

SUPPORTING INFORMATION

A Surface Enhanced Raman Spectroscopy Based Quantitative Bioassay on Aptamer-Functionalized Nanopillars Using Large-Area Raman Mapping

Jaeyoung Yang,^{†,#} Mirko Palla,^{†,||,#} Filippo Giacomo Bosco,^{⊥,#} Tomas Rindzevicius,[⊥] Tommy Sonne Alstrøm,[‡] Michael Stenbæk Schmidt,[⊥] Anja Boisen,[⊥] Jingyue Ju,^{||} and Qiao Lin^{†,*}

[†]*Department of Mechanical Engineering, Columbia University, New York, NY 10027, United States*

[⊥]*Department of Micro & Nanotechnology, Technical University of Denmark, Lyngby, 2800, Denmark*

[‡]*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, 2800, Denmark*

^{||}*Department of Chemical Engineering, Columbia University, New York, NY 10027, United States*

*** To whom correspondence should be addressed. Email: qlin@columbia.edu**

CONTENTS

1. HOTSPOT DISTRIBUTION THEORY	S3
2. EXPERIMENTAL PROCEDURES	S4
2.1. Manual Instrumentation	S4
2.2. Analysis of Manual Spectral Data	S4
3. CONTROL EXPERIMENTS	S5
3.1. Surface functionalization-dependent controls	S5
3.2. Analyte-dependent controls	S5
4. FIGURES AND TABLES	S6
REFERENCES	S20

1. HOTSPOT DISTRIBUTION THEORY

Based on equation (1) in the main text and a theoretical framework describing enhancement factor distribution around a single SERS hotspot¹, the probability density function $p(F)$ that a probe molecule at a random position experiences a given enhancement F was derived as:

$$p(F) \approx AF^{-(1+k)} \quad (\text{I})$$

which is a so-called long-tail distribution similar to the Pareto distribution. According to the rigorous definition of F considering single molecular (SM) point of view;² it is reasonable to assume that F is proportional to the measured intensity I for the same event:

$$F \propto I \quad (\text{II})$$

Thus, analogous to equation (II) an analytical expression can be developed for the intensity distribution $d(I)$ for the single hotspot model corresponding to a truncated Pareto distribution (TPD). It is truncated, since the distribution does not extend to $I \rightarrow \infty$, but has a maximum value at I_{max} corresponding to the largest experimentally obtained intensity value in the sample. Therefore, we can approximate the intensity distribution by a TPD with four parameters: k , A , I_{max} and I_{min} , which provides a very accurate description for the large intensity regime.

$$d(I) \approx AI^{-(1+k)} \quad (\text{III})$$

where k is a parameter which determines how fast the enhancement decreases when moving away from the hotspot; A is an indicator of how probable it is for a molecule to be located close to the hotspot; and I_{max} , I_{min} are the maximum, minimum intensities of the hotspot respectively and can be viewed as the global “strength” extremes of the hotspot. The intensity distribution function, $d(I)$ has the property that $d(I) \rightarrow 0$ as $I \rightarrow \infty$, which gives rise to the power rate convergence in the long-tail of the TPD.

2. EXPERIMENTAL PROCEDURES

2.1. Manual Instrumentation

All manually collected SERS spectra were recorded using a Jobin-Yvon LabRam ARAMIS Raman microscope (Horiba, Japan) in a standard backscattering configuration with a 633 nm excitation laser. The laser beam was focused onto the sample using a 50x long working-distance (NA = 0.5) dry objective (Nikon, Japan). All manual spectra in this work were obtained with an exposure time of 1 s and at 0.3 mW laser power before the objective.

2.2. Analysis of Manual Spectral Data

A data set of N spectra (in this work, typically $N = 10$) acquired at randomly selected regions on the same substrate was obtained. Data are presented as background removed averages of such a data set. Spectra were processed using the Savitzky-Golay fourth derivative method (window size of 25 data points), which can effectively reduce or eliminate possible false correlations resulting from a constant offset or broadband background.^{3,4}

3. CONTROL EXPERIMENTS

3.1. Surface functionalization-dependent controls

Several control experiments were performed to confirm the specific detection capability of aptamers. Functionalization dependent controls were performed on substrates functionalized without aptamers and with nonspecific aptamers (immunoglobulin E (IgE)-specific aptamers) and compared with substrates functionalized with vasopressin-specific aptamers; each substrate was reacted with 1 nM TVP sample. Using the Raman mapping on the substrate, 1000 spectra were collected 3 times (on 3 independent substrates) at each condition and top 20% of the SERS signals at each diagnostic TAMRA peak were integrated and averaged for intensity comparison. Our results (Figure S1a) demonstrated that substrates functionalized with vasopressin specific aptamers have superior specificity to TVP to the others. Signals from two control substrates could be attributed to nonspecific adsorption to surfaces and electrostatic attraction to charged DNA strands.

3.2. Analyte-dependent controls

Competition-induced analyte-dependent controls were compared between samples with TVP only, TVP with VP (unlabeled vasopressin which induces competition binding to vasopressin-specific aptamer with TVP) and TVP with DVP (D-enantiomer of VP which has little affinity to the vasopressin-specific aptamer, which minimally interfere TVP-aptamer binding). The results (Figure S1b) showed that the SERS signals from TVP were suppressed only when the sample includes VP, implying that TVP-aptamer binding must be competed with VP-aptamer one.

4. FIGURES AND TABLES

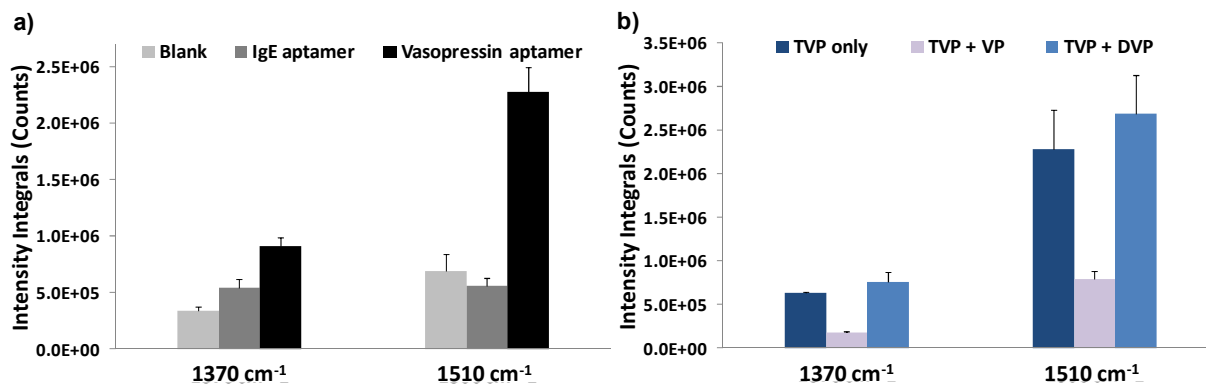


Figure S1 – Experimental results for various controls are shown. (a) Surface functionalization-dependent controls were performed on 3 kinds of substrate without aptamer (light grey) and with nonspecific aptamers (dark grey) / specific aptamers (black). (b) Competition-induced analyte-dependent controls were tested for 3 samples, TVP only (dark blue), TVP with VP (light blue) and TVP with DVP (blue). Analyte concentration is constant at 1 nM for all control experiments.

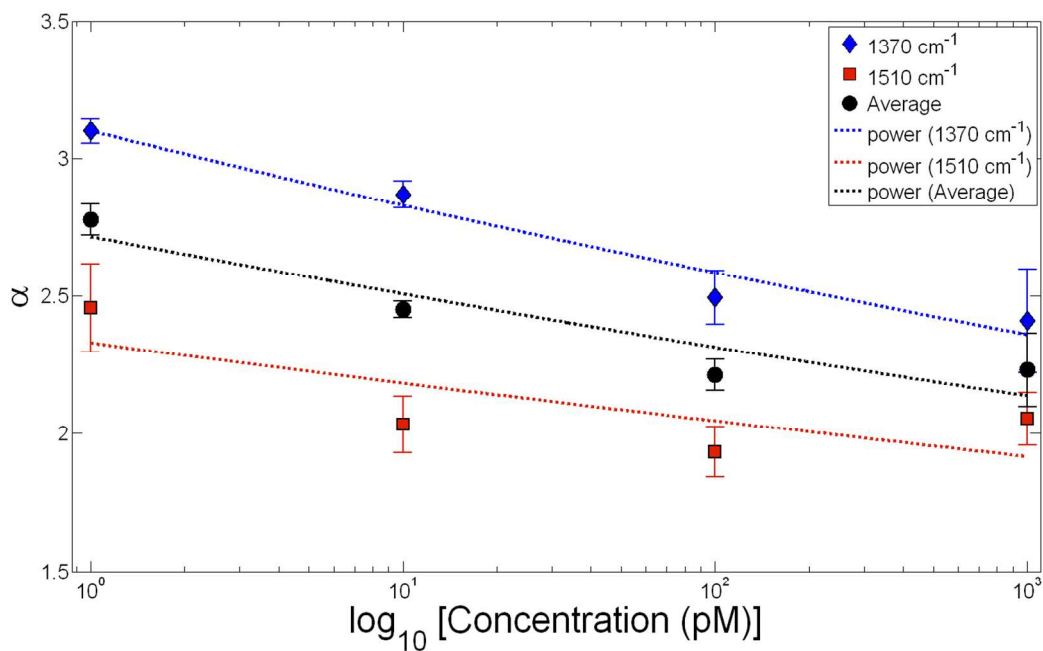


Figure S2 – Power fit exponent (α) of the SERS intensity distribution as a function of analyte concentration for the two diagnostic TAMRA peaks 1370 cm^{-1} (blue), 1510 cm^{-1} (red) as well as their averages (black). All curves are fitted with a power function (dotted lines). Note that this figure is a semi-log plot and error bars represent 2 standard deviations.

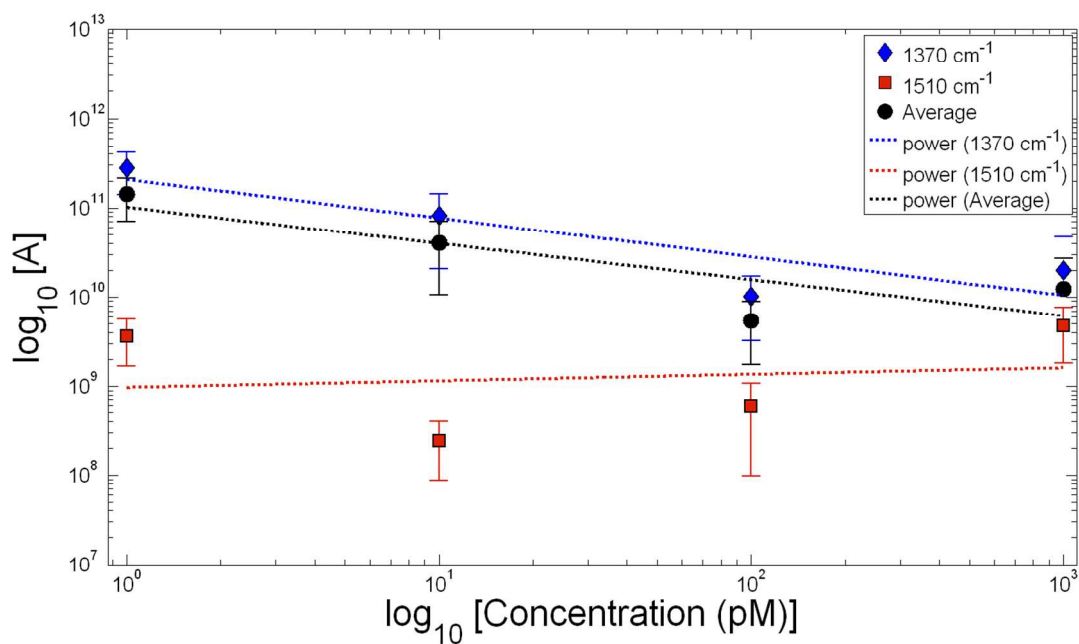


Figure S3 – Power fit of hotspot-closeness parameter (A) of the SERS intensity distribution as a function of analyte concentration for the two diagnostic TAMRA peaks 1370 cm^{-1} (blue) and 1510 cm^{-1} (red) as well as their averages (black). All curves are fitted with a power function (dotted lines). Note that this figure is a log-log plot and error bars represent 2 standard deviations.

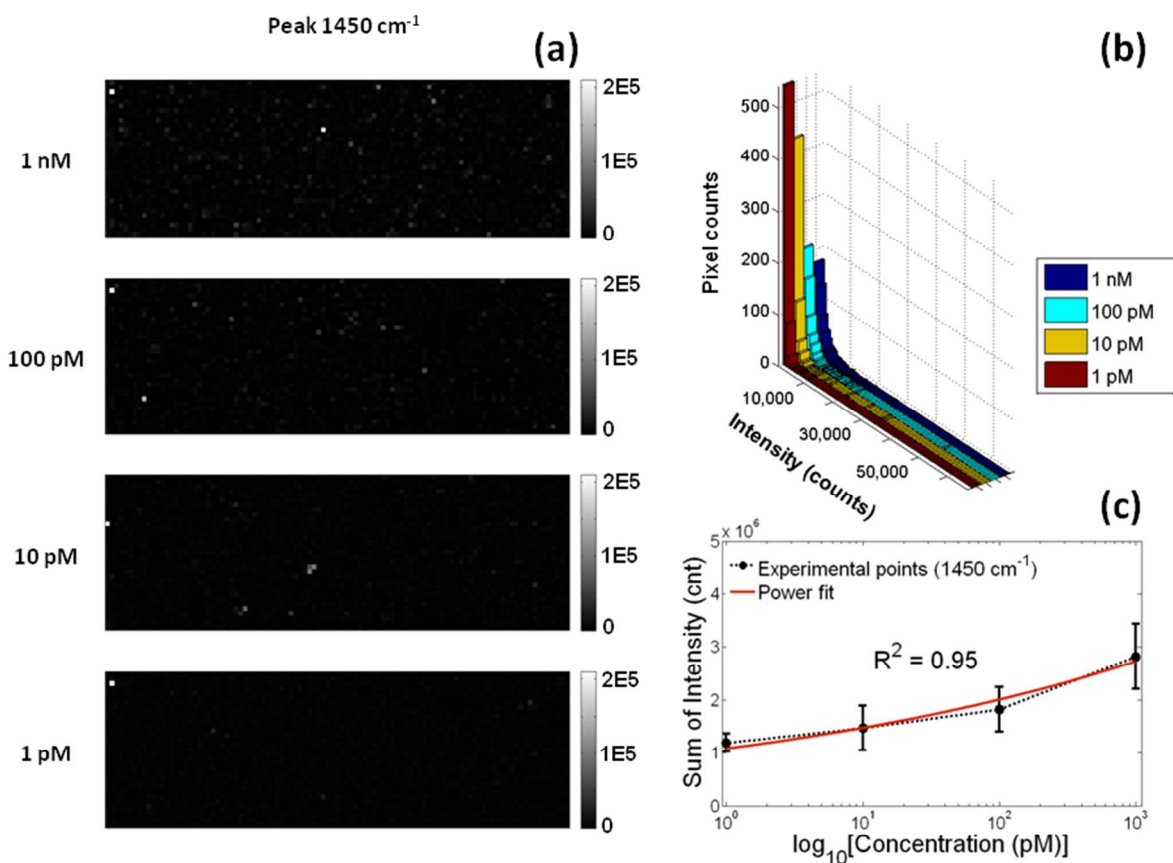


Figure S4 – (a) Intensity maps of TAMRA peak 1450 cm⁻¹ showing increased intensity in a series of mapping experiments for quantification of intensities with 1 nM, 100 pM, 10 pM and 1 pM TVP. A substrate area of about 101 $\mu\text{m} \times 33 \mu\text{m}$ is shown in each case, along with the intensity scale bar labeled with the maximum intensity values (in arbitrary units). The Raman map pixel size is 1 $\mu\text{m} \times 1 \mu\text{m}$. (b) Histogram representation of SERS intensity distributions of hotspots at different concentrations of TVP. Four histograms are displayed in overlaid configuration corresponding to various analyte concentrations. (c) Experimental hotspot intensity integral plots for the additional TVP peak 1450 cm⁻¹ as a function of analyte concentration on a semi-log scale (dotted black line). The error bars represent 2 standard deviations on three different samples for the same concentration. The curve fits a power function (solid red line).

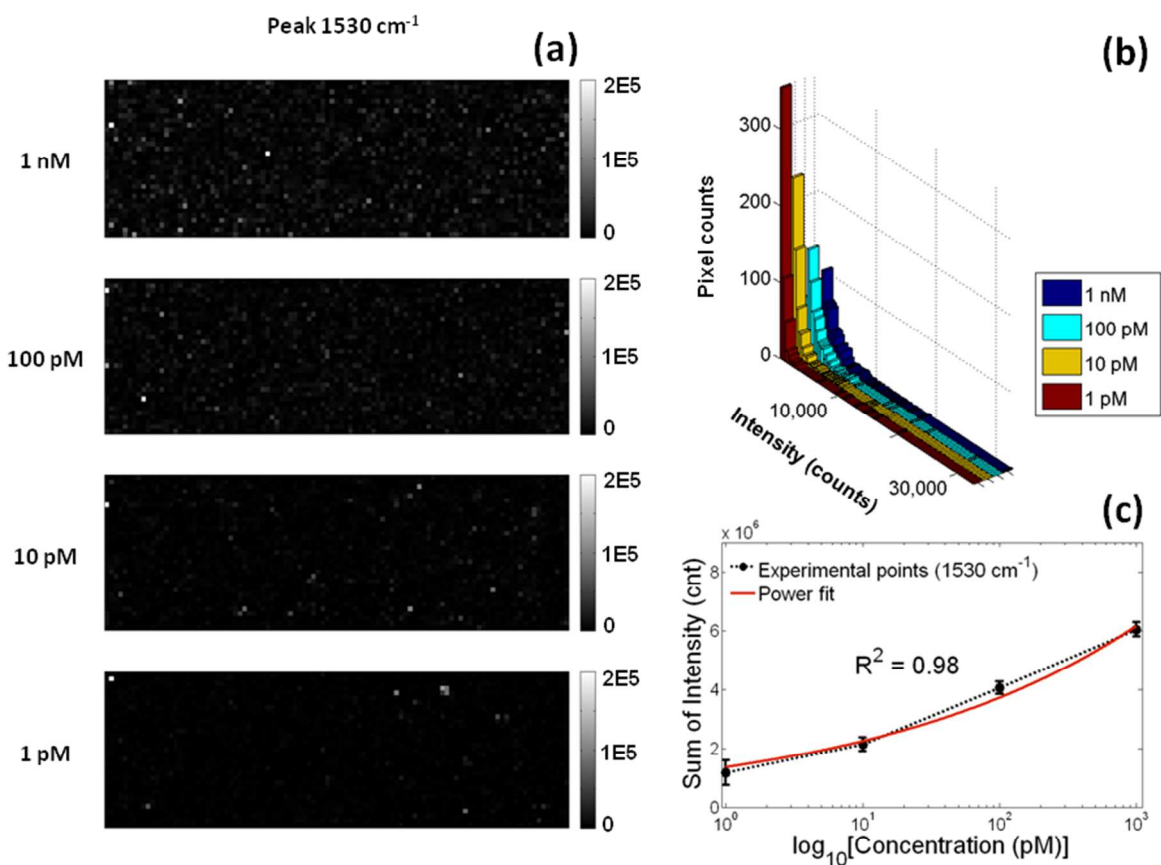


Figure S5 – (a) Intensity maps of TAMRA peak 1530 cm⁻¹ showing increased intensity in a series of mapping experiments for quantification of intensities with 1 nM, 100 pM, 10 pM and 1 pM TVP. A substrate area of about 101 μm × 33 μm is shown in each case, along with the intensity scale bar labeled with the maximum intensity values (in arbitrary units). The Raman map pixel size is 1 μm × 1 μm. (b) Histogram representation of SERS intensity distributions of hotspots at different concentrations of TVP. Four histograms are displayed in overlaid configuration corresponding to various analyte concentrations. (c) Experimental hotspot intensity integral plots for the additional TVP peak 1530 cm⁻¹ as a function of analyte concentration on a semi-log scale (dotted black line). The error bars represent 2 standard deviations on three different samples for the same concentration. The curve fits a power function (solid red line).

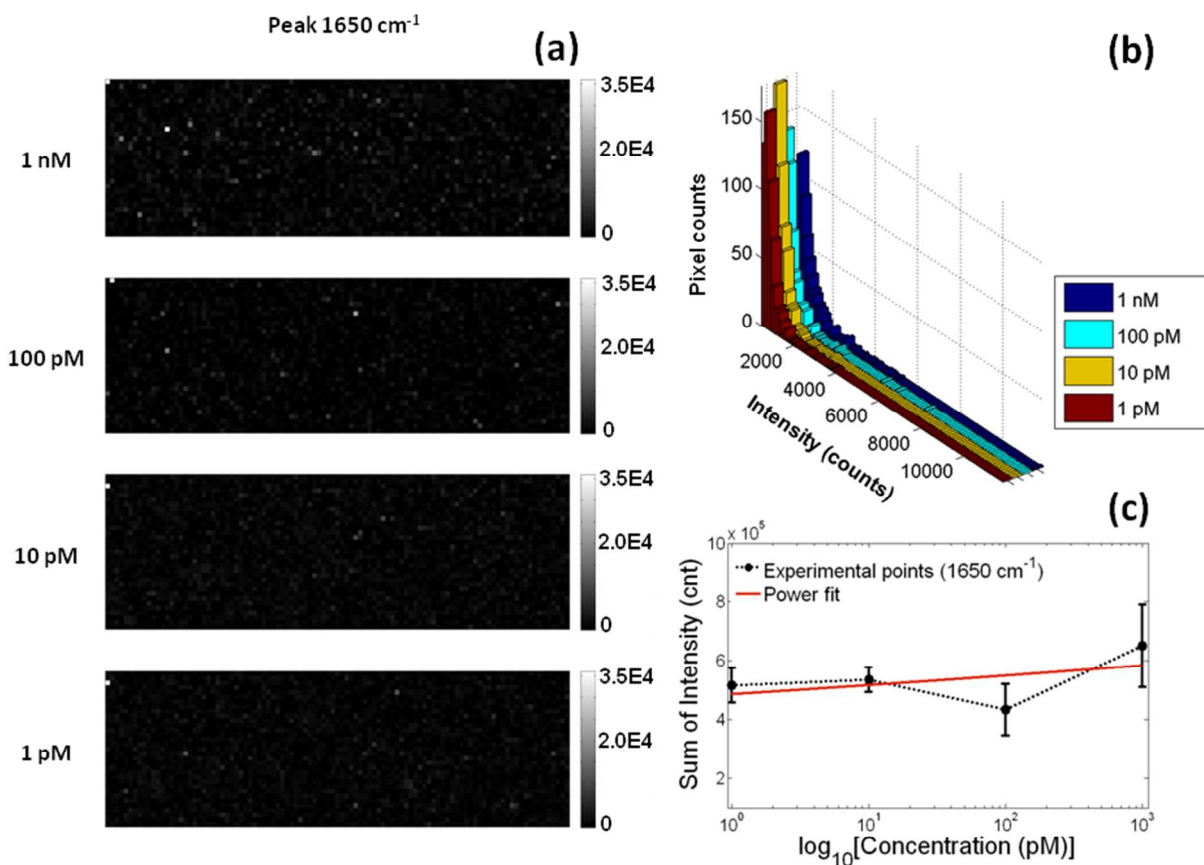


Figure S6 – (a) Intensity maps of TAMRA peak 1650 cm⁻¹ showing increased intensity in a series of mapping experiments for quantification of intensities with 1 nM, 100 pM, 10 pM and 1 pM TVP. A substrate area of about 101 μm × 33 μm is shown in each case, along with the intensity scale bar labeled with the maximum intensity values (in arbitrary units). The Raman map pixel size is 1 μm × 1 μm. (b) Histogram representation of SERS intensity distributions of hotspots at different concentrations of TVP. Four histograms are displayed in overlaid configuration corresponding to various analyte concentrations. (c) Experimental hotspot intensity integral plots for the additional TVP peak 1650 cm⁻¹ as a function of analyte concentration on a semi-log scale (dotted black line). The error bars represent 2 standard deviations on three

different samples for the same concentration. The curve does not fit a power function (solid red line) as expected potentially due to the suppression of the 1650 cm^{-1} vibrational mode.

TABLE S-I

P-values of the Two Sample Kolmogorov-Smirnov Test Evaluating SERS Mapping Consistency.

Test	Frequency	1 vs 2	1 vs 3	2 vs 3
1d-KS	1370 cm ⁻¹	0.32	0.38	0.99
1d-KS	1510 cm ⁻¹	0.30	0.99	0.29
2d-KS	Combined	> 0.2	> 0.2	> 0.2

During a mapping experiment, individual substrates were measured at three different locations, collecting 3333 spectra (Raman maps 1, 2 and 3) for each measurement. Each measurement on a substrate functionalized with vasopressin-specific aptamer (treatment) of one mapping area is compared to another mapping area of the same functionalization condition. For the treatment samples, the KS significance test verified that the measurements from the different areas were not statistically different, which confirms the reproducibility of the experimental method. The highest *p*-value is shown. In the combined row the maps for the two diagnostic peaks (1370 cm⁻¹ and 1510 cm⁻¹) have been collapsed into a composite map.

TABLE S-II

Coefficient of Determination of Pareto Fit for SERS Intensity Distribution for the Two Diagnostic TAMRA Peaks.

Diagnostic Peaks (cm⁻¹)	Concentration (pM)	Coefficient of Determination (R^2)
1370	10 ⁰	0.9768
	10 ¹	0.9882
	10 ²	0.9821
	10 ³	0.9671
1510	10 ⁰	0.9915
	10 ¹	0.9654
	10 ²	0.9812
	10 ³	0.9677

Every entry is the average of three independent measurements.

TABLE S-III

Power Fit Exponent (α) as a Function of Analyte Concentration for the Two Diagnostic TAMRA Peaks.

Peaks (cm⁻¹)	Scaling Factor (<i>b</i>)	Exponent (<i>d</i>)
1370	3.0926	-0.0389
1510	2.3051	-0.0256

The curve was fitted with a power function in general form: $\alpha(C) = bC^d$, where b and d are constant real numbers and C is a variable concentration of analyte.

TABLE S-IV

Hotspot-closeness Parameter (A) as a Function of Analyte Concentration for the Two Diagnostic TAMRA Peaks.

Peaks (cm⁻¹)	Scaling Factor (e)	Exponent (f)
1370	2.0818E+11	-0.4339
1510	9.7343E+08	0.0728

The curve was fitted with a power function in general form: $A(C) = eC^f$, where e and f are constant real numbers and C is a variable concentration of analyte.

TABLE S-V

Comparison of Experimentally and Theoretically Obtained Intensity Integral Values for the Two Diagnostic TAMRA Peaks.

Diagnostic Peaks (cm⁻¹)	Concentration (pM)	Experimental Value (cnt)	Standard Deviation (cnt)	Theoretical Prediction (cnt)	Percentage Error (%)
1370	10 ⁰	5.3043E+05	5.6975E+04	4.9024E+05	7.57 ± 8.96
	10 ¹	6.3391E+05	6.3732E+04	6.1553E+05	2.89 ± 8.87
	10 ²	9.9935E+05	7.3192E+04	8.5471E+05	14.47 ± 5.83
	10 ³	1.1153E+06	1.7055E+04	1.2421E+06	11.36 ± 1.67
1510	10 ⁰	5.6972E+05	3.0161E+04	8.0056E+05	40.51 ± 7.06
	10 ¹	1.0492E+06	2.0027E+05	9.2563E+05	11.77 ± 14.14
	10 ²	2.2775E+06	1.1520E+05	1.3303E+06	41.58 ± 2.81
	10 ³	3.0173E+06	3.4747E+05	2.5592E+06	15.18 ± 8.75

Every experimental entry is the average of three independent measurements and the standard deviation is based on these values.

TABLE S-VI

P-values for the Two Sample Kolmogorov-Smirnov Test Comparing Vasopressin-specific Aptamer and IgE-specific Aptamer Functionalization.

Test	Frequency	1 vs IgE	2 vs IgE	3 vs IgE	Combined
1d-KS	1370 cm ⁻¹	< 0.001	< 0.001	< 0.001	< 0.001
1d-KS	1510 cm ⁻¹	0.23	< 0.001	< 0.001	< 0.001
2d-KS	Combined	< 0.001	< 0.001	< 0.001	< 0.001

Each measurement on the substrate functionalized with vasopressin-specific aptamer is compared to each of the substrates functionalized with IgE-specific aptamers. The highest *p*-value is shown. For TAMRA frequency 1510 cm⁻¹ map 1 cannot be distinguished from the control samples. In the combined column the maps have been collapsed into a single measurement.

TABLE S-VII

2 Sigma Error Values as a Percentage of Unit Area Intensities.

Diagnostic Peaks (cm ⁻¹)	Concentration (pM)	2 Sigma (%)
1370	10 ⁰	1.85
	10 ¹	1.36
	10 ²	2.31
	10 ³	4.26
1510	10 ⁰	1.99
	10 ¹	1.63
	10 ²	1.13
	10 ³	1.99
Combined	10 ⁰	1.34
	10 ¹	3.04
	10 ²	1.87
	10 ³	1.90

The error bars are all less than 5% of measured data averaged over 100 randomly selected unit areas. This indicates that the uncertainty in our reported measurement is very low, thus the multimodal analysis appears to be extremely robust and accurate.

REFERECNES

1. 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*, 204701.
2. Etchegoin, P. G.; Meyer, M.; Blackie, E.; Le Ru, E. C. Statistics of Single-Molecule Surface Enhanced Raman Scattering Signals: Fluctuation Analysis with Multiple Analyte Techniques. *Anal. Chem.* **2007**, *79*, 8411-8415.
3. Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627-1639.
4. Zhang, D.; Ben-Amotz, D. Enhanced Chemical Classification of Raman Images in the Presence of Strong Fluorescence Interference. *Appl. Spectrosc.* **2000**, *54*, 1379-1383.