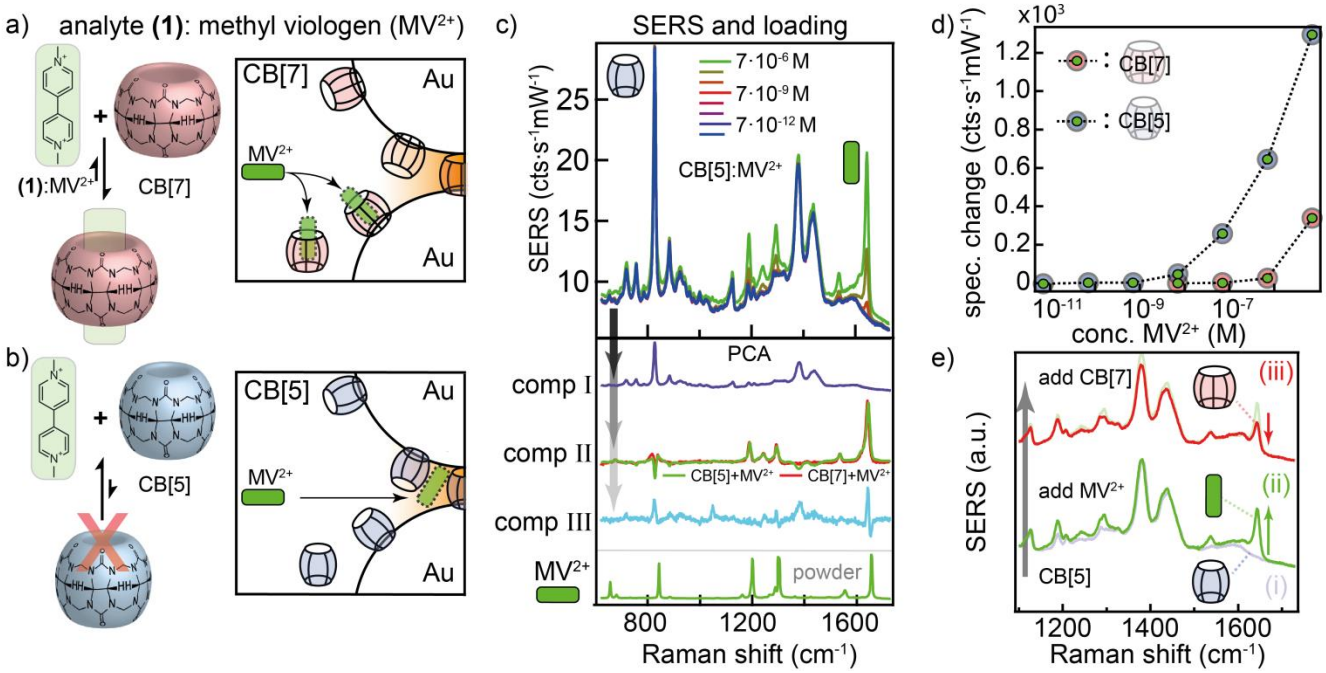

Figure 2: Analyte incorporation mechanisms in plasmonic hot-spots. **a**) Methyl viologen (MV^{2+}) has a strong binding affinity towards CB[7], binding also outside the plasmonic hot-spots, effectively lowering the probed MV^{2+} concentration. **b**) CB[5] is too small to bind MV^{2+} inside, but the constricted hotspot volume (orange shaded) binds analytes interstitially. **c**) (top) SERS spectra for MV^{2+} using CB[5] for different MV^{2+} concentrations down to picomolar. (bottom) Principal component analysis (PCA) components from CB[5]:MV^{2+} concentration series, matching CB[5] (comp I) and MV^{2+} bulk Raman (comp II). **d**) Integrated spectral changes *vs* MV^{2+} concentration for AuNP aggregates formed with CB[5] and CB[7]. **e**) SERS spectra showing the effect of adding (i) CB[5], then (ii) MV^{2+} resulting in a clear new peak at 1650 cm^{-1} and subsequently (iii) CB[7], lowering the intensity of the peak at 1650 cm^{-1} as CB[7] scavenges analytes away from the hot-spot.

Analyte binding mechanisms. The amphiphilic nature of CB[n] allows the larger variants (n=7 and 8) to sequester a range of molecules in their hydrophobic cavity, binding them to the substrate.[11,12,25] However we show here that at very small analyte concentrations, binding sites outside the hot-spots, arising from excess CB[7,8] molecules in solution and attached to the substrate, scavenge analytes away from the plasmonic hot-spots, thus effectively lowering the Raman scattering intensity for a given concentration (Figure 2**a**, box).

The AuNP metal surfaces are coated with a layer of hydrophobic CB[n] molecules (Figure 3), and water, as well as a mixed coating of trisodium citrate (hydrophilic) and citric acid (hydrophilic) used for colloidal charge stabilization. Bringing two such Au surfaces close together around the hot-spot creates a local environment particularly dense in local molecular interactions that no longer resemble a continuous solid-liquid interface. This change in environment seems to enhance binding of amphiphilic analytes from the aqueous phase due to the close proximity of both hydrophilic and hydrophobic sites

(Figure 2**b**). Such host-guest type of association is similar to that of the CB[7] system (Figure 2a) but instead occurs through interstitial incorporation. When using the smaller CB[5] molecule, size selection prevents the binding of anything larger than methane or methanol inside the small CB volume.[25,36] This prevents analytes from adsorbing at sites outside the hot-spot leaving only the interstitial incorporation mechanism to capture analytes (Figure 2b).

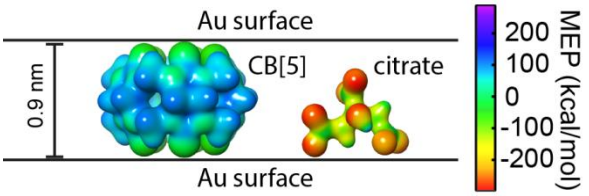

Figure 3: Calculated molecular electrostatic potential maps in implicit water for both CB[5] and citrate showing a strong negative potential for citrate and neutral/positive potential for CB[5].

Methyl viologen (MV^{2+}) is an amphiphilic analyte too large to fit in CB[5] but with a large affinity to CB[7]. When added to AuNPs aggregated using CB[5] spacers, a set of distinct MV^{2+} peaks appears between 1200-1300 cm^{-1} and at 1650 cm^{-1}, evident from nanomolar concentrations upwards, demonstrating interstitial incorporation (Figure 2**b**, top). This is in line with earlier observations for ethanol/methanol sensing using CB[5].[36] Principal component analysis (PCA) is used to isolate the spectral changes and identify their corresponding chemical moieties (Figure 2**c**, bottom). PCA allows correlated variables (in this case spectral features) to be identified and through orthogonal transformations combined into uncorrelated linear combinations of spectra. These transformed combinations are called principal components (here referred to as 'comp'). The PCA loading plot for comp I closely matches the characteristic CB[5] spectrum in Figure 1**d**, and comp II can be closely matched to the powder Raman spectrum of MV^{2+} (bottom trace in Figure 2**c**). The obtained comp II for CB[5] (green trace) and CB[7] (red trace) are nearly identical, eliminating possible additional differences between the binding mechanisms that could contribute to the enhancement factor but which would change the spectral shape or intensity ratios such as analyte orientation or binding into the metal surface (bottom Figure 2**c** comp II).[37]

Multiplying the obtained PCA score for comp II with the absolute counts integrated over the full spectral range of the loading plot for comp II provides a measure of the change in the SERS spectra upon MV^{2+} addition in $cts \cdot s^{-1} \cdot mW^{-1}$ (Figure 2**d**). Comparing the SERS changes between aggregates formed with CB[5] and CB[7] clearly shows stronger spectral peaks for CB[5]. At sub-micromolar concentrations changes are visible only for the CB[5] aggregates, showing an improvement in the LOD by more than an order of magnitude in spite of the smaller spacer's inability to directly bind MV^{2+}. To demonstrate this scavenging effect more clearly, MV^{2+} was added to CB[5]-AuNP aggregates, giving a clear set of SERS peaks (Figure 2**e**, lower trace), and subsequently CB[7] was added resulting in

a reduction of the MV^{2+} peaks (Figure 2**e**, upper trace). Both experiments confirm that the excess binding sites introduced by CB[7] scavenge analytes away from the hot-spots.

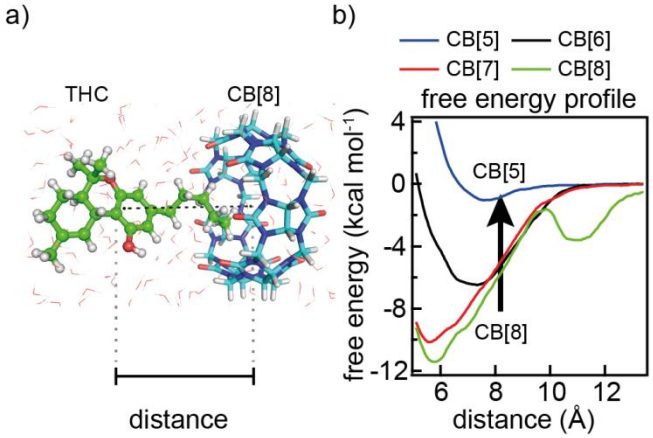

Figure 4: Molecular dynamics simulations of Δ^9-tetrahydrocannabinol (THC) interacting with different-sized CB[*n*] spacers. **a**) Scheme depicting the biasing coordinate used for the umbrella sampling free energy calculations for a THC molecule entering the CB[*n*] cavity, with explicit water. **b**) Free energy profiles calculated for each THC-CB[*n*] complex as a function of centre-of-mass distance showing a free energy dip of -9 and -11 kcal·mol^{-1} for THC-CB[7] and THC-CB[8] complexes respectively, decreased binding affinity for CB[6], and no favourable binding free energy for CB[5].

Drug detection. We studied this system in more detail by varying both the CB[*n*] spacer size and the chemical nature of the analyte molecule. To demonstrate that this improvement of LOD is not unique to MV^{2+} and to showcase the robustness of this technique, a set of controlled substances were explored. The chosen substances were selected for their interest in healthcare and substance control, and would typically require at least nanomolar sensitivities to accurately determine their concentrations in urine after consumption.[38] Here we use Δ^9-tetrahydrocannabinol (THC) the principal psychoactive constituent in cannabis (chemical structure shown in Figure 5**a**) and several synthetic analogues with different chemical structures designed to induce similar psychotropic effects. Molecular dynamics simulations were performed using umbrella sampling to model the THC molecule binding into the cavity of the CB[*n*] spacers (Figure 4**a**). Free energy profiles along the association coordinate were generated as a function of the center-of-mass distance for each THC-CB[*n*] complex (Figure 4**b**). This shows that both CB[7] and CB[8] have a highly favorable binding to THC, whereas the binding free energy gain is nearly halved for CB[6] and almost non-existent for CB[5], showing clearly the effect of reducing the spacer cavity size on analyte binding. We calculate the binding energy for each system from more accurate DFT calculations (see methods) to model the interacting complexes, showing that weaker binding affinities are indeed predicted as the size of the spacer is reduced (Figure 5**a**).

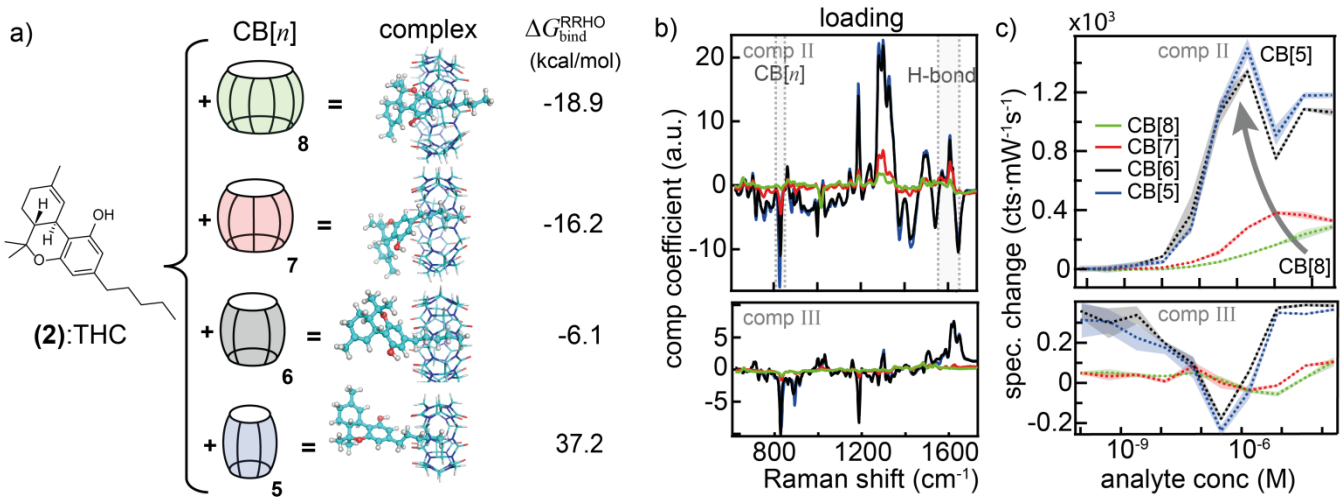

Figure 5: Influence of analyte binding mechanism on analyte detection. **a**) THC binding affinities to each of the CB[*n*] spacers, modelled using DFT calculations, see methods for details. **b**) Experimental PCA loading plots from concentration series of each THC-CB[*n*] complex showing (top) comp II: THC, and (bottom) comp III: unassigned molecular interactions. **c**) PCA scores for each of the four complexes show an increase in scores (proportional to signal strength) with decrease in CB[*n*] spacer size (arrow).

To experimentally probe how these differences in binding affect analyte detection a concentration series of THC, diluted in methanol, was measured using SERS substrates prepared with each of the different CB[*n*] spacers (Figure 5**a,b**). A significantly higher analyte component II coefficient for the loading plots was found when using the smaller CB[5] and CB[6] compared to their larger homologues CB[7] and CB[8] with a more rapid increase and higher maximum counts with the same concentration for the smaller spacers. In component III a range of peaks appear around 1600 cm^{-1}, which we tentatively assign to hydrogen-bonding related interactions (from tri-sodium citrate, methanol, water, or THC), indicative of analyte binding within the complex environment.[36] When comparing the spectral changes for each of the spacers, CB[7] and CB[8] show analyte detection at sub-micromolar concentrations, but a clear enhancement of spectral changes and lower LOD is observed for CB[5] and CB[6] (Figure 5**c**), in line with the earlier observations (Figure 2**c**). This again shows that the analyte is selectively incorporated within the substrate hotspots independent of direct binding within the spacer, and that eliminating excess binding improves the detectivity of the THC molecule.

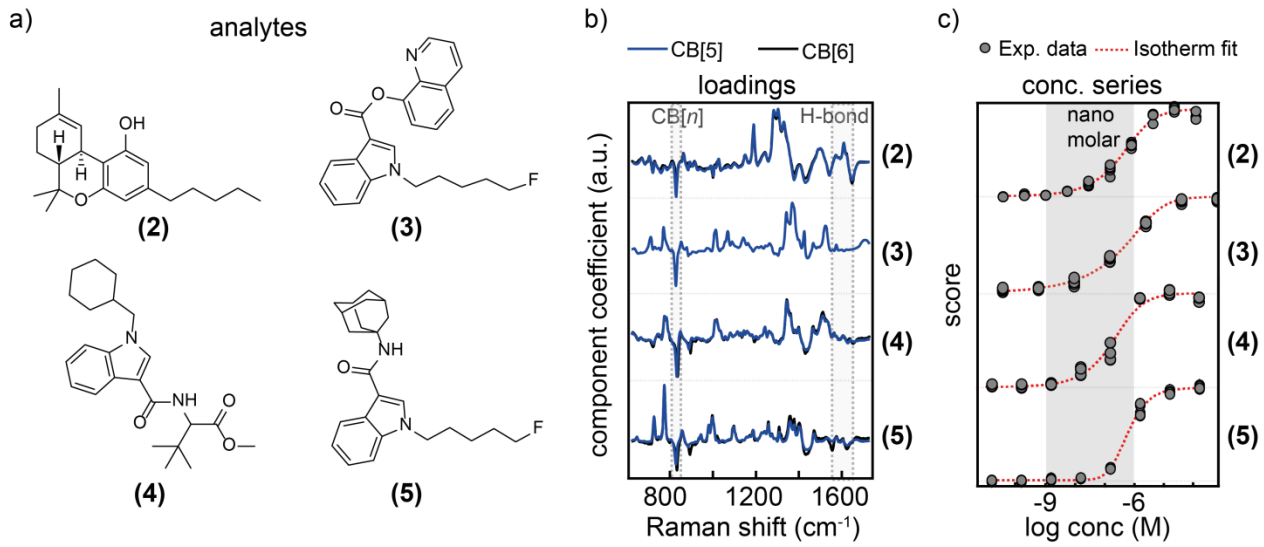

Figure 6: Non-specific binding of plasmonic hot-spots. **a**) Four different analytes: THC: **(2),** and three synthetic analogues **(3)**, **(4)** and **(5)**. **b**) PCA loading plots showing distinct spectra for each compound, with little difference whether CB[5] or CB[6] is used. **c**) PCA scores and Langmuir isotherm fits for each of the components shows LODs clearly in the nanomolar regime with compounds **(2-4)** showing LODs near or below 1 nanomolar concentrations.

Since the observed interstitial binding is independent of the CB[n] spacer cavity at low concentrations, these SERS substrates allow for more ubiquitous analyte incorporation. This makes such substrates a powerful new tool when probing for a range of different analytes such as the many synthetic analogues of THC that have appeared on consumer markets in recent years.[39–42] To demonstrate that these substrates can indeed incorporate different compounds, a concentration series of three synthetic analogues of THC are also measured (Figure 6**a**). When comparing the loading plots for each of the compounds, other than the characteristic CB[n] peak at 830 cm^{-1} and varying peaks between 1550-1700 cm^{-1}, each compound provides a clearly distinct spectrum acting as a unique fingerprint identifier (Figure 6**b**). The demonstrated non-specificity to analytes makes this method of sensing highly suitable for routine screening of such compounds. The technique readily copes with rapid changes in chemical structures, required when probing for such compounds.[42] Comparing the PCA results, it is clear that all compounds can be readily detected at nanomolar concentrations, which is well below typical clinical levels (see Figure 6**c**).[38,42,43] To obtain an estimate of the LOD for each compound a Hill-Langmuir isotherm was fitted to the PCA scores (see ESI section 4-6 for details) using:

$$\text{PCA score} = A\frac{1}{1+\left(\frac{K_\text{d}}{[\text{analyte}]}\right)^N}$$

where A is the saturation value, K_d the dissociation coefficient, [analyte] the analyte concentration and N the Hill coefficient. The residuals on these fits from the noise in the SERS spectra allow estimation of the concentration at which the highest peak would be discernable from the noise (ESI section 6). This provides an insight into the LOD for each

analyte. Although in practice LODs are expected at slightly higher concentrations since several peaks need to clear the noise threshold (>0.03 cts·mW^{-1}·s^{-1}) for a spectrum to be distinct and recognizable (Table 1).

Table 1. Estimated limits of detection based on Hill-Langmuir fit and spectral noise

Analyte	concentration @ signal>noise
(2) Δ^9-tetrahydrocannabinol (THC)	0.34(±0.02)·10^{-9} M
(3) 5F-PB-22	0.05(±0.01)·10^{-9} M
(4) MMB-CHMICA	0.40(±0.09)·10^{-9} M
(5) 5F-AKB48	26.0(±0.03)·10^{-9} M

To confirm these estimated LODs are truly realistic, spectral changes at analyte concentrations near the LOD are compared to the noise threshold (see ESI section 6 for details). The high reproducibility of the SERS spectra allows for the reference to be reliably subtracted from the raw data revealing spectral changes arising with the addition of the analyte and its carrier solvent, as shown for analyte (**2**:THC) in Figure 7.

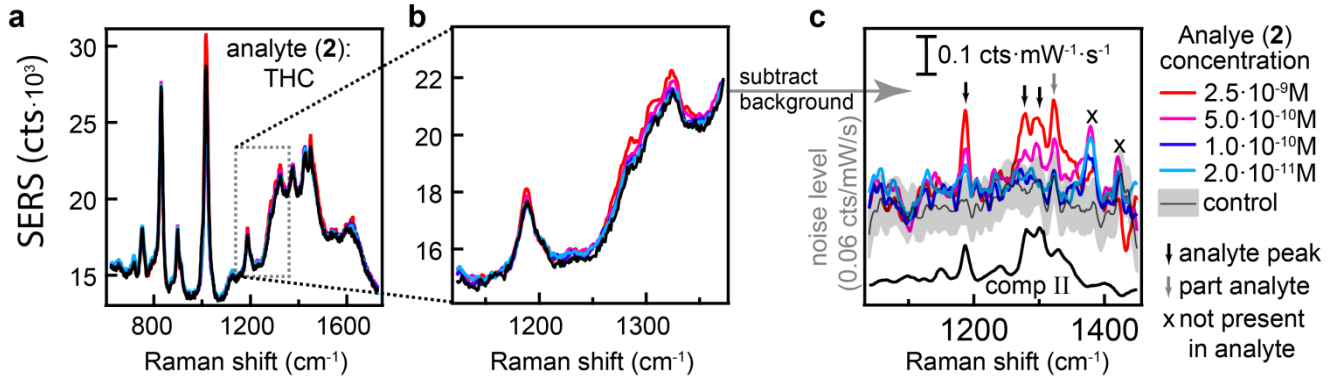

Figure 7: Validation of the LOD for analyte (2). a) SERS Spectra of CB[5]:AuNP aggregates with 4 different analyte concentrations (2.5 nM, 0.5 nM, 0.1 nM, 0.02 nM). Zoomed-in region of interest showing small spectral changes. c) SERS spectra with background subtracted, showing peaks for analyte (**2**) exceeding the noise threshold for 2.5 nM and 0.5 nM concentrations (arrows).

At 2.5 nM and 0.5 nM the analyte peaks are still recognizable and exceed the noise threshold (Figure 7**c**), while at 0.1 nM the signal has dropped into the noise. This is in good agreement with the derived LOD of 0.34 nM, showing that using a Hill-Langmuir fit with PCA scores is a suitable technique to approximate LODs. Such low LODs are typically the preserve of immunoassay SERS substrates tailored to detect a specific-analyte.[20] Interestingly a higher LOD is observed for compound (**5**), and is paired with a higher Hill coefficient (see SI Table S2) indicating a stronger competitive binding occurs for this analyte. Exploring in detail what determines this difference in LOD will further push understanding of the complex interactions present in self-assembled plasmonic nanogaps, and is the subject of ongoing research. However it is clear is that the chemical

environment of plasmonic gaps can be exploited for interstitial analyte incorporation and that eliminating excess binding sites has a drastic effect on improving the LODs. On this basis, new strategies can be developed for existing substrates to remove or passivate excess binding. Such strategies can for example involve multiple washing steps to remove excess binding sites, or adding ions or large molecules to block these sites, leaving only hot-spots exposed.

Conclusion

We have demonstrated an interstitial analyte incorporation mechanism in self-assembled colloidal SERS substrates and used it to show the effects of analyte 'theft' by indiscriminate binding on the limits of detection. We have shown that for tetrahydrocannabinol (THC) and all three tested synthetic analogues, weaker binding of molecular spacers results in higher SERS signals and lower LODs, reaching sub-nanomolar concentrations. These findings highlight that for SERS-based detection of analytes at very low concentrations, indiscriminate binding of target molecules should be eliminated where possible, as this has a detrimental effect on signal strengths and when successful can increase the limits of detection by more than an order of magnitude.

Methods

Concentration series: Tetrahydrocannabinol (1 mg/mL in methanol) and methyl viologen dichloride were purchased from Sigma Aldrich, the synthetic analogues **3**-**5** were provided by Tic Tac Communications, and all chemicals were used as received. The different analyte concentrations were prepared by volumetric dilution of analytes using either water (for MV^{2+}) or methanol (Laboratory reagent grade, Fisher Scientific) as solvent. Vials containing the diluted analyte concentrations were sealed and used within 1 hour of preparation to minimize effects of solvent evaporation.

Formation of SERS substrates: 60 nm AuNP suspensions were purchased from BBI Solutions (citrate capped, optical density OD1) and stored at 7 °C. Prior to use the AuNP suspension was allowed to reach room temperature. CB[*n*] molecular spacers were synthesized and separated according to the procedure described in reference [25]. To induce self-assembly 7 µL of a 1 mM solution of CB[*n*] was added to the bottom of a black polystyrene 96-well plate (Thermo Fisher Scientific). 300 µl of AuNP suspension was added and allowed to aggregate for 10 minutes.

Analyte detection: The CB[5], CB[6], CB[7] and CB[8] concentration series were measured using the same stock solutions, freshly prepared from a 1 mg/ml solution in methanol using volumetric dilution with a suitable carrier solvent (methanol for the synthetic cannabinoids, water for methyl viologen). Specifically 1 mL of analyte (**2**) at 1 mg/ml in MeOH was added to an empty 5ml volumetric flask and filled to the appropriate volume using MeOH. The new concentration in the flask (now 0.2 mg/ml) was stored in a sealed container and 1 ml was drawn for the next dilution step. For SERS measurements 20 µL of analyte solution was added to the aggregated suspension, mixed, and allowed to homogenize for 2 more minutes. SERS spectra were taken on a commercial Renishaw Raman setup using either a 532 nm, a 633 nm or a 785 nm laser, with typical quantitative measurements taken using a 785 nm laser at 119 mW, by combining 3 iterations with 10 second

integration time. For focusing and collection a 5x 0.15 NA Olympus objective was used giving an estimated spot size of 0.4 mm^3. To demonstrate reproducibility typical measurements were performed at least in threefold, meaning three unique samples were created by combining CB[n] and AuNPs and adding the desired analyte concentration from a stock solution.

Principal component analysis: Prior to principal component analysis (PCA) a linear background was subtracted from each of the spectra using the lowest point in the spectra. The WaveMetrics Igor implementation of PCA was used to calculate the loading plots and scores for each of the components. The PCA results were rotated as described in reference [36].

Finite difference time domain simulations (FDTD): FDTD simulations were performed using Lumerical FDTD Solutions v8.12. The Au NP chains were modeled as linear assemblies of core-shell spheres with a core diameter of 60 nm of Au and dielectric shell of 0.9 nm with refractive index 1.45. The dielectric function of gold was taken from Johnson and Christy. The structure was illuminated with a broadband plane wave (TFSF source) polarized along the chain length. The scattering and near field intensities were obtained from inbuilt cross-section and near-field monitors. The narrow gaps of the plasmonic chains were simulated by using multiple meshing of the narrow gaps and nanoparticles. The calculations were converged at 0.3 nm meshing for the gaps along the dimer axis of the NPs and with dx=dy=dz=1 nm meshing throughout the NP volume. Care was taken to ensure there were no staircasing artefacts in defining the curved surface of nanoparticles. We have previously shown the importance of meshing in the accurate determination of field volumes and their contribution to near-fields.[44]

Density functional theory (DFT) calculations: Gas phase and subsequent continuum solvent geometry optimizations of the complexes (THC@CB[n], n=5-8), host (CB[n], n=5-8), and guest (THC) molecules were performed using the hybrid B3LYP exchange-correlation functional in combination with the split-valence double-zeta polarized basis set, 6-31G* and including Grimme's D3 dispersion correction with Becke-Johnson damping.[45] Continuum solvent geometry optimizations were performed using the SMD continuum models parametrized for water. The gas phase potential energies of the THC@CB[n], n=5-8 complexes were corrected for basis set superposition error, which is significant due to the incompleteness of the present basis set. For the accurate description of the low frequency modes an ultrafine DFT integration grid was used. No symmetry restrictions were imposed during the geometry optimization procedure. Frequency calculations with SMDsolvent model[46] were performed at the same level of theory to obtain the association Gibbs free energies, $G_0^{\mathrm{RRHO/QH}}$(l) and enthalpies, H_0^{RRHO}(l) in the rigid rotor/harmonic-oscillator (RRHO) and quasi-harmonic (mixture of RRHO and free rotor vibrational entropies along with translational entropy correction based on the free space accessible to the solute) [47,48] (QH) approximation and including zero-point vibrational energy at 298 K and 1 atm. Final continuum solvent solution phase association Gibbs free energies ($\Delta G_{\mathrm{bind}}^{\mathrm{RRHO/QH}}$), and enthalpies ($\Delta H_{\mathrm{bind}}^{\mathrm{RRHO/QH}}$) were calculated by adding the counterpoise correction, δE^{CP}(g):

$$\Delta G_{\mathrm{bind}}^{\mathrm{RRHO/QH}} = \Delta G_0^{\mathrm{RRHO/QH}}(\mathrm{l}) + \delta E^{\mathrm{CP}}(\mathrm{g})$$
$$\Delta H_{\mathrm{bind}}^{\mathrm{RRHO/QH}} = \Delta H_0^{\mathrm{RRHO/QH}}(\mathrm{l}) + \delta E^{\mathrm{CP}}(\mathrm{g})$$

where Δ represents that the supramolecular approach $\Delta X = X(\mathrm{complex}) - X(\mathrm{host}) - X(\mathrm{guest})$ has been used. The association free energies are summarized in Table S1. All standard DFT calculations were performed by the Gaussian 09[49] *ab initio* program package.

Free energy profiles of association: Molecular dynamics (MD) simulations were performed with the NAMD 2.9[50] program using the CHARMM36[51] force field. The THC@CB[*n*], *n*=5-8 complexes were solvated in a pre-equilibrated TIP3P cubic water box of edge 65 Å. The resulting systems contain 8689, 8685, 8669 and 8660 H_2O molecules for the *n*=5, 6, 7 and 8, CB[*n*] analogues, respectively. Our MD protocol consisted of: (1) energy minimization over 15000 steps; (2) equilibration over 1 ns in the NPT ensemble (ρ = 1.01325 bar, T = 303.15 K) with the RMSD of heavy atoms in CB[*n*], *n*=5-8 and THC constrained to their initial position using a force constant of 1 kcal/(mol·Å^2); (3) 2 ns run in the NPT ensemble; (4) umbrella sampling (US) production runs of 5 ns in the NPT ensemble for each umbrella window with a spring constant of 100 kcal/(mol·Å^2). Temperature and pressure were held constant at 303.15 K and 1 atm, respectively. Constant temperature was set by a Langevin thermostat with a damping coefficient of 1 ps^{-1}. All of the bonds and angles involving hydrogen atoms were constrained by the SHAKE[52] algorithm. We used the particle mesh Ewald method[53] for the long-range electrostatics in combination with a 12 Å cutoff for the evaluation of the non-bonded interactions. Trajectories were run with a time step of 2 fs and the collective variable employed in US were printed out in each step and used for the analysis. The umbrella bias for the host-guest association process was defined as the distance between center of mass (COM) of CB[*n*], *n*=5-8 and the COM of the THC ligand. We used the dynamic histogram analysis method (DHAM)[54] to compute the free energy profiles along the association coordinate.

Supporting Information

Supporting Information Available: The following files are available free of charge.

ESI.pdf

The supporting information contains the following items: (1) a high resolution FDTD modelling of the field enhancement and the resulting SERS intensity probability distribution. (2) An example calculation of analyte theft in a colloidal substrate. (3) DFT calculations for the binding energies between THC and CB[*n*]. (4) A step-by-step analysis of the performed PCA analysis. (5) A demonstration of analyte identification using a simple Pearson correlation and comparisons to bulk powder spectra. (6) Explanation about the reproducibility and noise estimation of the system and Langmuir-Hill fit parameters.

Corresponding Author

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

Notes

The authors declare no competing financial interest.

Acknowledgements

We acknowledge financial support from EPSRC Grants EP/L027151/1, EP/L015978/1 and EP/N020669/1) and the European Research Council grant BioNet 757850. B.d.N acknowledges support from the Leverhulme Trust and Isaac Newton Trust, C.C. acknowledges support from NPL (PO443073), R.C. acknowledges support from the Dr. Manmohan Singh scholarship from St. John's

College, M.K. thanks the European Commission for a Marie Skłodowska-Curie Fellowship (SPARCLEs, 7020005), S.J.B. thanks the European Commission for a Marie Skłodowska-Curie Fellowship (NANOSPHERE, 658360).

References

(1) Cialla, D.; März, A.; Böhme, R.; Theil, F.; Weber, K.; Schmitt, M.; Popp, J. Surface-Enhanced Raman Spectroscopy (SERS): Progress and Trends. *Anal. Bioanal. Chem.* **2012**, *403* (1), 27–54.

(2) Sharma, B.; Cardinal, M. F.; Kleinman, S. L.; Greeneltch, N. G.; Frontiera, R. R.; Blaber, M. G.; Schatz, G. C.; Duyne, R. P. V. High-Performance SERS Substrates: Advances and Challenges. *MRS Bull.* **2013**, *38* (8), 615–624.

(3) Sharma, B.; Frontiera, R. R.; Henry, A.-I.; Ringe, E.; Van Duyne, R. P. SERS: Materials, Applications, and the Future. *Mater. Today* **2012**, *15* (1), 16–25.

(4) Guerrini, L.; Graham, D. Molecularly-Mediated Assemblies of Plasmonic Nanoparticles for Surface-Enhanced Raman Spectroscopy Applications. *Chem. Soc. Rev.* **2012**, *41* (21), 7085–7107.

(5) Schlücker, S. Surface-Enhanced Raman Spectroscopy: Concepts and Chemical Applications. *Angew. Chem. Int. Ed. 53* (19), 4756–4795.

(6) Le Ru, E. C.; Etchegoin, P. G. Rigorous Justification of the |E|4 Enhancement Factor in Surface Enhanced Raman Spectroscopy. *Chem. Phys. Lett.* **2006**, *423* (1), 63–66.

(7) Le Ru, E. C.; Meyer, M.; Etchegoin, P. G. Proof of Single-Molecule Sensitivity in Surface Enhanced Raman Scattering (SERS) by Means of a Two-Analyte Technique. *J. Phys. Chem. B* **2006**, *110* (4), 1944–1948.

(8) Radziuk, D.; Moehwald, H. Prospects for Plasmonic Hot Spots in Single Molecule SERS towards the Chemical Imaging of Live Cells. *Phys Chem Chem Phys* **2015**, *17* (33), 21072–21093.

(9) Le Ru, E. C.; Grand, J.; Sow, I.; Somerville, W. R. C.; Etchegoin, P. G.; Treguer-Delapierre, M.; Charron, G.; Félidj, N.; Lévi, G.; Aubard, J. A Scheme for Detecting Every Single Target Molecule with Surface-Enhanced Raman Spectroscopy. *Nano Lett.* **2011**, *11* (11), 5013–5019.

(10) McNay, G.; Eustace, D.; Smith, W. E.; Faulds, K.; Graham, D. Surface-Enhanced Raman Scattering (SERS) and Surface-Enhanced Resonance Raman Scattering (SERRS): A Review of Applications. *Appl. Spectrosc.* **2011**, *65* (8), 825–837.

(11) Kasera, S.; Herrmann, L. O.; Barrio, J. del; Baumberg, J. J.; Scherman, O. A. Quantitative Multiplexing with Nano-Self-Assemblies in SERS. *Sci. Rep.* **2014**, *4*, 6785.

(12) Kasera, S.; Biedermann, F.; Baumberg, J. J.; Scherman, O. A.; Mahajan, S. Quantitative SERS Using the Sequestration of Small Molecules Inside Precise Plasmonic Nanoconstructs. *Nano Lett.* **2012**, *12* (11), 5924–5928.

(13) Le Ru, E. C.; Etchegoin, P. G.; Meyer, M. Enhancement Factor Distribution around a Single Surface-Enhanced Raman Scattering Hot Spot and Its Relation to Single Molecule Detection. *J. Chem. Phys.* **2006**, *125* (20), 204701.

(14) Fang, Y.; Seong, N.-H.; Dlott, D. D. Measurement of the Distribution of Site Enhancements in Surface-Enhanced Raman Scattering. *Science* **2008**, *321* (5887), 388–392.

(15) Yonzon, C. R.; Haynes, C. L.; Zhang, X.; Walsh, Joseph T.; Van Duyne, R. P. A Glucose Biosensor Based on Surface-Enhanced Raman Scattering: Improved Partition Layer, Temporal Stability, Reversibility, and Resistance to Serum Protein Interference. *Anal. Chem.* **2004**, *76* (1), 78–85.

(16) Alharbi, O.; Xu, Y.; Goodacre, R. Detection and Quantification of the Opioid Tramadol in Urine Using Surface Enhanced Raman Scattering. *Analyst* **2015**, *140* (17), 5965–5970.

(17) A. Alvarez-Puebla, R.; M. Liz-Marzán, L. Traps and Cages for Universal SERS Detection. *Chem. Soc. Rev.* **2012**, *41* (1), 43–51.

(18) Alharbi, O.; Xu, Y.; Goodacre, R. Simultaneous Multiplexed Quantification of Nicotine and Its Metabolites Using Surface Enhanced Raman Scattering. *Analyst* **2014**, *139* (19), 4820–4827.
(19) C. Bantz, K.; F. Meyer, A.; J. Wittenberg, N.; Im, H.; Kurtuluş, Ö.; Hoon Lee, S.; C. Lindquist, N.; Oh, S.-H.; L. Haynes, C. Recent Progress in SERS Biosensing. *Phys. Chem. Chem. Phys.* **2011**, *13* (24), 11551–11567.
(20) Wang, G.; Park, H.-Y.; Lipert, R. J.; Porter, M. D. Mixed Monolayers on Gold Nanoparticle Labels for Multiplexed Surface-Enhanced Raman Scattering Based Immunoassays. *Anal. Chem.* **2009**, *81* (23), 9643–9650.
(21) Jones, C. L.; Bantz, K. C.; Haynes, C. L. Partition Layer-Modified Substrates for Reversible Surface-Enhanced Raman Scattering Detection of Polycyclic Aromatic Hydrocarbons. *Anal. Bioanal. Chem.* **2009**, *394* (1), 303–311.
(22) Xie, Y.; Wang, X.; Han, X.; Xue, X.; Ji, W.; Qi, Z.; Liu, J.; Zhao, B.; Ozaki, Y. Sensing of Polycyclic Aromatic Hydrocarbons with Cyclodextrin Inclusion Complexes on Silver Nanoparticles by Surface-Enhanced Raman Scattering. *Analyst* **2010**, *135* (6), 1389–1394.
(23) Abalde-Cela, S.; Hermida-Ramón, J. M.; Contreras-Carballada, P.; Cola, L. D.; Guerrero-Martínez, A.; Alvarez-Puebla, R. A.; Liz-Marzán, L. M. SERS Chiral Recognition and Quantification of Enantiomers through Cyclodextrin Supramolecular Complexation. *ChemPhysChem 12* (8), 1529–1535.
(24) Pande, S.; Ghosh, S. K.; Praharaj, S.; Panigrahi, S.; Basu, S.; Jana, S.; Pal, A.; Tsukuda, T.; Pal, T. Synthesis of Normal and Inverted Gold−Silver Core−Shell Architectures in β-Cyclodextrin and Their Applications in SERS. *J. Phys. Chem. C* **2007**, *111* (29), 10806–10813.
(25) 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.
(26) 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.
(27) 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.
(28) Chen, T.; Wang, H.; Chen, G.; Wang, Y.; Feng, Y.; Teo, W. S.; Wu, T.; Chen, H. Hotspot-Induced Transformation of Surface-Enhanced Raman Scattering Fingerprints. *ACS Nano* **2010**, *4* (6), 3087–3094.
(29) Lee, A.; Andrade, G. F. S.; Ahmed, A.; Souza, M. L.; Coombs, N.; Tumarkin, E.; Liu, K.; Gordon, R.; Brolo, A. G.; Kumacheva, E. Probing Dynamic Generation of Hot-Spots in Self-Assembled Chains of Gold Nanorods by Surface-Enhanced Raman Scattering. *J. Am. Chem. Soc.* **2011**, *133* (19), 7563–7570.
(30) Chen, T.; Du, C.; Tan, L. H.; Shen, Z.; Chen, H. Site-Selective Localization of Analytes on Gold Nanorod Surface for Investigating Field Enhancement Distribution in Surface-Enhanced Raman Scattering. *Nanoscale* **2011**, *3* (4), 1575–1581.
(31) Lee, T.-C.; A. Scherman, O. Formation of Dynamic Aggregates in Water by cucurbit[5]uril Capped with Gold Nanoparticles. *Chem. Commun.* **2010**, *46* (14), 2438–2440.
(32) Esteban, R.; Taylor, R. W.; Baumberg, J. J.; Aizpurua, J. How Chain Plasmons Govern the Optical Response in Strongly Interacting Self-Assembled Metallic Clusters of Nanoparticles. *Langmuir* **2012**, *28* (24), 8881–8890.
(33) Taylor, R. W.; Esteban, R.; Mahajan, S.; Aizpurua, J.; Baumberg, J. J. Optimizing SERS from Gold Nanoparticle Clusters: Addressing the Near Field by an Embedded Chain Plasmon Model. *J. Phys. Chem. C* **2016**, *120* (19), 10512–10522.

(34) Romero, I.; Aizpurua, J.; Bryant, G. W.; Abajo, F. J. G. de. Plasmons in Nearly Touching Metallic Nanoparticles: Singular Response in the Limit of Touching Dimers. *Opt. Express* **2006**, *14* (21), 9988–9999.
(35) Carnegie, C.; Chikkaraddy, R.; Benz, F.; de Nijs, B.; Deacon, W. M.; Horton, M.; Wang, W.; Readman, C.; Barrow, S. J.; Scherman, O. A.; et al. Mapping SERS in CB:Au Plasmonic Nanoaggregates. *ACS Photonics* **2017**, *4* (11), 2681–2686.
(36) Nijs, B. de; Kamp, M.; Szabó, I.; J. Barrow, S.; Benz, F.; Wu, G.; Carnegie, C.; Chikkaraddy, R.; Wang, W.; M. Deacon, W.; et al. Smart Supramolecular Sensing with Cucurbit[N]Urils: Probing Hydrogen Bonding with SERS. *Faraday Discuss.* **2017**, *205* (0), 505–515.
(37) Feilchenfeld, H.; Chumanov, G.; Cotton, T. M. Photoreduction of Methylviologen Adsorbed on Silver. *J. Phys. Chem.* **1996**, *100* (12), 4937–4943.
(38) Karschner, E. L.; Schwilke, E. W.; Lowe, R. H.; Darwin, W. D.; Pope, H. G.; Herning, R.; Cadet, J. L.; Huestis, M. A. Do Δ9-Tetrahydrocannabinol Concentrations Indicate Recent Use in Chronic Cannabis Users? *Addict. Abingdon Engl.* **2009**, *104* (12), 2041–2048.
(39) Kikura-Hanajiri, R.; Uchiyama, N.; Kawamura, M.; Goda, Y. Changes in the Prevalence of Synthetic Cannabinoids and Cathinone Derivatives in Japan until Early 2012. *Forensic Toxicol.* **2013**, *31* (1), 44–53.
(40) Spaderna, M.; Addy, P. H.; D'Souza, D. C. Spicing Things up: Synthetic Cannabinoids. *Psychopharmacology (Berl.)* **2013**, *228* (4), 525–540.
(41) Lindigkeit, R.; Boehme, A.; Eiserloh, I.; Luebbecke, M.; Wiggermann, M.; Ernst, L.; Beuerle, T. Spice: A Never Ending Story? *Forensic Sci. Int.* **2009**, *191* (1–3), 58–63.
(42) Namera, A.; Kawamura, M.; Nakamoto, A.; Saito, T.; Nagao, M. Comprehensive Review of the Detection Methods for Synthetic Cannabinoids and Cathinones. *Forensic Toxicol.* **2015**, *33* (2), 175–194.
(43) Behonick, G.; Shanks, K. G.; Firchau, D. J.; Mathur, G.; Lynch, C. F.; Nashelsky, M.; Jaskierny, D. J.; Meroueh, C. Four Postmortem Case Reports with Quantitative Detection of the Synthetic Cannabinoid, 5F-PB-22. *J. Anal. Toxicol.* **2014**, *38* (8), 559–562.
(44) Kongsuwan, N.; Demetriadou, A.; Chikkaraddy, R.; Benz, F.; Turek, V. A.; Keyser, U. F.; Baumberg, J. J.; Hess, O. Suppressed Quenching and Strong-Coupling of Purcell-Enhanced Single-Molecule Emission in Plasmonic Nanocavities. *ACS Photonics* **2018**, *5* (1), 186–191.
(45) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.
(46) Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Use of Solution-Phase Vibrational Frequencies in Continuum Models for the Free Energy of Solvation. *J. Phys. Chem. B* **2011**, *115* (49), 14556–14562.
(47) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem. – Eur. J.* **2012**, *18* (32), 9955–9964.
(48) Mammen, M.; Shakhnovich, E. I.; Deutch, J. M.; Whitesides, G. M. Estimating the Entropic Cost of Self-Assembly of Multiparticle Hydrogen-Bonded Aggregates Based on the Cyanuric Acid·Melamine Lattice. *J. Org. Chem.* **1998**, *63* (12), 3821–3830.
(49) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. Gaussian 09, Revision D. 01. **2009**.
(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.
(51) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145.

(52) Andersen, H. C. Rattle: A “velocity” Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52* (1), 24–34.
(53) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.
(54) Rosta, E.; Hummer, G. Free Energies from Dynamic Weighted Histogram Analysis Using Unbiased Markov State Model. *J. Chem. Theory Comput.* **2015**, *11* (1), 276–285.