Supporting Information

Zavaleta et al. 10.1073/pnas.0813327106

SI Text

Comparison Between Spectroscopic Nanoimaging Techniques: Quantum Dots vs. SERS Nanoparticles. The multiplexing capabilities of each of these nanoimaging techniques was evaluated, and showed that the Maestro (CRi, Woburn, MA) system was unable to unmix four colocalized quantum dots at relatively high concentrations using its real component analysis algorithm. Instead, the software incorrectly computed that nine distinct spectra and one autofluorescence spectrum were embedded within the raw data in different areas of the mouse. To compare the in vivo multiplexing capabilities of SERS nanoparticles to quantum dots, we i.v. injected four quantum dots (Qdot 525, Qdot 655, Qdot 705, and Qdot 800) simultaneously of equal concentrations (≈8 μM). We chose these particular quantum dots because they had large spectral differences with the best chance of being spectrally unmixed. However, the Maestro system was unable to unmix the four colocalized quantum dots given at doses ≈1000 times more concentrated than our SERS nanoparticle (0.8 nM) setup, which was able to successfully unmix five different SERS nanoparticles colocalized in the liver (Fig. 3). In addition, under similar conditions and concentrations, Raman imaging in conjunction with SERS nanoparticles revealed a sensitivity of ≈8.125 pM, which represents an important advantage over the sensitivity of conventional fluoroscopy imaging devices in conjunction with quantum dots (≈11 nM using IVIS (Caliper Life Sciences, Mountain View, CA) and Maestro Imaging Systems).

Computer Methods for Discriminating Spectra. Nanoplex software assumes that the spectrum to be analyzed comes from a mixture of known Raman-active components. By this we mean that the pure spectra of all components in the mixture are known. From the physics of Raman spectroscopy, this means that the spectrum to be analyzed is, at least approximately, a weighted sum (that is, a linear combination or superposition) of these pure spectra. Treating the spectra as vectors, this means that we can bring the powerful machinery of linear algebra to bear on the problem. If **S** represents the captured spectrum to be analyzed and $\mathbf{C_1} \ldots \mathbf{C_n}$ are the known pure components, the algebraic statement is:

$$\mathbf{S} = w_1\mathbf{C}_1 + w_2\mathbf{C}_1 + \ldots + w_n\mathbf{C}_n + \mathbf{R} \qquad [1]$$

where **R** is the *residual* error, which we are trying to minimize for the optimal for choices of $w_1 \ldots w_n$.

The Nanoplex software component listbox contains the components and their properties. There are two types of components: taggants and backgrounds. A taggant is typically the pure component spectra of a Nanoplex SERS nanoparticle, and a background is something behind or containing the taggant such as the Raman signal of the mouse tissue.

As mentioned before, Nanoplex software uses direct classical least squares (DCLS) regression. Equation (**1**) in the previous section is a regression equation, and the DCLS solution of the equation is the vector of weights (*aka* signals) (w_1, w_2, ..., w_n) that minimizes the L_2 norm of **R**. This is the same as minimizing the sum of the squares of the coordinates of **R**. Solving for these DCLS weights is very quick. The weighted sum of the pure components = **S–R** is referred to as the *best fit* to **S**. The vectors $\mathbf{C_1} \ldots \mathbf{C_n}$ **are** called the *regressors*. If **R** is 0, then the fit is perfect, but this can never happen in practice (unless **S** is actually one of the regressors). Nanoplex software reports the quality of the fit as

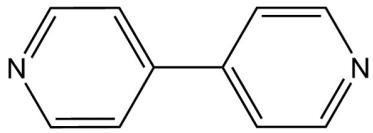
$$\text{Relative Fitting Error} = \text{norm2}(\mathbf{S}-\mathbf{R})/\text{norm2}(\mathbf{S})$$

The Relative Fitting Error (**RFE**) is a number between 0 and 1, but Nanoplex software reports it as a percentage. Geometrically, it is the sine (angle between **S** and the subspace generated by $\mathbf{C_1} \ldots \mathbf{C_n}$). The more **S** is separated from this subspace, the larger the angle, and the larger the **RFE**.

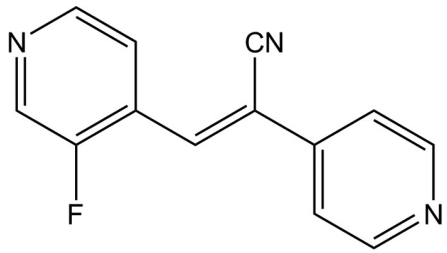
Polynomial Regressor. Eq. **1** above is actually an oversimplification. In practice, we have found that the captured spectrum **S** often has these two types of *additive disturbances*: instrumental intensity (Y) drift (including a constant offset) and uncalibrated chemicals that produce a slowly changing spectrum (such as autofluorescence of murine tissue). Fortunately, these disturbances are *peak free*, so we can model them with well-understood peak-free functions. Nanoplex software uses polynomials for this, making the regression equation:

$$\mathbf{S} + w_1\mathbf{C}_1 + w_2\mathbf{C}_1 + \ldots + w_n\mathbf{C}_n + c_o\mathbf{P}_0 + \ldots + c_d\mathbf{P}_d + \mathbf{R} \qquad [2]$$

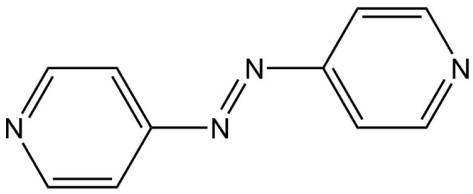
where $\mathbf{P_i}$ is a polynomial of degree i, and $\mathbf{P_0}, \ldots, \mathbf{P_d}$ are orthogonal on the selected wave number range. Thus the linear combination of these polynomials is also a polynomial of degree d. Eq. **2** is also easily solved by DCLS, just like Eq. **1** above.



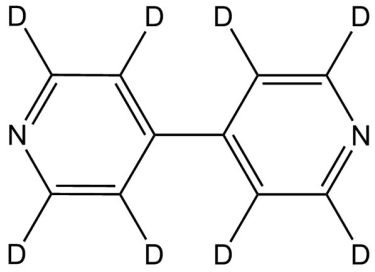
S420
4,4'-dipyridyl



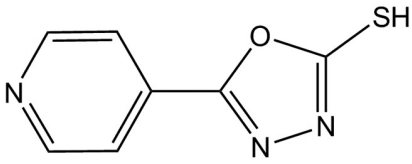
S466
1-(4-pyridyl)-1-cyano-2-(2-fluoro-4-pyridyl)-ethylene



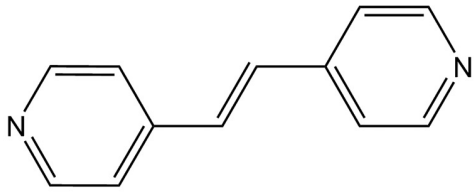
S481
4-Azobis(pyridine)



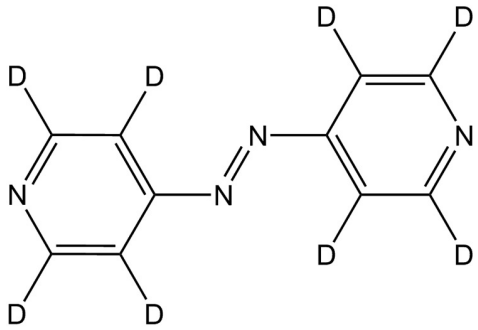
S421
d8-4,4'-dipyridyl



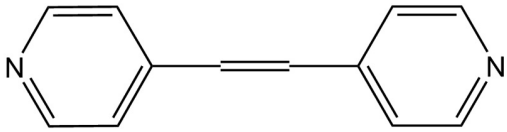
S403
5-(4-pyridyl)-1,3,4-oxadiazole-2-thiol



S440
Trans –1,2-Bis(4-pyridyl)-ethylene



S482
d8-4-Azobis(pyridine)



S470
1,2-di(4-pyridyl) acetylene

Fig. S1. Chemical structures of eight of the 10 Raman active tags adsorbed onto the gold core of the SERS nanoparticles used in this study. Each molecule has a different configuration of bonds that vibrate differently, resulting in various Raman spectra ideal for multiplexing. SERS 421 is the deuterated form of SERS 420, and SERS 482 is the deuterated form of SERS 481, where D = deuterium. The chemical structures of S663 and S661 are proprietary and therefore are not disclosed here.