APPENDIX A

Supporting Information in Chapter 2

A.1 Figures

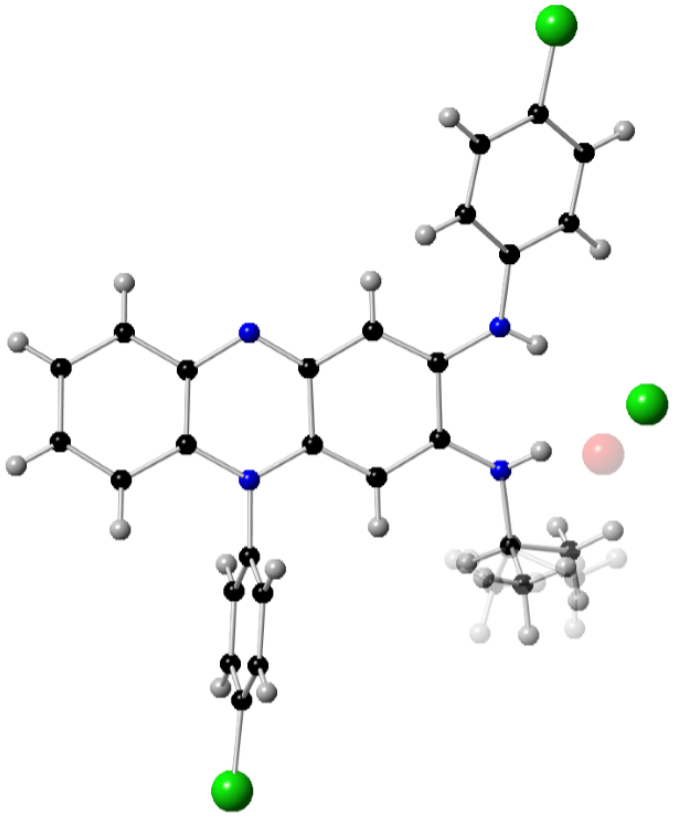


Figure S-1. Asymmetric unit of CFZ-HCl including the water molecule.

The disorder in the propyl group showed as a transparent overlay. The water has ~13% occupancy and the presence of water is likely due to the solvent used in the crystallization solution or HCl.

A.2 Tables

Table S-1. Crystallographic data for CFZ-HCl.

Empirical formula	$C_{27}H_{23.26}Cl_3N_4O_{0.13}$
Formula weight (mg/ml)	512.18
Temperature (K)	100
Wavelength (Å)	1.54178
Crystal system	Orthorhombic
Space group	Pbca
a (Å)	10.266
b (Å)	19.828
c (Å)	24.156
α (°)	90
β (°)	90
ϒ (°)	90
Volume ($Å^3$)	4917.1
Z	8
Density (calc., mg/m^3)	1.384
Absorption coefficient (mm^{-1})	3.562
F(000)	2122
Crystal size (mm^3)	0.758 x 0.269 x 0.014
Theta range for data collection (°)	3.659 to 68.333
Index ranges	-12<=h<=12 -23<=k<=21 -29<=l<=29
Reflections collected	28160
Independent reflections	4500 [R(int) = 0.0352]
Completeness to theta max	99.8 %
Absorption correction	Integration
Max. and min. transmission	0.9615 and 0.3840
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	4500 / 134 / 365
Goodness-of-fit on F^2	1.069
Final R indices [I>2sigma(I)]	R1 = 0.0392, wR2 = 0.0982
R indices (all data)	R1 = 0.0422, wR2 = 0.1010
Extinction coefficient	0.00261
Largest diff. peak and hole ($e.Å^{-3}$)	0.685 and -0.415

APPENDIX B

Supporting Information in Chapter 3

B.1 Supporting Methods

B.1.1 Elaborated Experimental Setup and Raman Microanalysis

Our protocol prepares cells as thin dry dispersions of cellular material on the surface of chips. Multiple preparation parameters were tested, such as chip-washing solutions and drying conditions (i.e., air-dried vs. freeze-dried), and the optimal procedure was used for analyses and reported herein. The Raman image acquisition methodology (continuous area scan) was optimized to minimize any photo-degradation of cell samples during long duration imaging; photo-degradation during image acquisition for this method presents as a sequential decrease in Raman signal-to-noise ratio across the area of the cell, which we did not observe with our optimized acquisition conditions. Based on our Raman imaging results of untreated macrophages (Figure 2A), the nuclei of each cell are consistently distinguishable from the cytoplasm and the presence of lipid-rich inclusions was detected in a subpopulation of cells. Based on these observations, we believe that the major structural features of interest (i.e., nucleus, cytoplasm, and lipid-rich inclusions) are conserved via our preparation protocol. Regarding alteration of spectral features: our pure component reference spectra were dissolved in appropriate solvent (with residual water) and allowed to air dry on the surface of silicon chips before acquisition of Raman spectra. In this way, the spectral features of our reference components closely match those acquired from the actual cell samples.

B.1.2 Justification for Least Squares Regression Model

While PCA allows for mathematical decomposition of highly complex cell Raman spectra, the reduction of data dimensions often leaves the resulting principal component spectra unintelligible in terms of cell biology principles; because of this, it is particularly difficult to quantitatively extract the biochemical information of specific interest for the purposes of this project (Y.H. Ong et al. Optics Express. 20 (2012) 22158). Additionally, we have shown that least squares regression modelling with protein, lipid, and nucleic acid accounts for >90% of the spectral variance observed throughout most cells with the expected spatial segregation of these three components in nuclear and cytoplasmic regions, indicating that our statistical approach is indeed yielding useful and interpretable, spectral information. Least squares regression assumes that the acquired cell spectrum is a linear summation of the most abundant components present in sample (protein, lipid, and nucleic acid in the case of a cell) as expected based on the fundamental theory of Raman spectroscopy (K. Galler et al. Sci. Rep. 6 (2016) 24155).

B.1.3 Elaborated Stoichiometric Calibration Methodology

We generated quantitative calibration standards using mixtures of pure drug and phospholipid. The pure components were dissolved in methanol:water solution (5:1 volumetric ratio) and mixed according to stoichiometric ratios; water was included to mimic the presence of residual water in the BAL cell preparations on silicon. A small volume (~2.5uL) of each calibration solution was spotted onto separate silicon chips and the methanol was allowed to evaporate, leaving a thin film of mixed material deposited across the surface of the chip. The acquired calibration spectra were preprocessed as per same parameters as the cell sample spectra. Least squares regression modelling was used to quantitatively interpret the acquired calibration spectra and extract the % signal for each component (i.e., phospholipid and drug). The calculated % drug signal was shown to be linearly correlated with the actual mass ratio of drug in phospholipid. The acquired cell sample spectra were deconvoluted using regression modelling and the

% signals were obtained for each component (i.e., protein, lipid, nucleic acid, and drug). The % drug signal was then translated into % mass (wt/wt) according to the calibration curve. The actual mass ratios were calculated and reported for each to illustrate how our method can be used to infer the manner in which drug molecules may be distributing inside cells, which is also consistent with previously measured intracellular distribution properties of the test compounds.

B.2 Figures

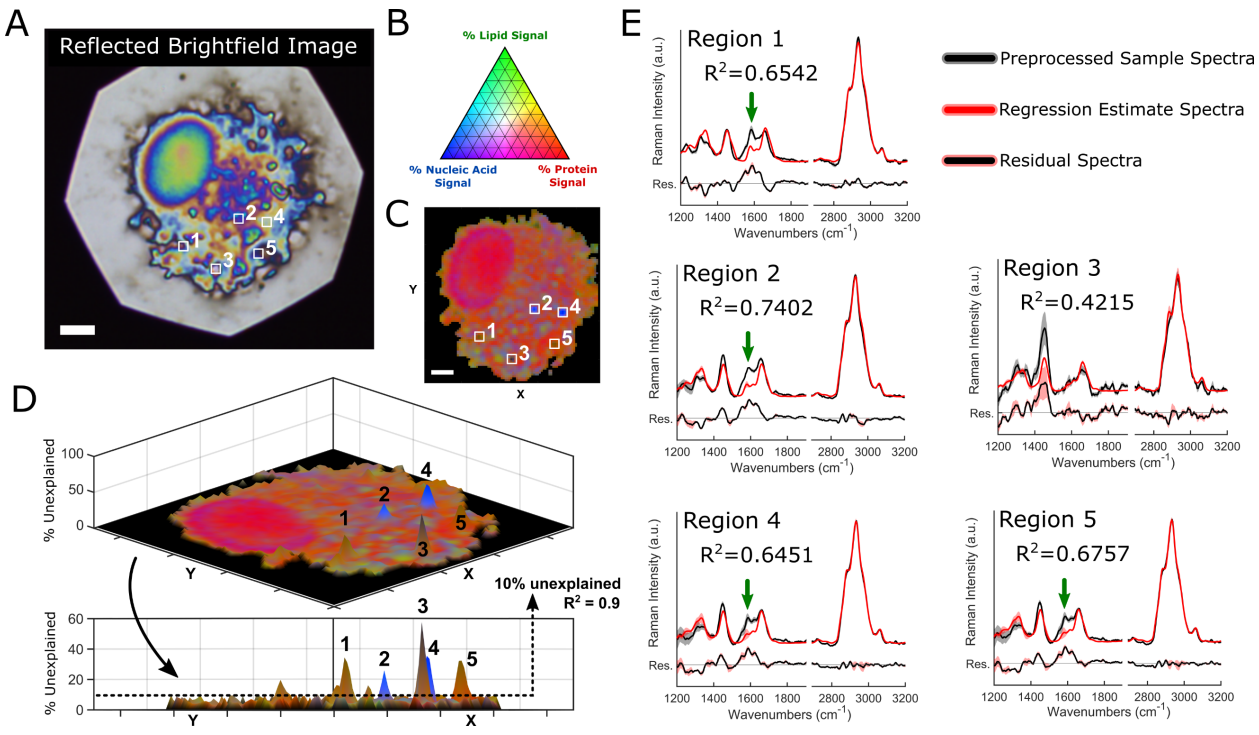


Figure S-2. Residual spectra analysis of single-cell Raman images.

(A) Reflected brightfield image of the sample scan area. Scale bar: 5 μm. **(B)** Pseudo-color ternary scale for visualization of component signals (%) throughout Raman image. **(C)** Compiled Raman image results for the analyzed cell. Scale bar: 5 μm. **(D)** Spectral analysis of cytoplasmic inclusions which yielded Raman signals that were unaccounted for by the regression model reference spectra; suggestive of unique chemical composition. **(E)** Extracted average preprocessed sample spectra, regression estimate spectra, and residual spectra from indicated regions of unexplained composition (standard deviation shown by shadow) with 1585 cm^{-1} Raman band indicated by green arrow.

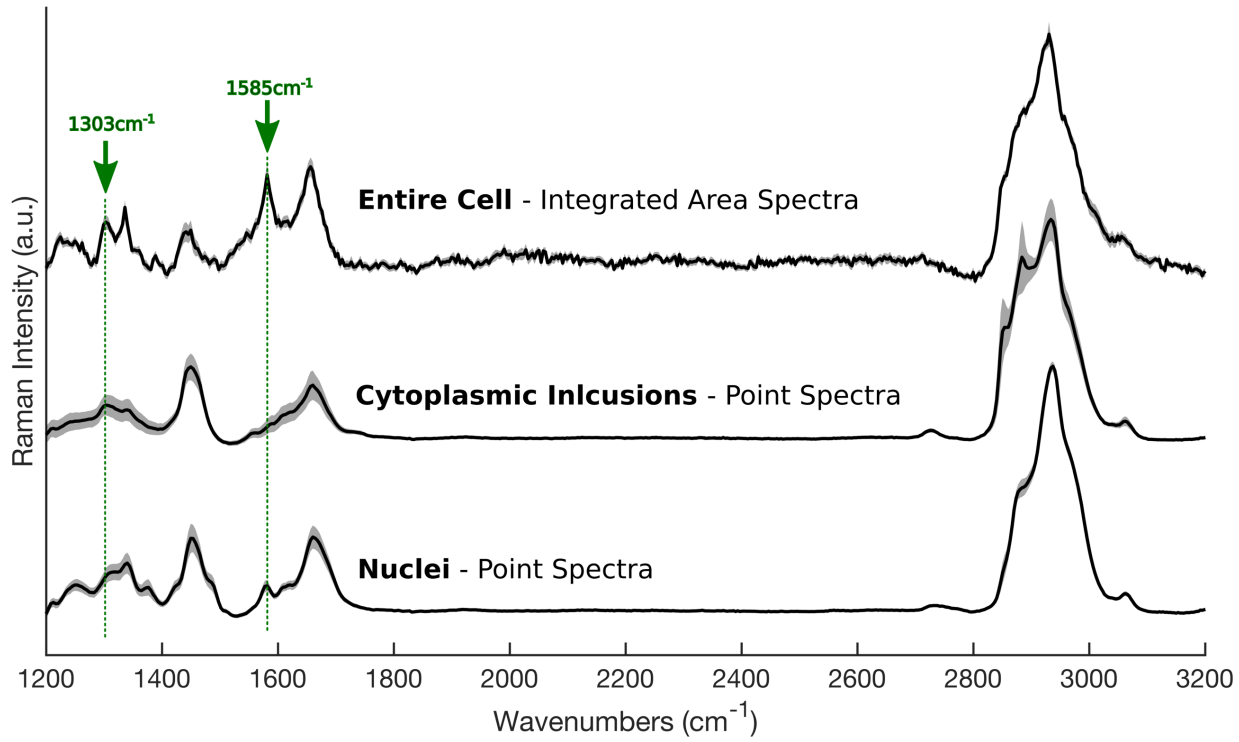


Figure S-3. Cytochrome spectral contributions revealed by integrated area scans.

Integrated Raman spectra acquired from across entire cell area (for each cell analyzed) reveal presence of cytochrome c spectral contributions at ~1303cm^{-1} and 1585cm^{-1} which were not present in point spectra acquired from cytoplasmic inclusions or nuclei of cells. Standard deviation shown by shadow (n=20 for integrated area spectra; n=120 point spectra from each group).

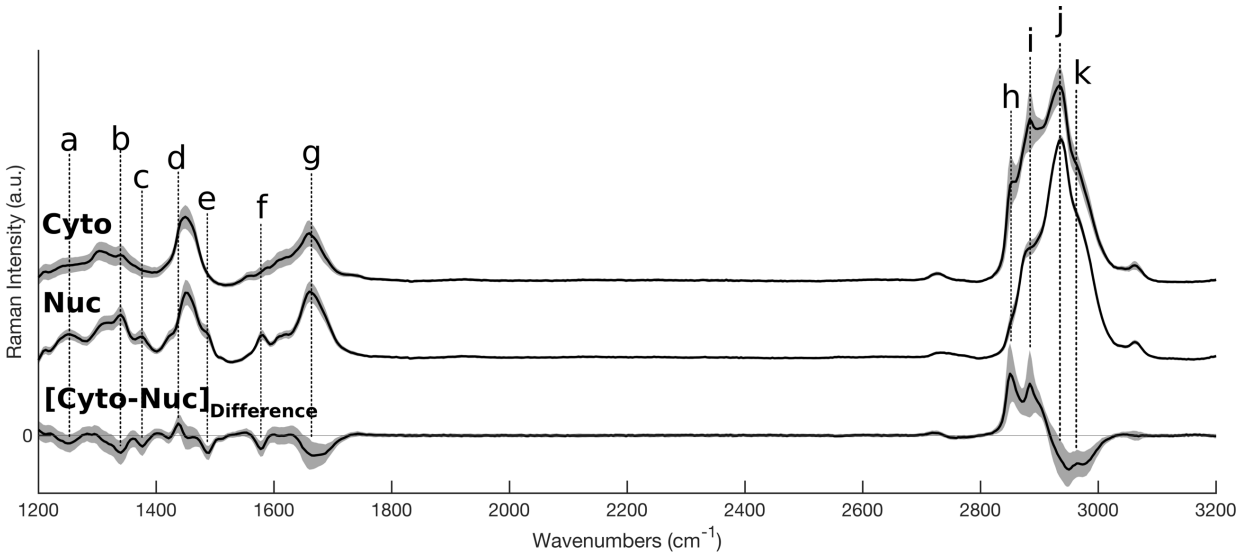


Figure S-4. Spectral differences between nuclei and cytoplasmic inclusions.

Raman point spectra reveal differences in chemical composition of nucleus and cytoplasmic inclusions in 120 individual alveolar macrophages; the major differences were attributed to vibrational modes of the most abundant biochemical components comprising typical eukaryotic cells. Standard deviation shown by shadow (n=120 point spectra from each group.

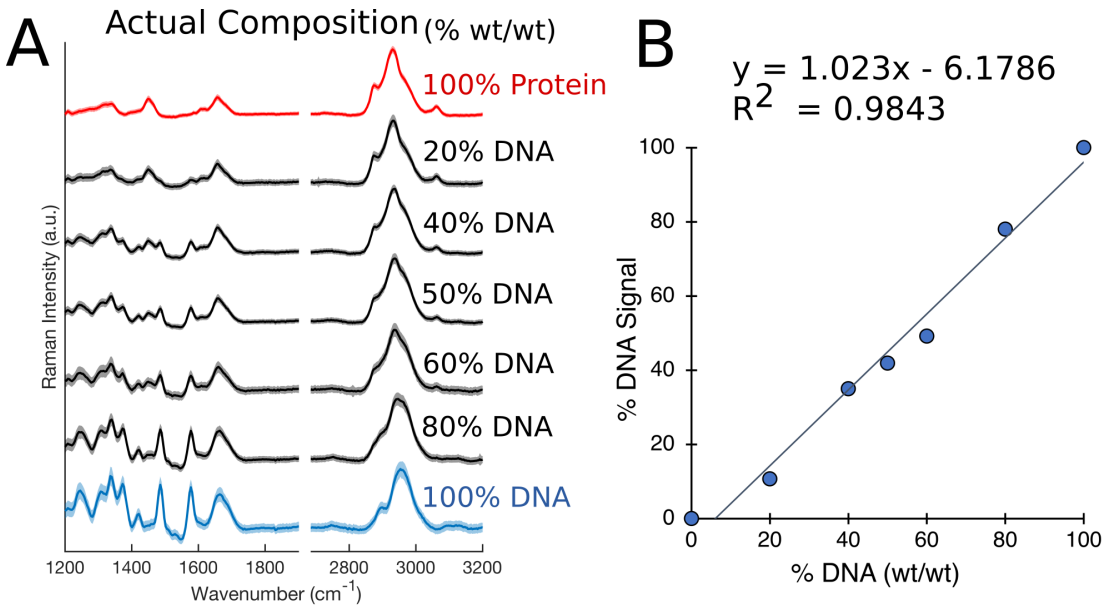


Figure S-5. Stoichiometric mixtures of protein and DNA.

Stoichiometric mixtures of protein and DNA enable quantitative determination of the relative mass ratio of protein and nucleic acid within isolated alveolar macrophages. **(A)** Overlay of Raman spectra acquired from each of the calibration samples (standard deviation shown by shadow). **(B)** Calibration curve showing strong correlation between % DNA signal acquired from regression modelling and actual % DNA (wt/wt) present in standard.

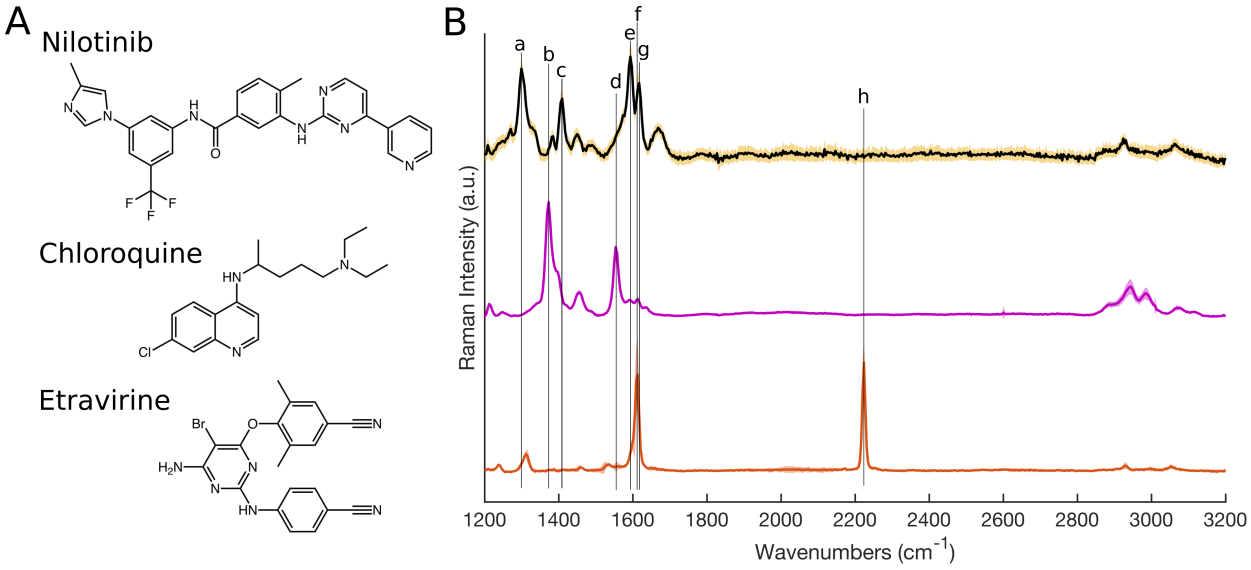


Figure S-6. Xenobiotic compound reference spectra.

(A) Molecular structures and **(B)** corresponding Raman spectra for xenobiotic compounds of interest used in the study with characteristic Raman shifts marked (a-h) for each compound (corresponding to tags listed in Table S-4).

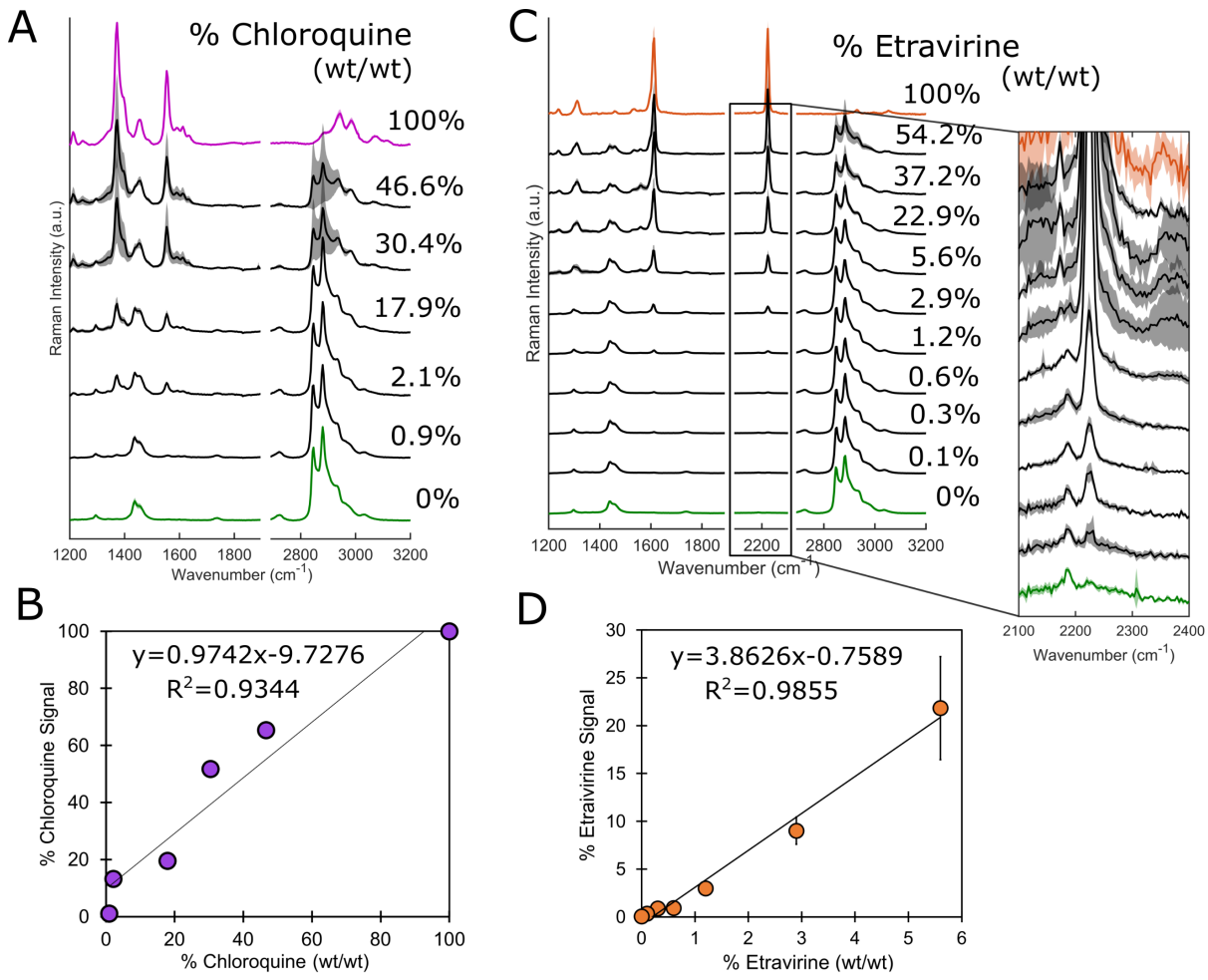


Figure S-7. Stoichiometric mixtures of pure drug and phospholipid enable quantitative interpretation of Raman spectra.

(A) Overlay of Raman spectra acquired from each of the calibration samples for chloroquine (standard deviation shown by shadow). **(B)** Calibration curve showing strong correlation between % chloroquine signal acquired from regression modelling and actual % chloroquine (wt/wt) present in standard. **(C)** Overlay of Raman spectra acquired from each of the calibration samples for etravirine (standard deviation shown by shadow). **(D)** Calibration curve showing strong correlation between % etravirine signal acquired from regression modelling and actual % etravirine (wt/wt) present in standard.

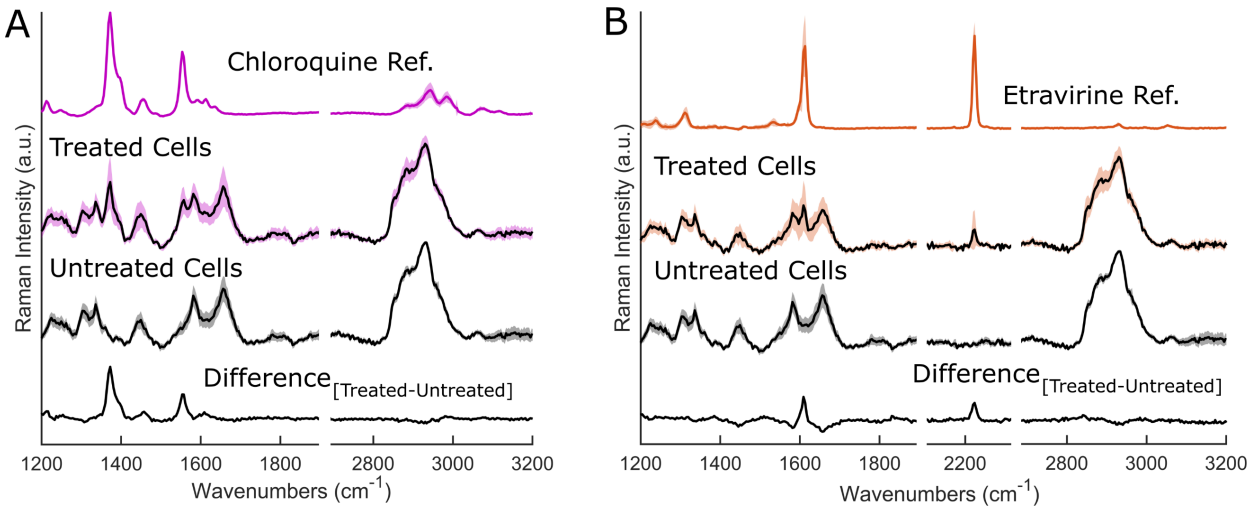


Figure S-8. Screening cells for chloroquine and etravirine accumulation.

Integrated area scans allow microanalysis on a single-cell basis and enable characterization of chemically-heterogeneous cell populations; difference spectra reveal spectral contributions from characteristic peaks of sequestered xenobiotics for **(A)** chloroquine-treated cells and **(B)** etravirine-treated cells.

B.3 Tables

Table S-2. Tentative Raman vibrational assignments for most abundant biomolecules present in typical eukaryotic cells. Tags correspond to Raman shifts marked in Figure S-2.

tag	Raman shift (cm^{-1})	vibrational assignment	biochemical component
a	1250	G, C (NH_2)	nucleic acid
b	1342	G, CH_3CH_2 wagging	nucleic acid, protein
c	1370	T, A, G (ring breathing)	nucleic acid
d	1437	CH_2 deformation	lipid
e	1485	G, A (ring breathing)	nucleic acid
f	1579	G, A, pyrimidine ring	nucleic acid
g	1655	Amide I, υ(C=O)	nucleic acid, protein
h	2850	υ_s(CH_2), fatty acids	lipid
i	2884	υ_s(CH_2), fatty acids	lipid
j	2933	C-H vibrations	protein
k	2956	υ_{as}(CH), deoxyribose	nucleic acid

Table S-3. Relevant physicochemical properties of xenobiotic compounds of interest used in bioaccumulation study.

compound	drug class	logP	amine pKa(s)	solubility (mg/mL)	volume of distribution, *Vd* (L/kg)
nilotinib	tyrosine kinase inhibitor	4.51	4.2, 5.92	0.0020	3
chloroquine	antimalarial	5.28	7.29, 10.32	0.0106	>100
etravirine	non-nucleoside reverse transcriptase inhibitor	5.54	3.49	0.0169	5.6

Table S-4. Tentative characteristic vibrational assignments for xenobiotic compounds of interest used in study. Tags correspond to Raman shifts marked in Figure 3-3A and S-6B.

tag	Raman shift (cm^{-1})	vibrational assignment	compound
a	1299	C-N stretching	nilotinib
b	1373	C-C stretching (quinoline ring)	chloroquine
c	1409	C-H wag	nilotinib
d	1554	C=C stretching	chloroquine
e	1593	υ_s(C=N), aromatic ring	nilotinib
f	1613	υ(CC), aromatic ring chain vibrations	etravirine
g	1617	υ(CC), aromatic ring chain vibrations	nilotinib
h	2224	υ(C≡N)	etravirine

APPENDIX C

Supporting Information in Chapter 4

C.1 Figures

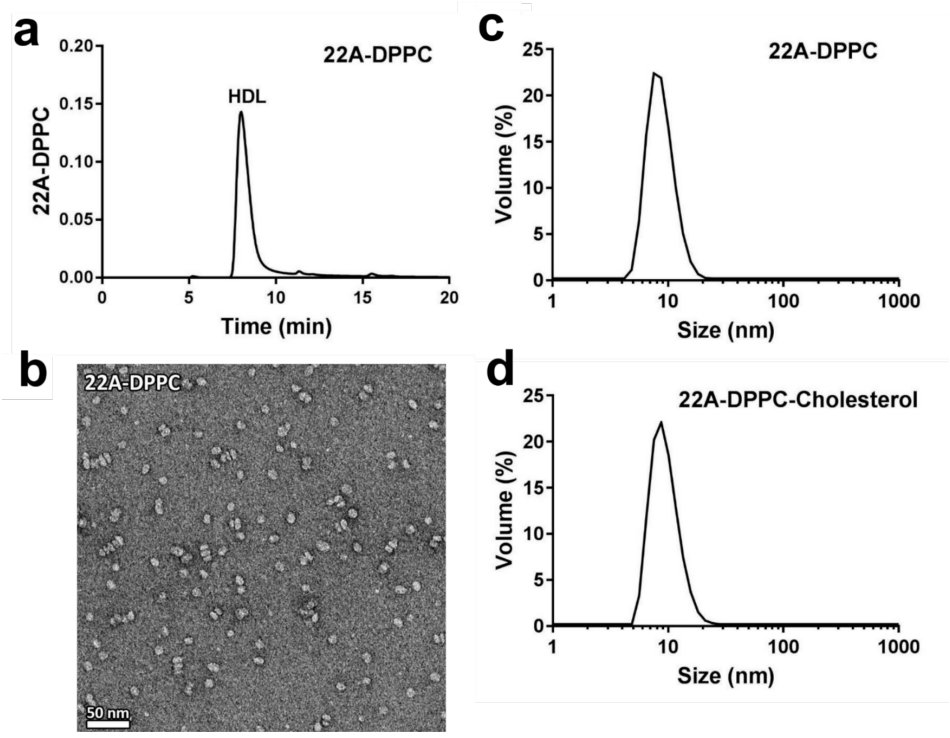


Figure S-9. Characterization of 22A-DPPC and 22A-DPPC-cholesterol HDLs.

(a) The purity of synthetic HDL formulations was assessed using gel permeation chromatography, **(b)** morphology was analyzed by transition electron microscopy, and **(c,d)** particle size distribution was determined by dynamic light scattering; particle sizes were calculated as volume intensity averages (n=3).

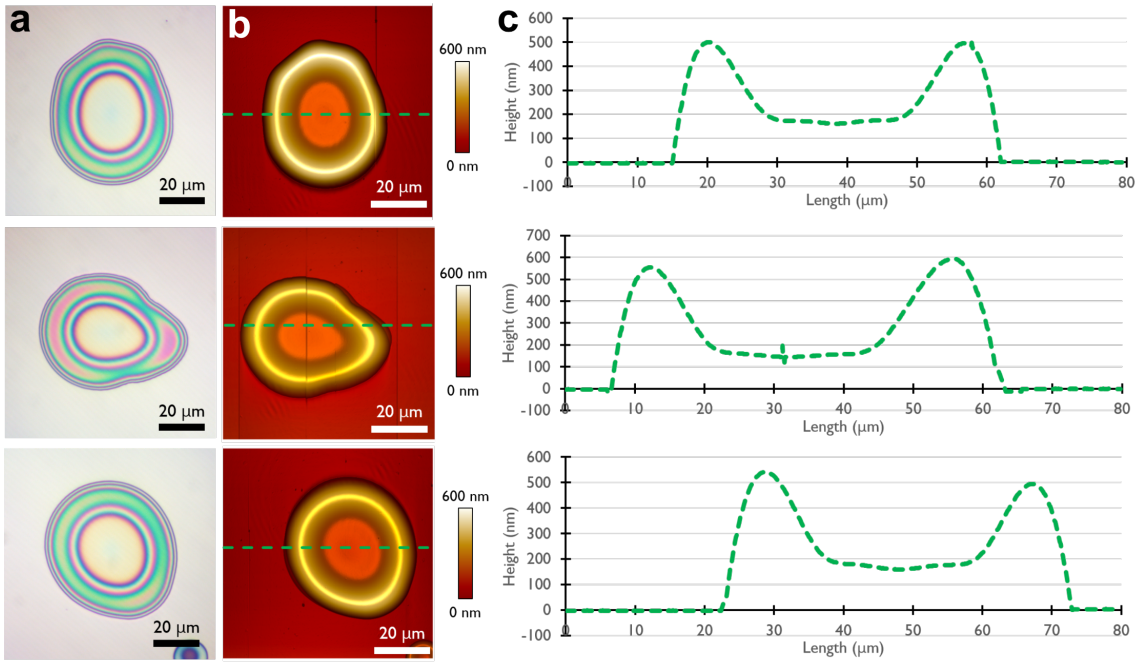


Figure S-10. AFM imaging of 1000pg protein micro-calibration standards.

(a) Reflected brightfield images of 1000pg albumin micro-calibration standards on silicon chips and **(b)** corresponding AFM topographical profiles with cross-section marked in green dashed line **(c)**.

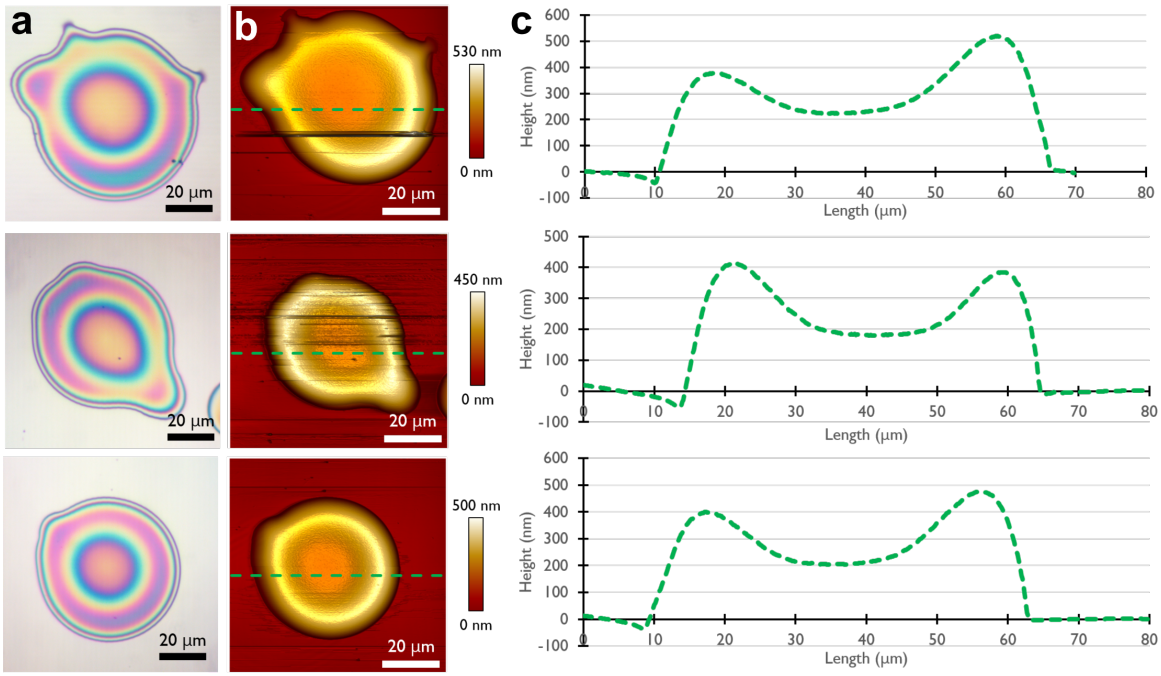


Figure S-11. AFM imaging of 900pg HDL micro-calibration standards.

(a) Reflected brightfield images of 900pg HDL micro-calibration standards (each containing 600pg of lipid and 300pg of peptide 22A) on silicon chips and **(b)** corresponding AFM topographical profiles with cross-section marked in green dashed line **(c)**.

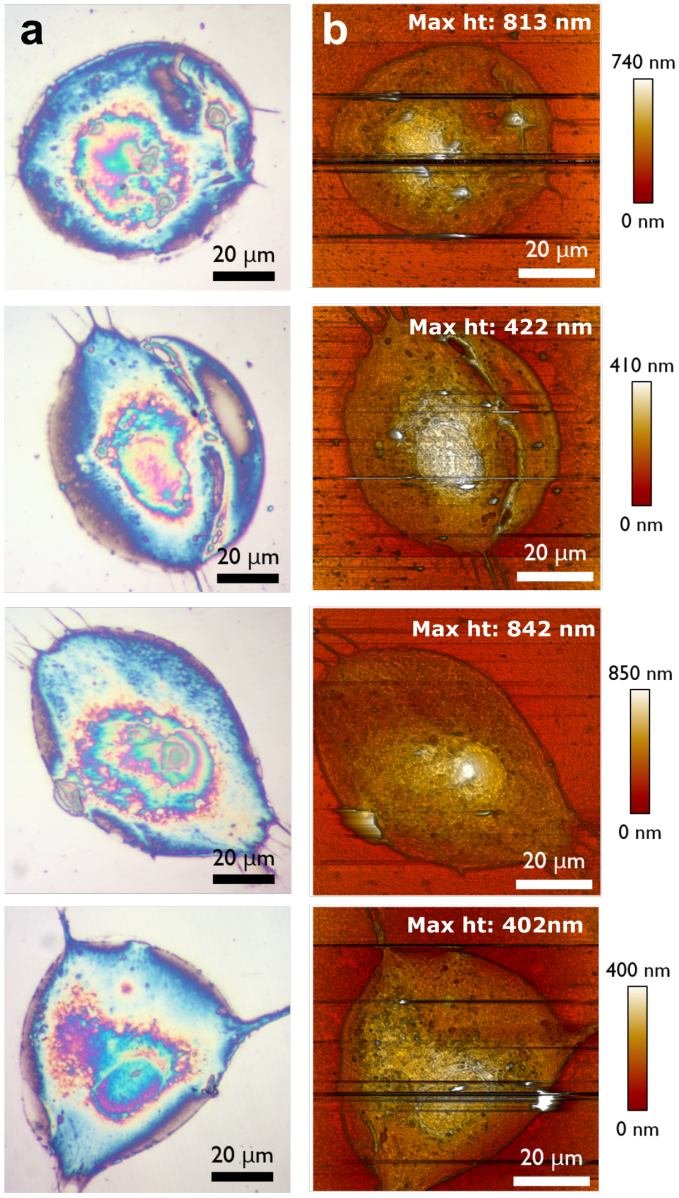


Figure S-12. AFM imaging of human skin fibroblast cell preparation.

(a) Reflected brightfield images of fibroblast cell preparations on silicon chips and **(b)** corresponding AFM topographical profiles.

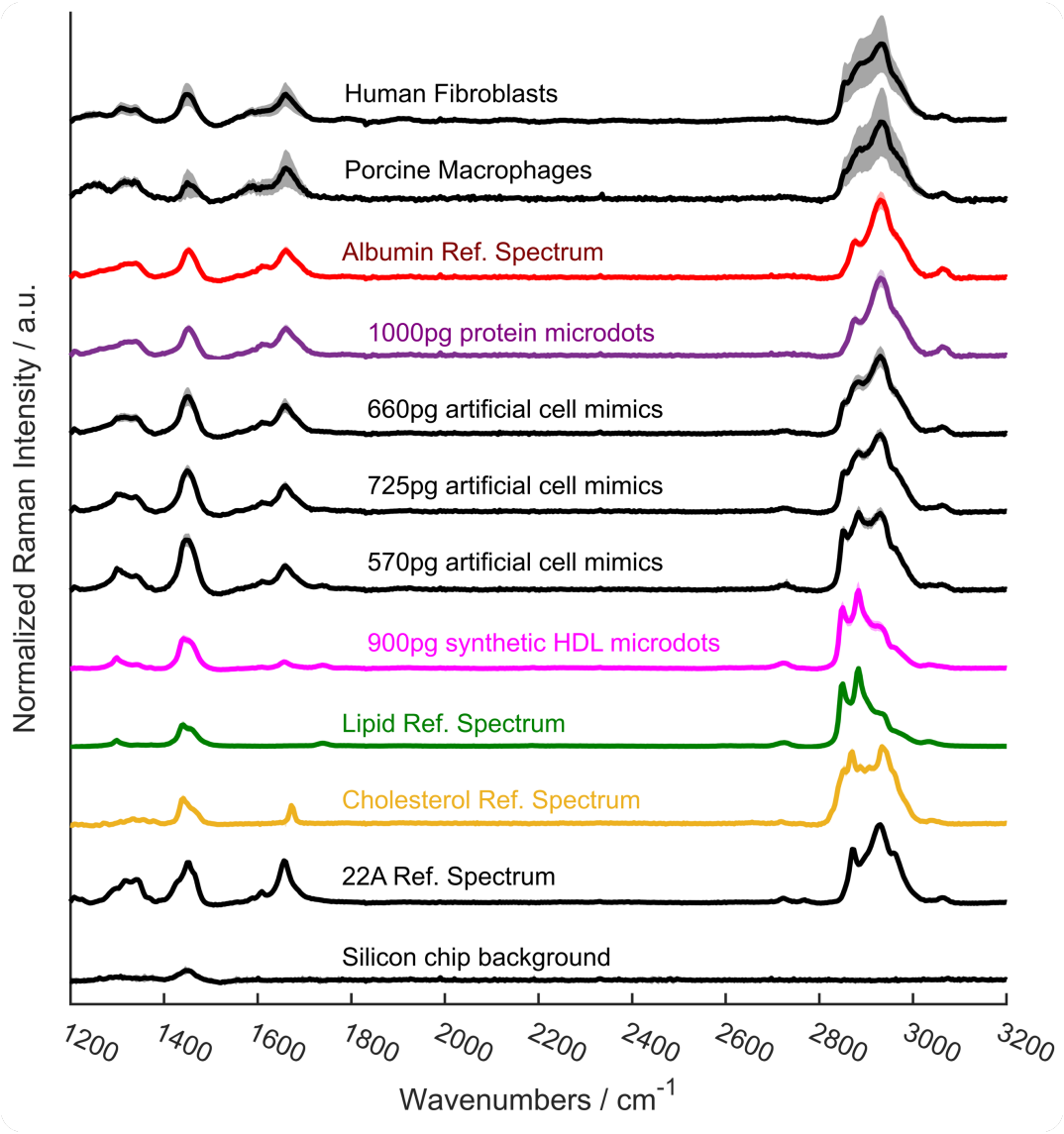


Figure S-13. Spectral overlay of cells, micro-calibration standards, and reference components.

Raman spectral overlay of all major signal contributors including cell samples, pure component reference materials, and calibration micro-calibration standards; spectra shown are averages with standard deviation shown in shadow.

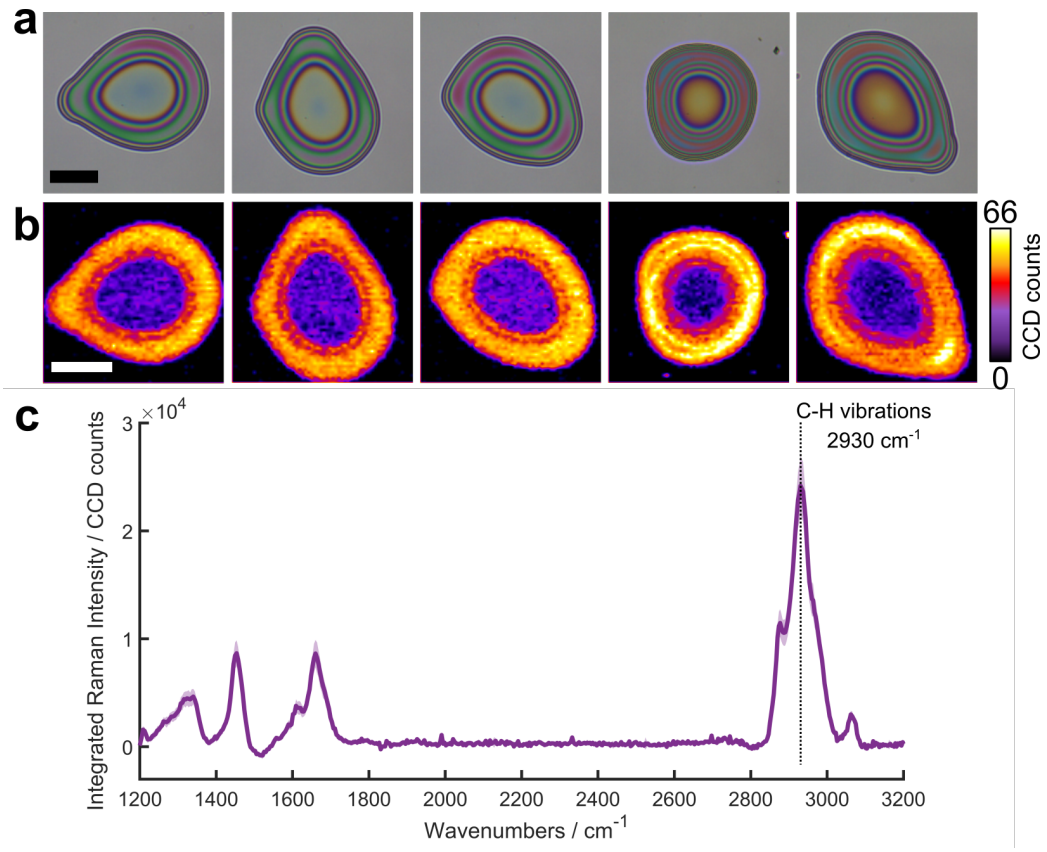


Figure S-14. Raman images of 1000pg protein micro-calibration standards.

(a) Reflected brightfield images of 1000pg albumin micro-calibration standards on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 1000pg micro-calibration standard dataset.

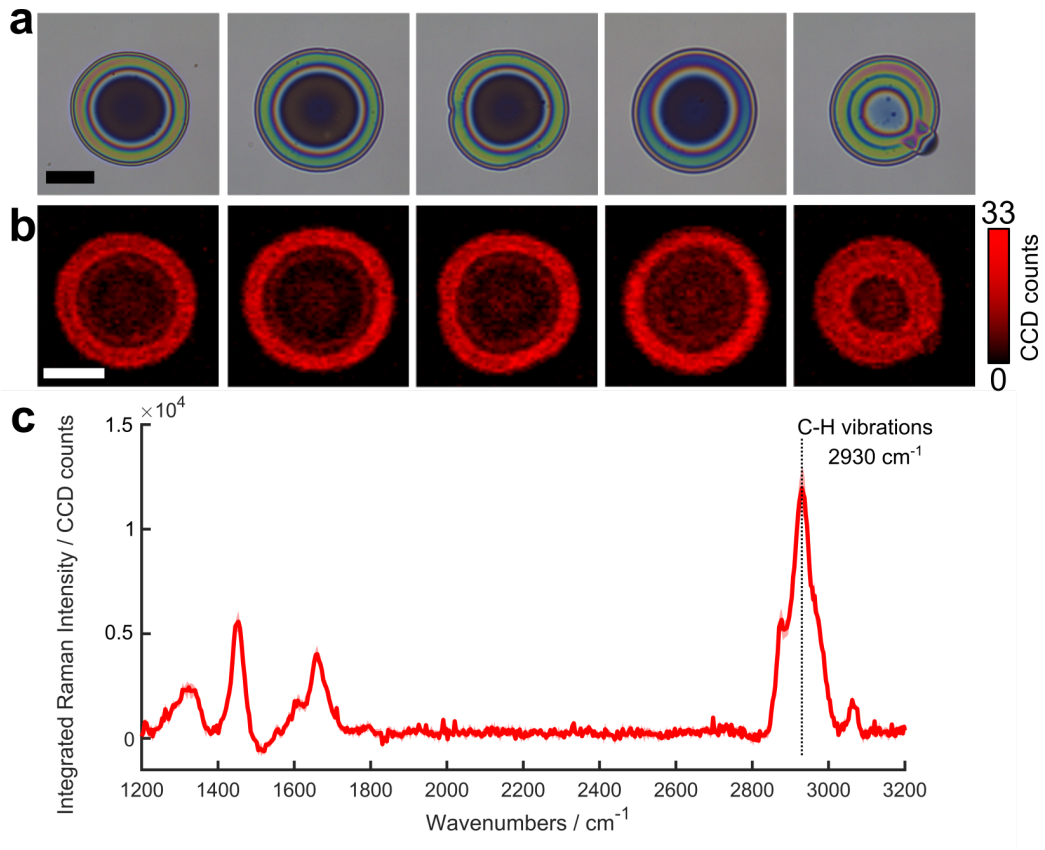


Figure S-15. Raman images of 500pg protein micro-calibration standards.

(a) Reflected brightfield images of 500pg albumin micro-calibration standards on silicon chips and **(b)** corresponding Raman area scans showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 500pg micro-calibration standard dataset.

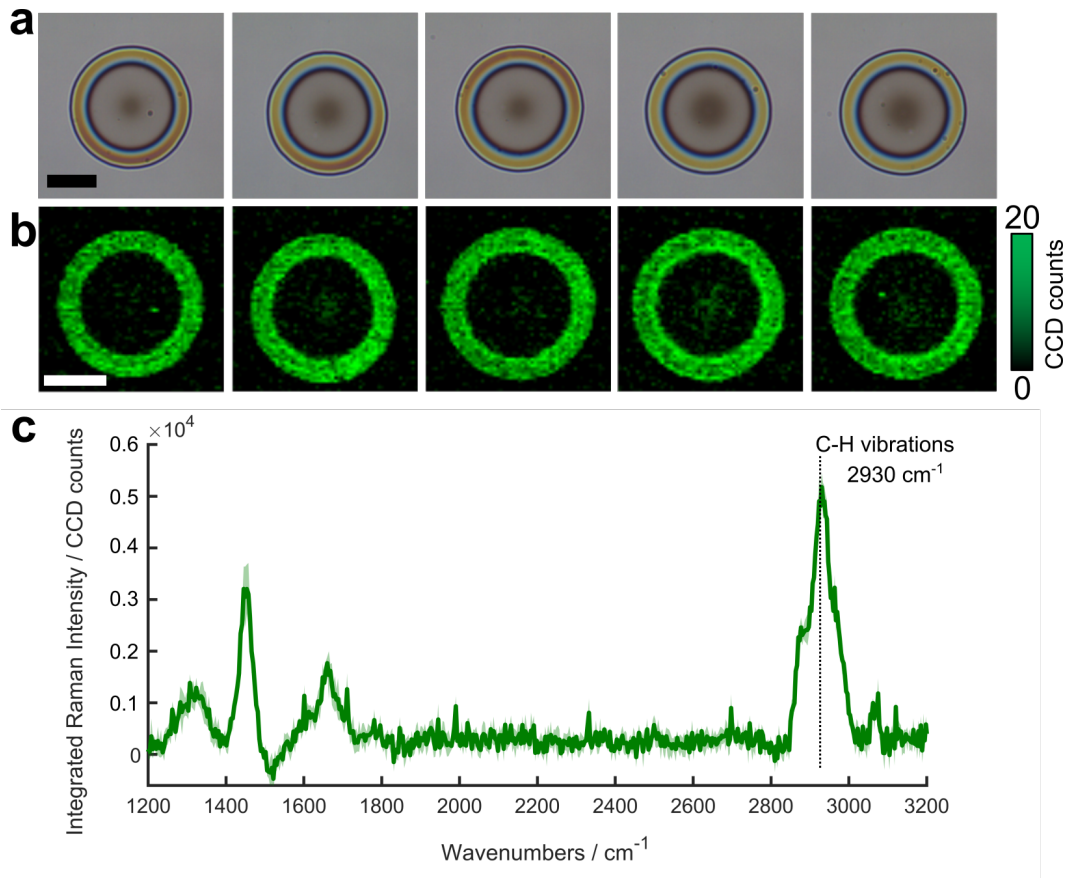


Figure S-16. Raman images of 250pg protein micro-calibration standards.

(a) Reflected brightfield images of 250pg albumin micro-calibration standards on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 250pg micro-calibration standard dataset.

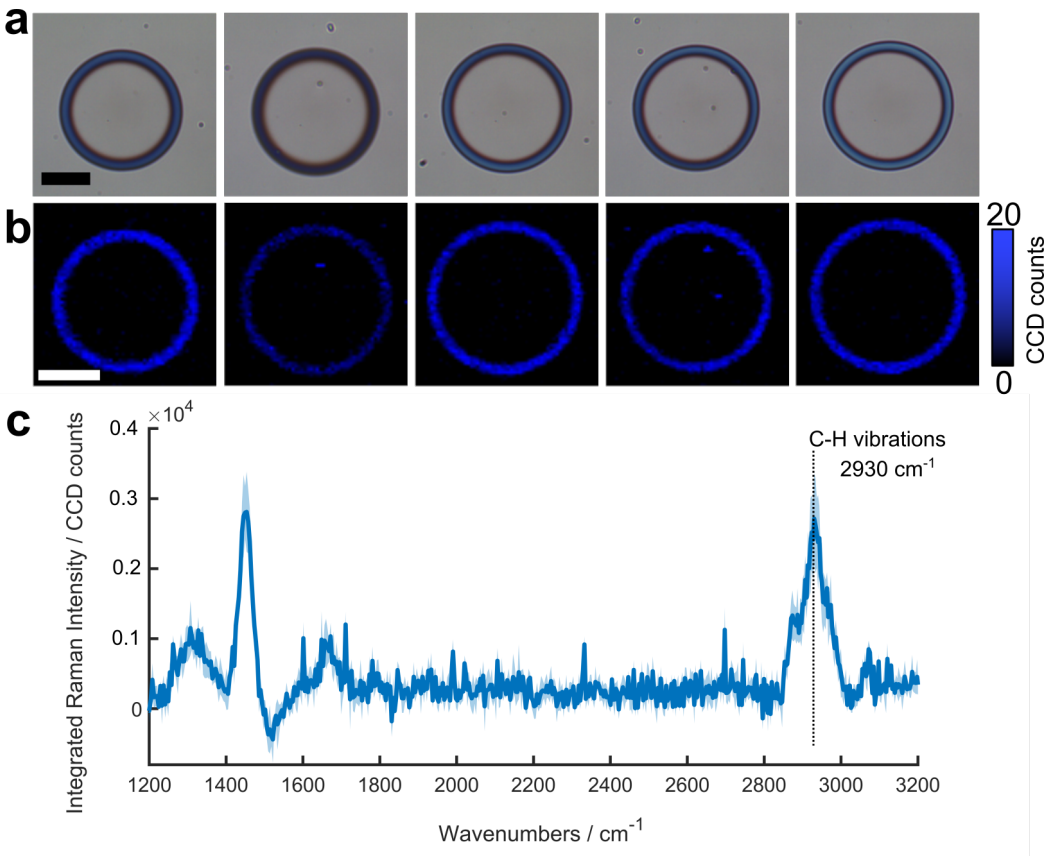


Figure S-17. Raman images of 100pg protein micro-calibration standards.

(a) Reflected brightfield images of 100pg albumin micro-calibration standards on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 100pg micro-calibration standard dataset.

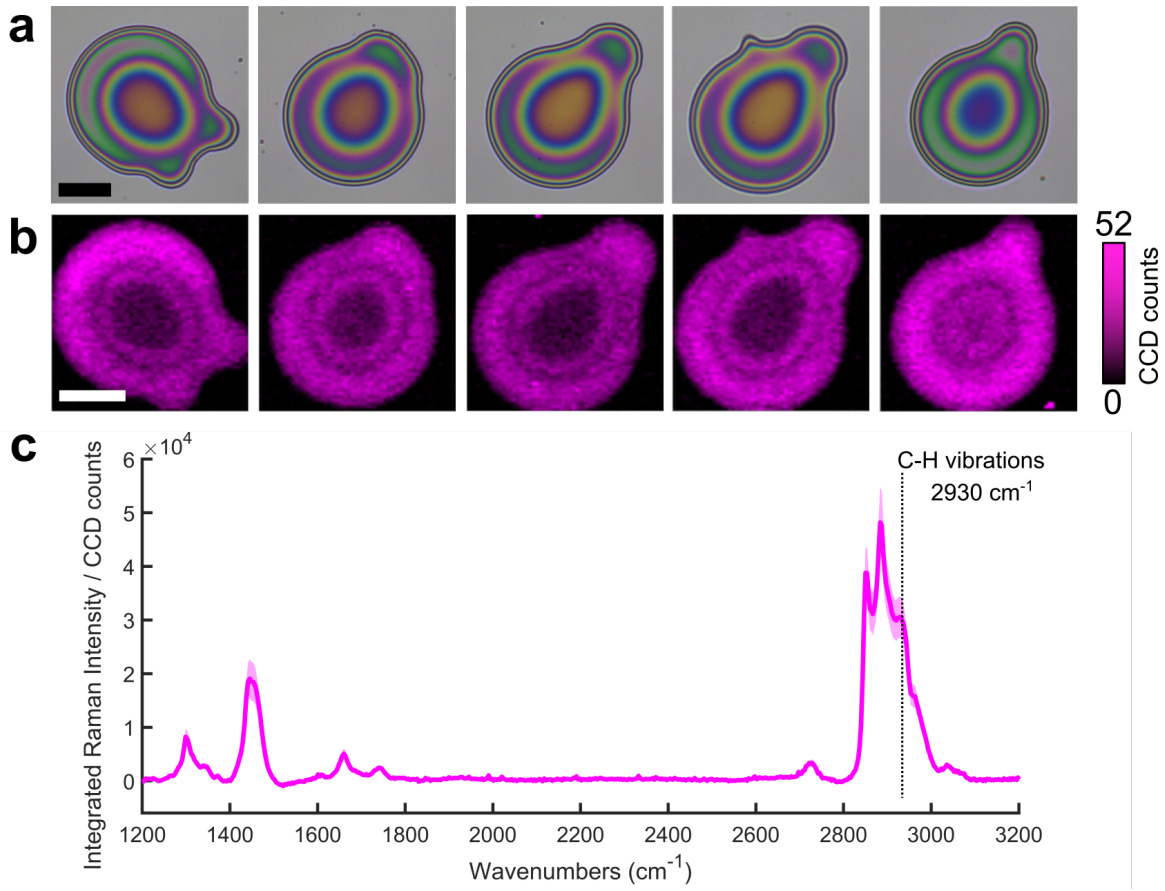


Figure S-18. Raman images of 900pg HDL micro-calibration standards.

(a) Reflected brightfield images of 900pg HDL micro-calibration standards (each containing 600pg of lipid) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 900pg micro-calibration standard dataset.

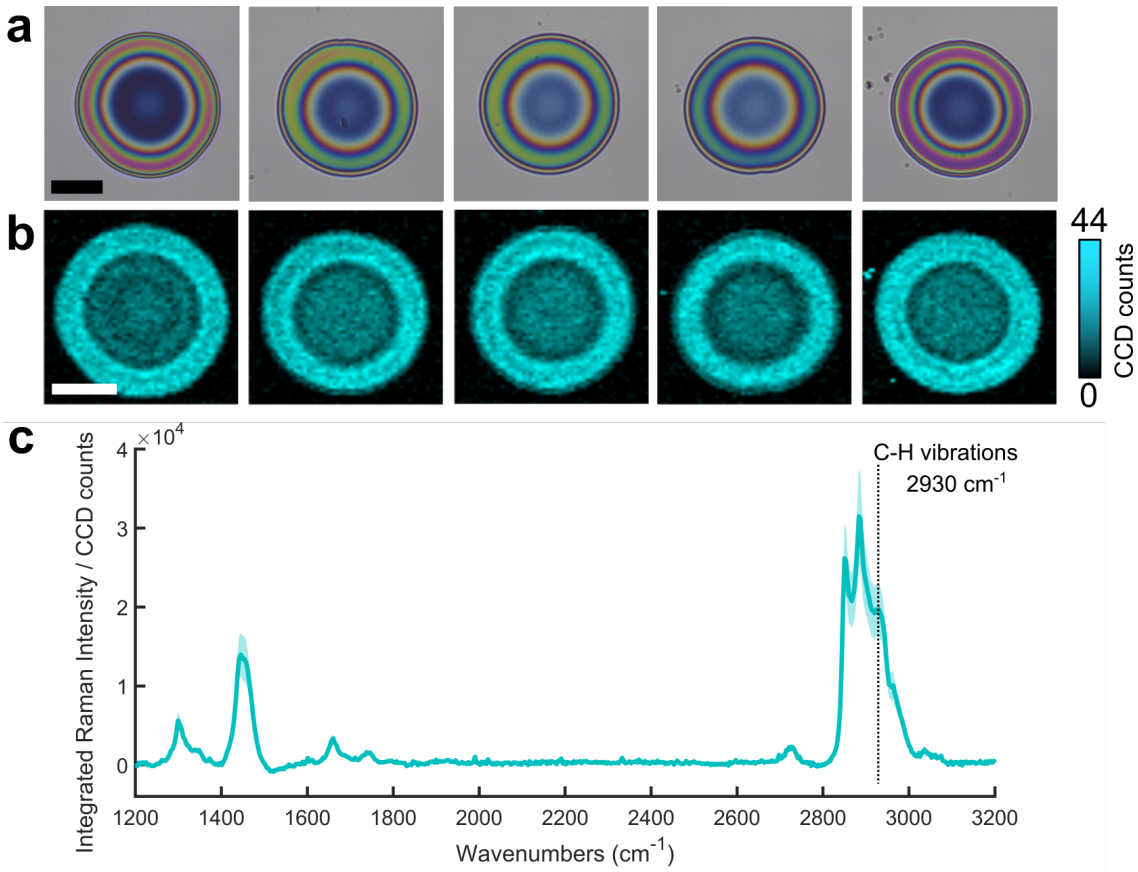


Figure S-19. Raman images of 450pg HDL micro-calibration standards.

(a) Reflected brightfield images of 450pg HDL micro-calibration standards (each containing 300pg of lipid) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 450pg micro-calibration standard dataset.

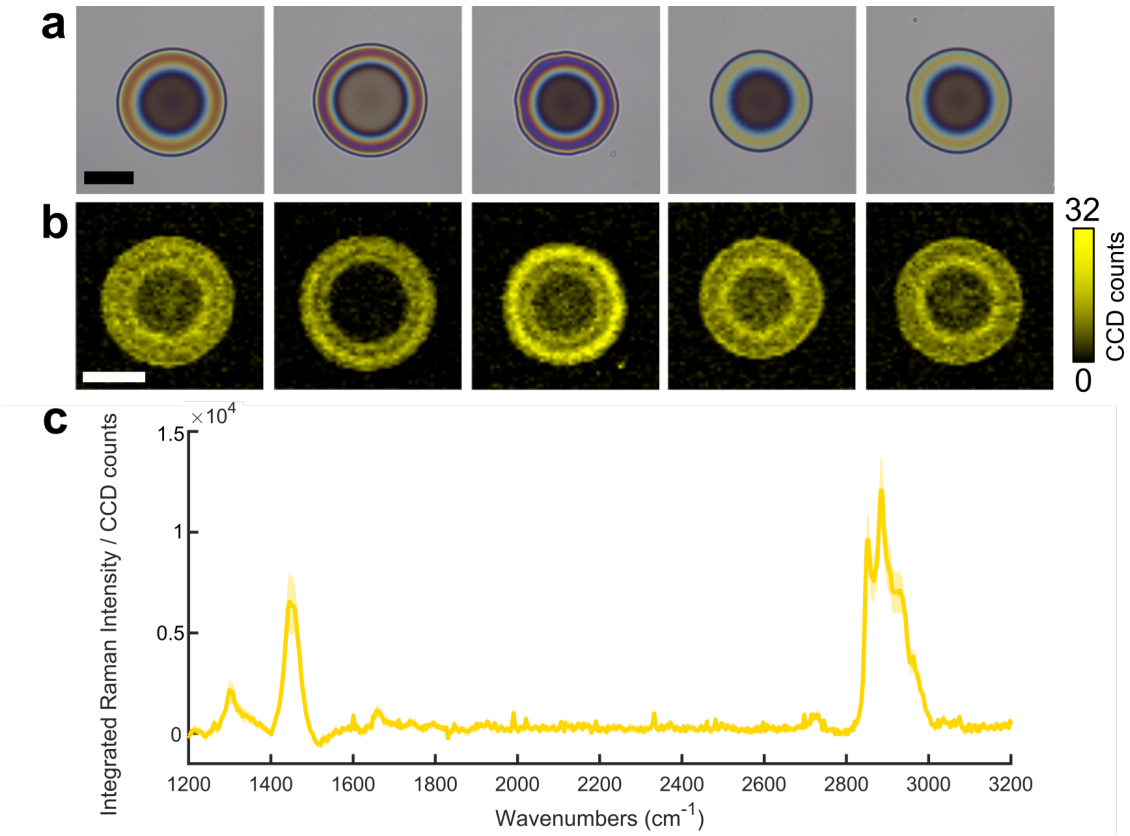


Figure S-20. Raman images of 225pg HDL micro-calibration standards.

(a) Reflected brightfield images of 225pg HDL micro-calibration standards (each containing 150pg of lipid) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 225pg micro-calibration standard dataset.

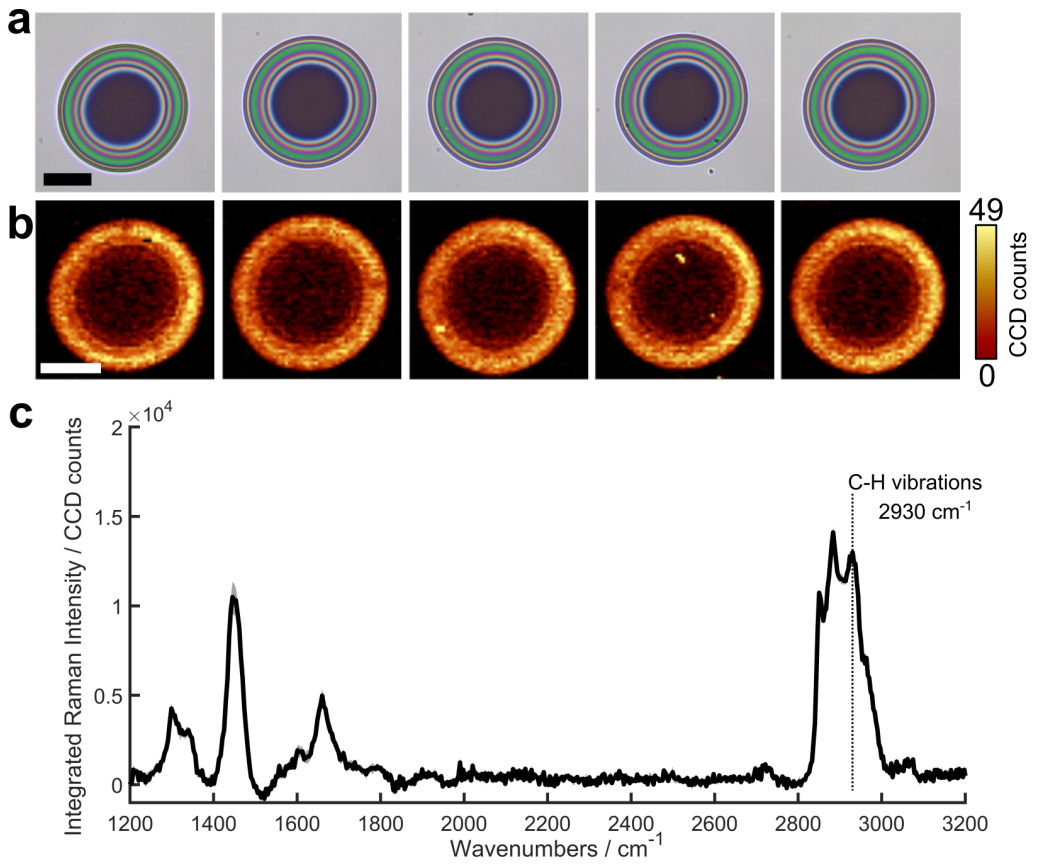


Figure S-21. Raman images of 570pg mixed composition micro-calibration standards.

(a) Reflected brightfield images of 570pg mixed composition micro-calibration standards (each containing 250pg of albumin, 200pg of lipid, 100pg of peptide 22A, and 20pg of cholesterol) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20μm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 570pg micro-calibration standard dataset.

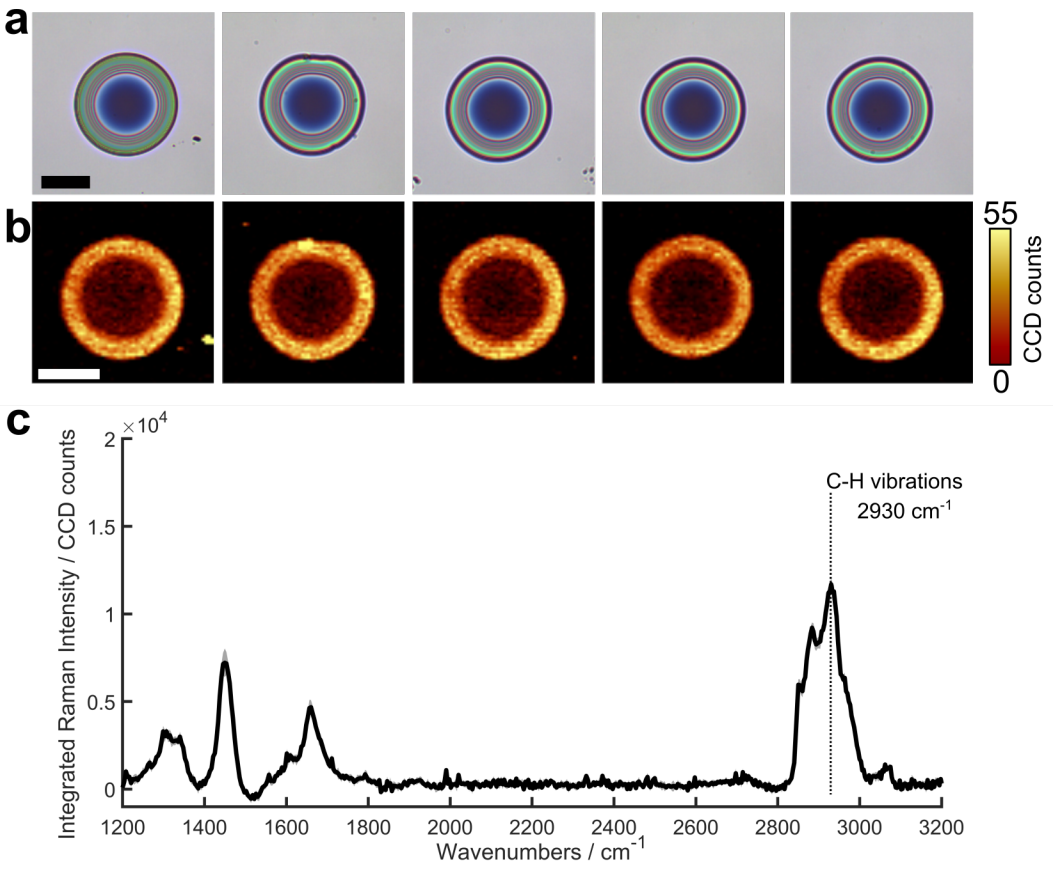


Figure S-22. Raman images of 725pg mixed composition micro-calibration standards.

(a) Reflected brightfield images of 725pg mixed composition micro-calibration standards (each containing 500pg of albumin, 150pg of lipid, and 75pg of peptide 22A) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 725pg micro-calibration standard dataset.

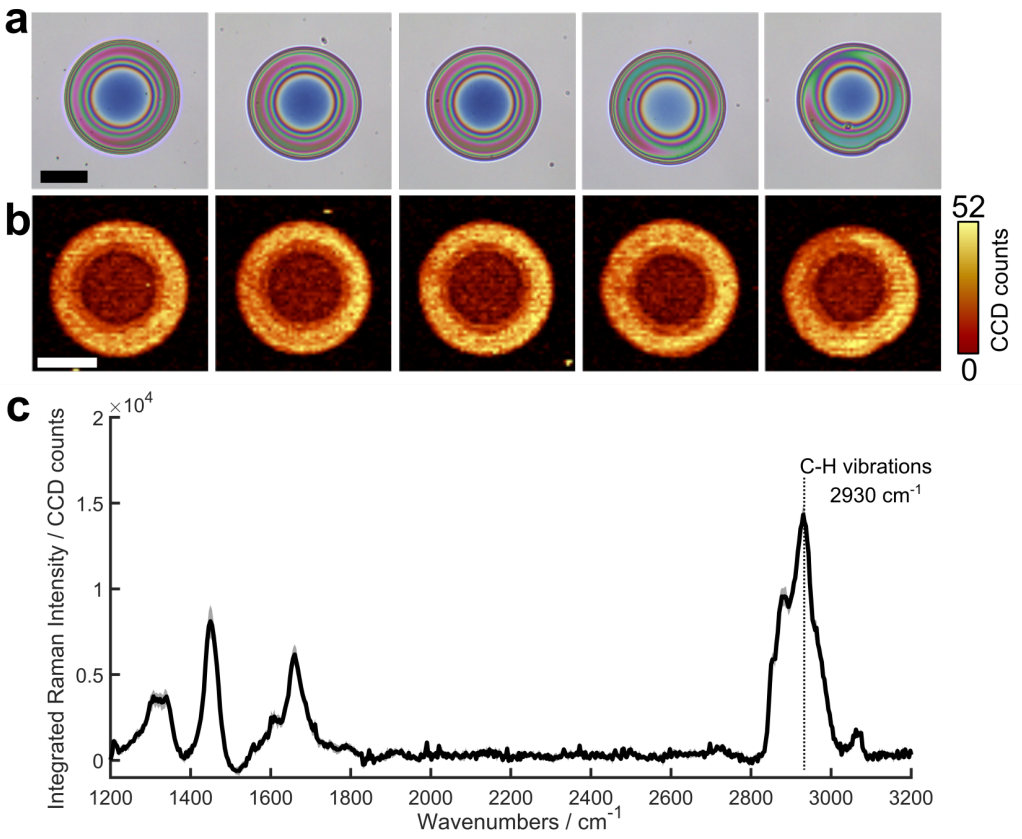


Figure S-23. Raman images of 660pg mixed composition micro-calibration standards.

(a) Reflected brightfield images of 660pg mixed composition micro-calibration standards (each containing 500pg of albumin, 100pg of lipid, 50pg of peptide 22A, and 10pg of cholesterol) on silicon chips and **(b)** corresponding Raman area scans for each showing signal intensity at 2930cm^{-1} and revealing material distribution; scale bars: 20µm. **(c)** Average integrated Raman spectrum (standard deviation shown by shadow) for the 660pg micro-calibration standard dataset.

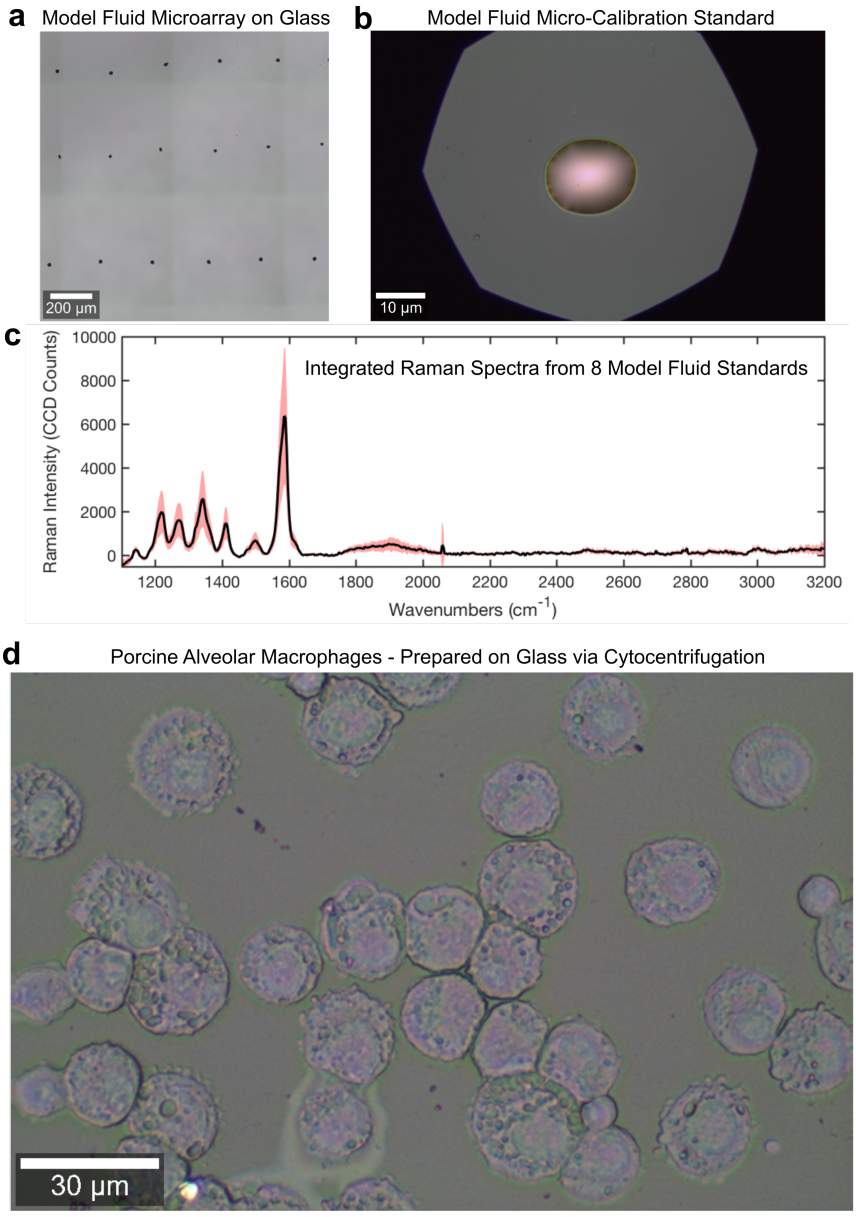


Figure S-24. Preliminary feasibility testing for inkjet printing and cell preparation on glass slides.

(a) Reflected brightfield image of inkjet-printed model fluid microarray on glass slide, **(b)** reflected brightfield image of a single model fluid micro-calibration sample, and **(c)** integrated Raman spectra acquire from 8 different model fluid samples. **(d)** Reflected brightfield image of untreated porcine alveolar macrophages prepared on glass slides via cytocentrifugation.

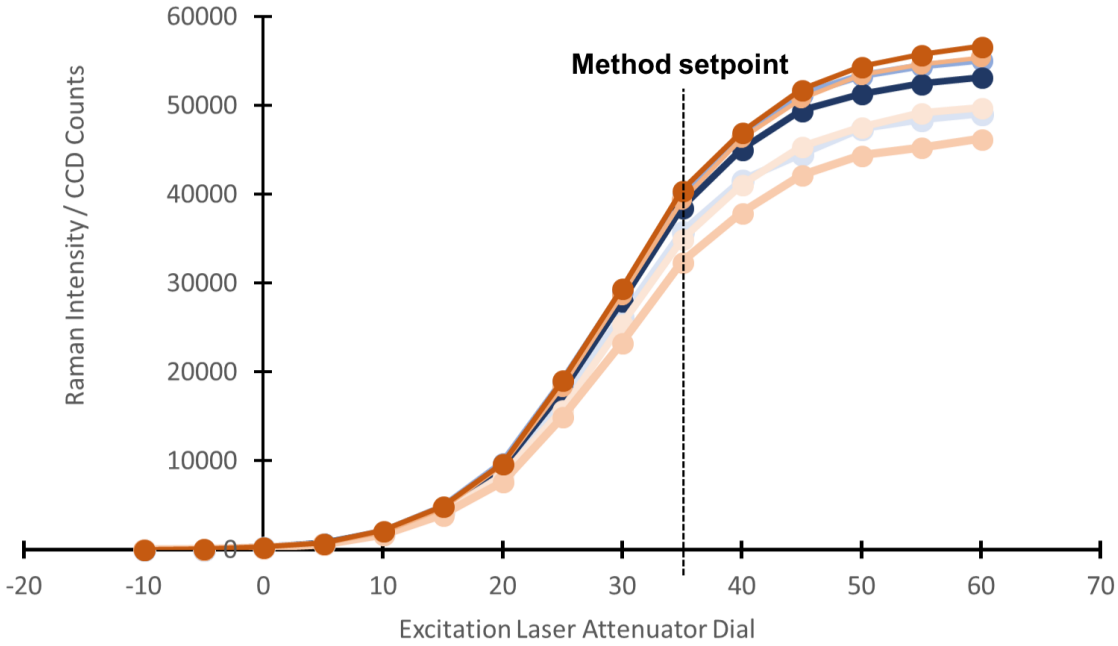
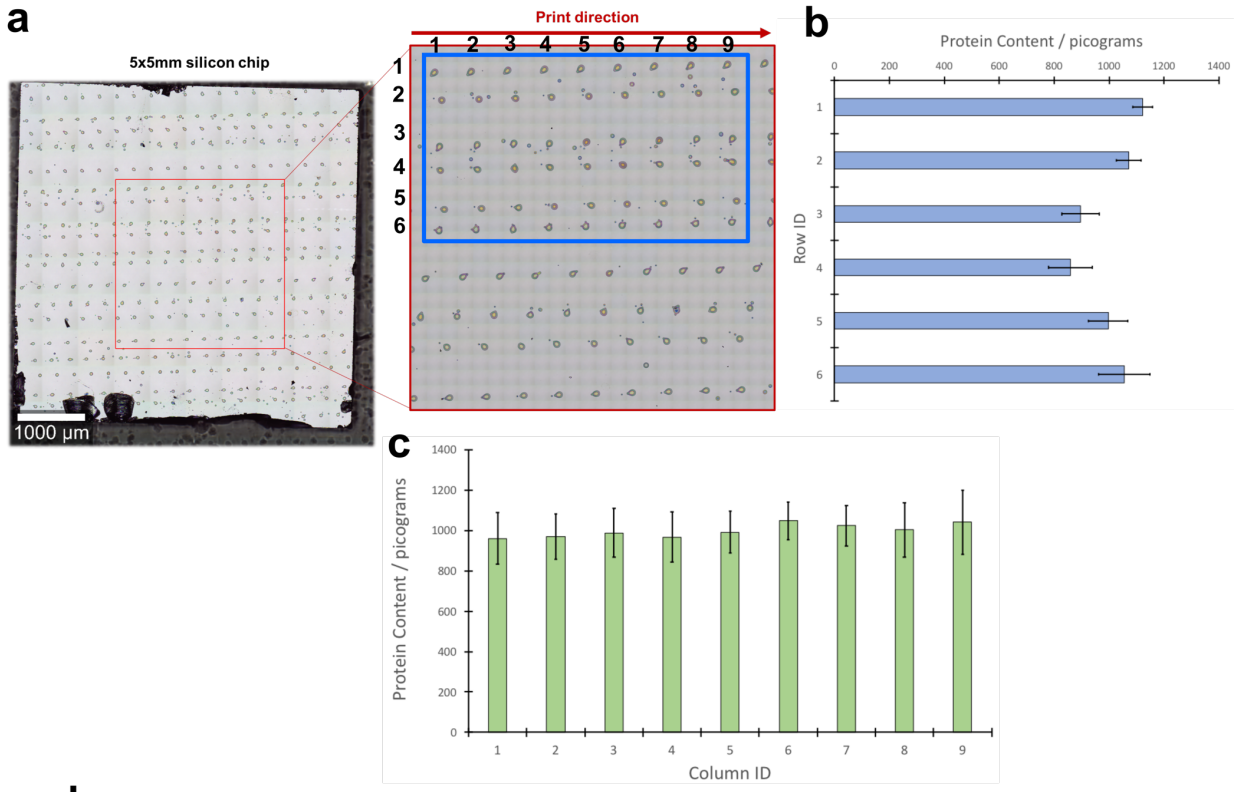


Figure S-25. Instrument daily check performance curves.

Instrument daily check performance curves (different colors/shades respectively represent different days/times); the laser is focused onto a clean surface of silicon chip and the z-plane is adjusted to acquire the optimal, or greatest, Raman signal intensity at 521cm^{-1} (the major Raman peak of silicon). The excitation laser attenuator dial is adjusted from completely closed to completely open in increments shown on dial and single spectra are acquired at each interval. By reporting the Raman intensity at 521cm^{-1} (shown on y-axis), the instrument performance was assessed daily. Instrument performance drift can be quantitatively measured at the method setpoint (the excitation intensity used for all Raman measurements reported in this project).



d

	1	2	3	4	5	6	7	8	9	Average	SD
1	1098	1100	1164	1082	1097	1119	1132	1120	1190	1122	35
2	1111	1021	1041	1079	1046	1139	1075	1009	1122	1071	46
3	871	893	912	839	1023	875	988	837	824	896	69
4	816	785	810	870	817	1055	847	850	880	859	80
5	1004	1006	972	854	921	1049	1073	1052	1045	997	72
6	867	1015	1035	1081	1051	1062	1032	1158	1195	1055	93
Average	961	970	989	967	993	1050	1025	1004	1043		
SD	127	112	121	124	104	93	99	135	158		

Figure S-26. Inkjet-printing nozzle-to-nozzle error analysis.

(a) Error analysis approach for 1000pg albumin microarray; 54 separate micro-calibration standards were analyzed via Raman microanalysis and the protein content (reported in units of picograms) was reported for each **(b)** row, **(c)** column, and **(d)** summarized in data table for each individual micro-calibration standard. A single row is deposited from a single nozzle; each row has a different nozzle.

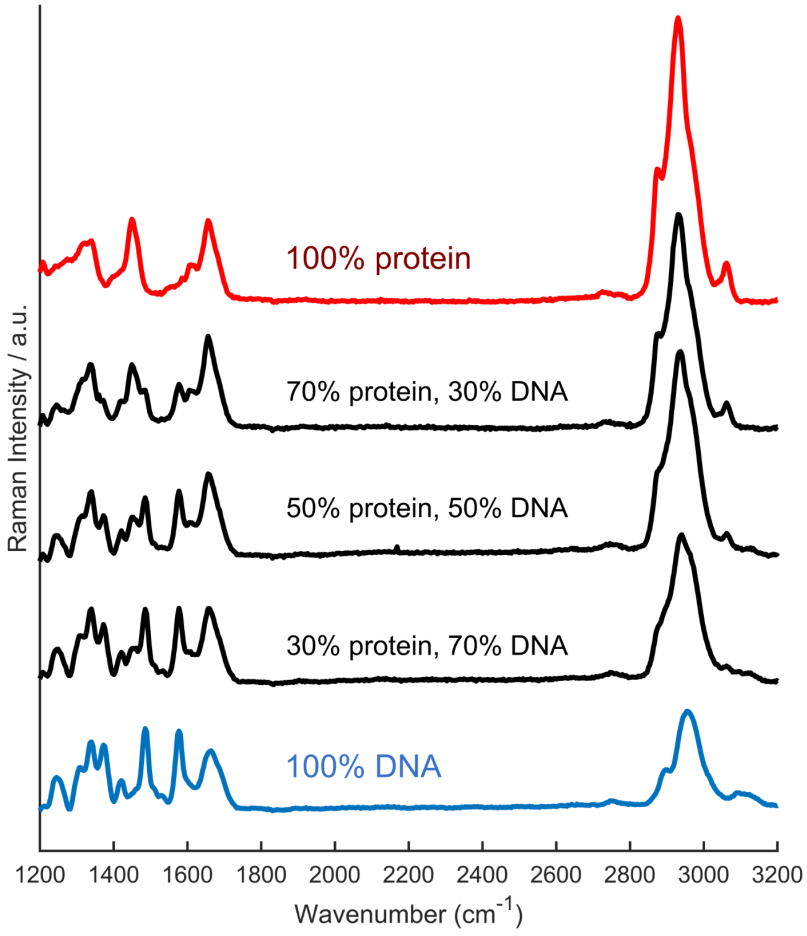


Figure S-27. Stoichiometric mixtures of protein and DNA.

Stoichiometric mixtures of albumin and DNA used to experimentally determine “carbohydrate & nucleic acid” ultraquantitative reference spectra scaling.

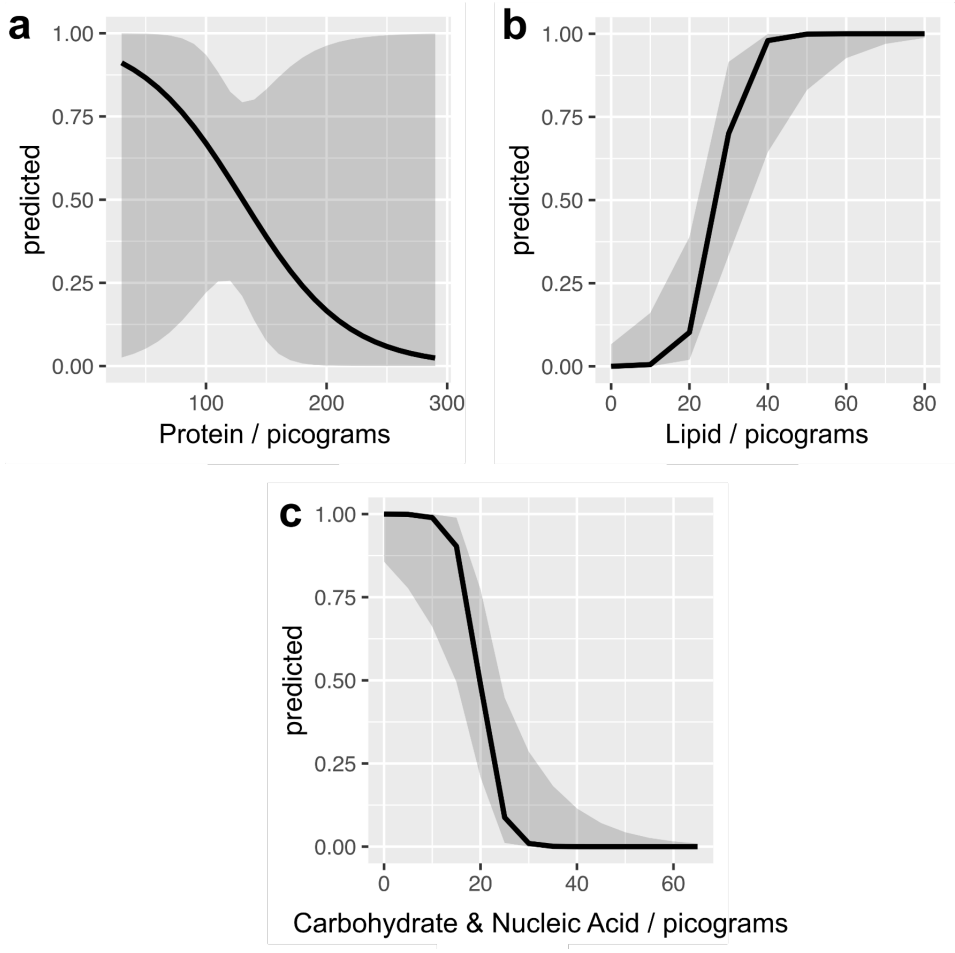


Figure S-28. Logistic regression results for alveolar macrophage ultraquantitative dataset.

Predictive probability plots for **(a)** protein, **(b)** lipid, and **(c)** carbohydrate and nucleic acid.

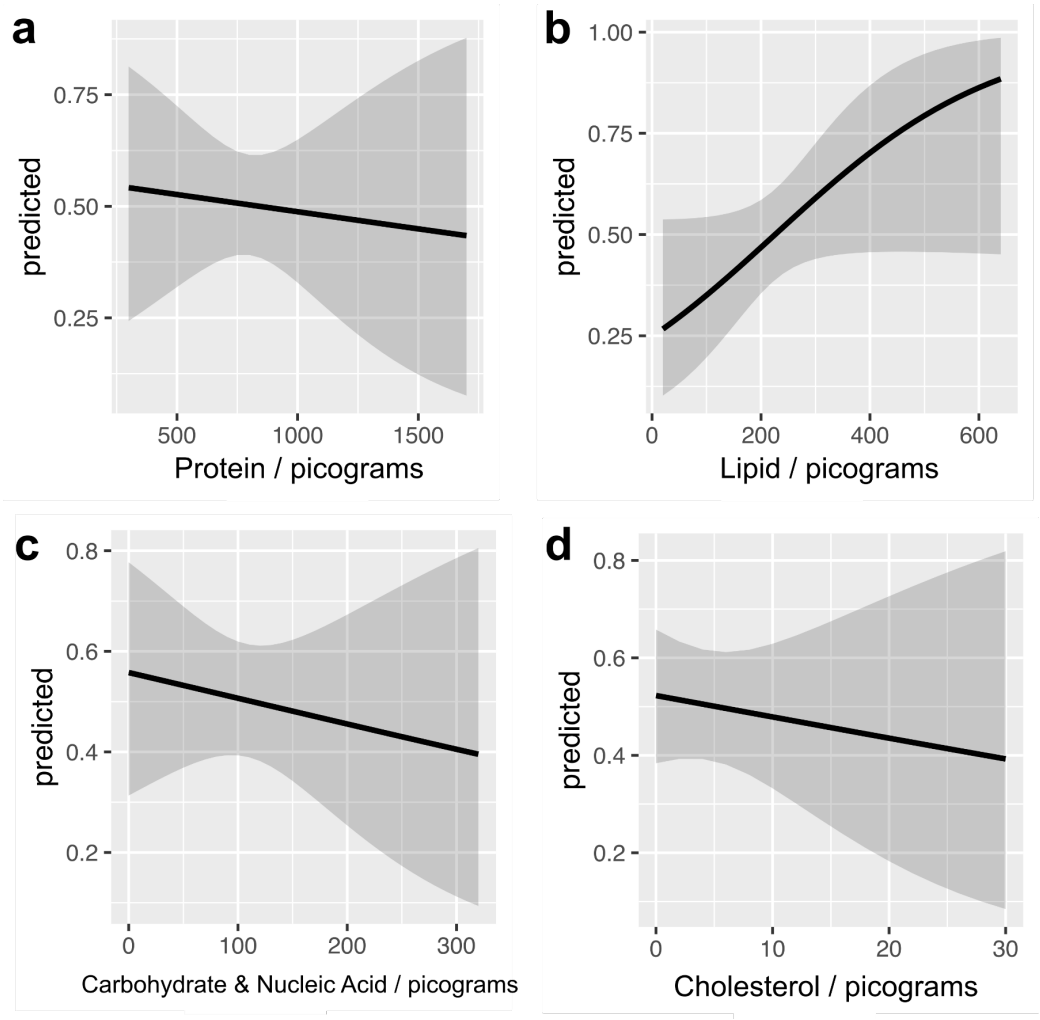


Figure S-29. Logistic regression results for fibroblast ultraquantitative dataset.

Predictive probability plots for **(a)** protein, **(b)** lipid, **(c)** carbohydrate and nucleic acid, and **(d)** cholesterol measurements.

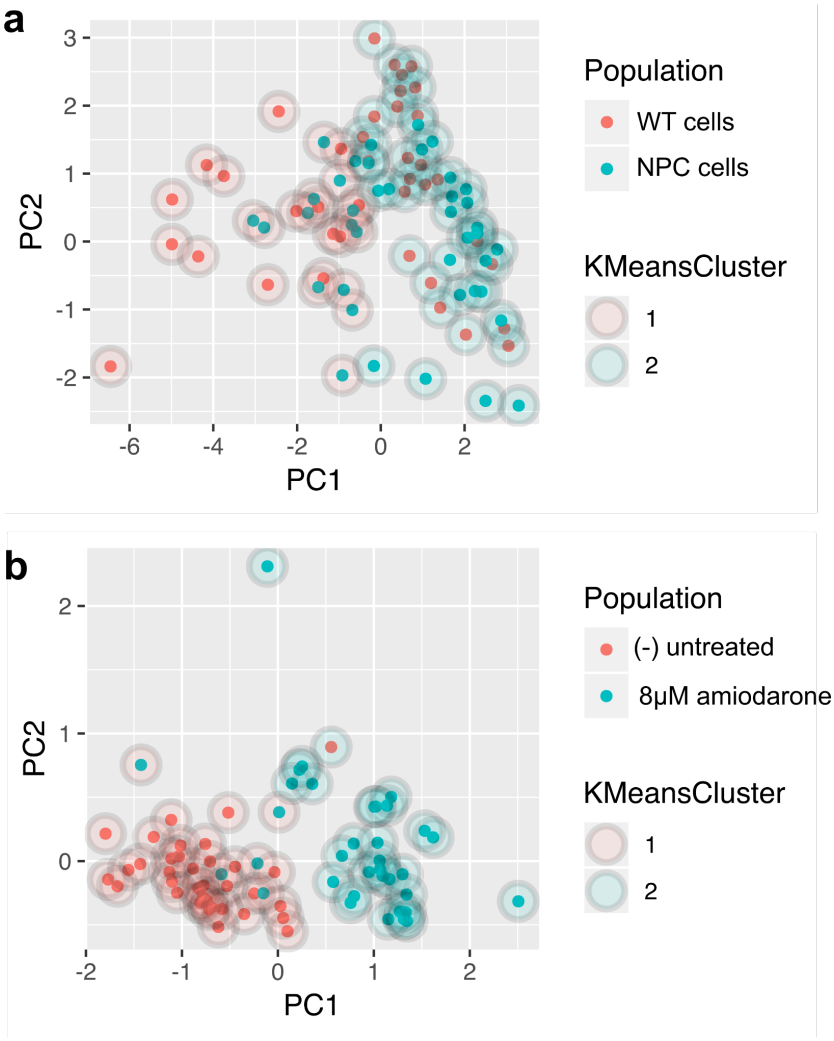


Figure S-30. K-means cluster analysis for cell population differentiation.

K-means clustering of PCA results datasets (PC scores) for **(a)** fibroblasts and **(b)** alveolar macrophages.

C.2 Tables

Table S-5. Logistic regression model results for alveolar macrophage ultraquantitative dataset (using picogram measurements as predictors).

	Estimate	Std. Error	z value	p-value
Intercept	3.73838145	1.59365893	2.3457852	0.018987049
Protein	-0.02320455	0.03158227	-0.7347335	0.462501832
Lipid	0.30189391	0.10791507	2.7975140	0.005149754
Carbohydrate & Nucleic Acid	-0.45733076	0.18154032	-2.5191691	0.011763215

Table S-6. Logistic regression model results for fibroblast ultraquantitative dataset (using picogram measurements as predictors).

	Estimate	Std. Error	z value	p-value
Intercept	-0.567338105	0.780569701	-0.7268257	0.46733275
Protein	-0.000308520	0.001242250	-0.2483558	0.80385910
Lipid	0.004917963	0.002678185	1.8363047	0.06631261
Cholesterol	-0.017560972	0.037796071	-0.4646243	0.64220055
Carbohydrate & Nucleic Acid	-0.002052581	0.004299290	-0.4774232	0.63306080

APPENDIX D

Supporting Information in Chapter 5

D.1 Tables

Table S-7. Directory of pig subjects from sepsis/ARDS experimental model.

	Date of Experiment	Injury Model	Initial P/F Ratio	Survival Time (hr)
Pig #1	11 December 2017	Sepsis	425	16
Pig #2	23 April 2018	Gastric Acid + Sepsis	529	10
Pig #3	26 June 2018	Gastric Acid + Sepsis	491	4
Pig #4	16 October 2018	Gastric Acid + Sepsis	536	3
Pig #5	22 January 2019	Gastric Acid	404	23

APPENDIX E

Supporting Information in Chapter 6

E.1 Supporting Methods

To assess automated algorithm reliability and probability of false negative and positive interpretations, the most abundant chemical components in the system were identified and utilized to create test sample sets. The components of interest were 4HPR, protein, lipid, residual blood, and glass. For representative protein and lipid materials, albumin (MP Biomedicals; cat. #810033) and 1,2-dioleoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids, Inc.; cat. #850375P) were used. Each component was dissolved in an appropriate solvent (4HPR/lipid in methanol and protein/blood in water), transferred to glass slide, and allowed to evaporate dry as pure component dispersions. Additionally, an untreated tissue section containing residual blood and an untreated tissue section washed free of blood (with PBS) were analyzed on glass to determine reliability of algorithm's scaled glass and blood subtraction. Large area Raman scans were acquired from the resulting test samples using the Raman acquisition methodology. These test datasets were processed through the algorithm and the error frequency was determined based on fraction of pixels interpreted as 4HPR-positive or 4HPR-negative.

E.2 Figures

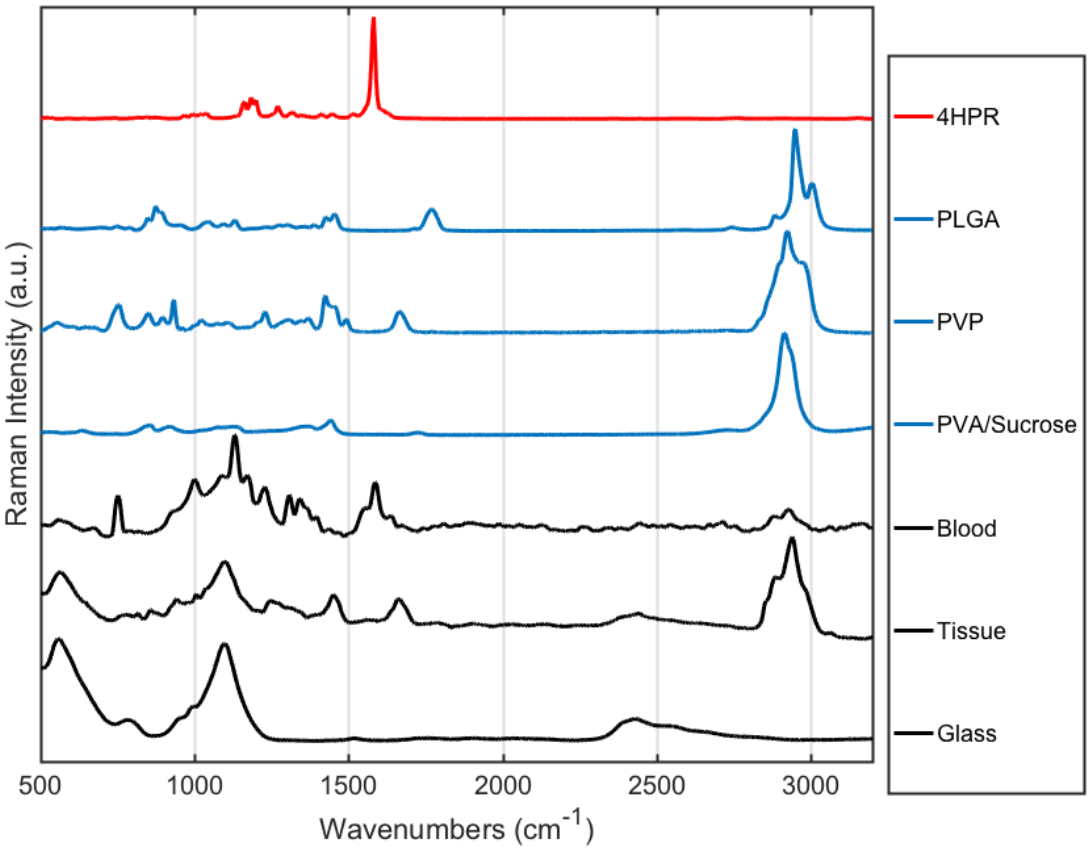


Figure S-31. Reference spectral overlay of signal contributors in tissue sections and millicylinder formulations.

Millicylinder excipients and biocomponent Raman reference spectra show no interference with 4HPR's peak at 1582 cm^{-1} except from blood. Spectral contributions from glass are apparent in unprocessed tissue spectra. Spectra are shown for 4HPR, PVP, PLGA, PVA/Sucrose, rabbit blood, and untreated tissue.

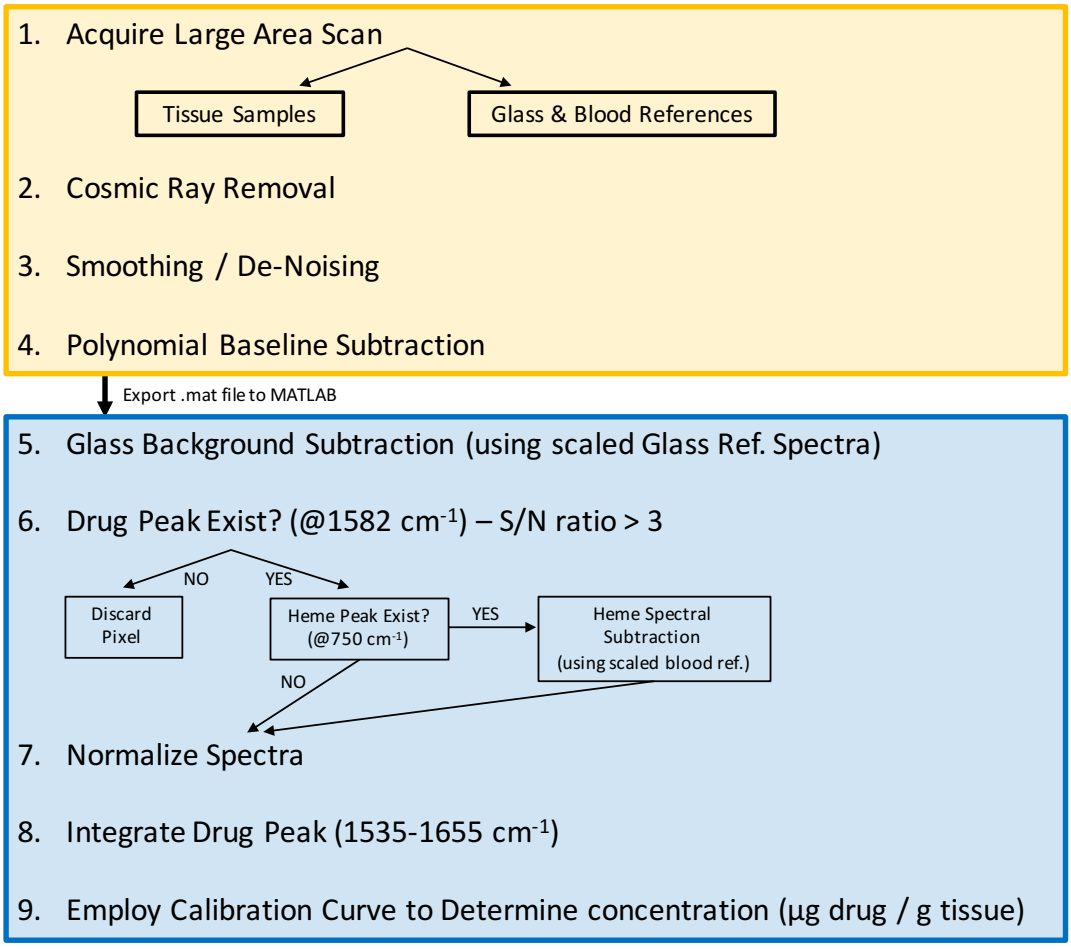


Figure S-32. Raman hyperspectral image processing algorithm overview.

Data acquisition and initial pre-processing performed using WiTec ProjectFOUR software (shown on top in yellow); automated MATLAB algorithm processing steps (shown on bottom in blue).

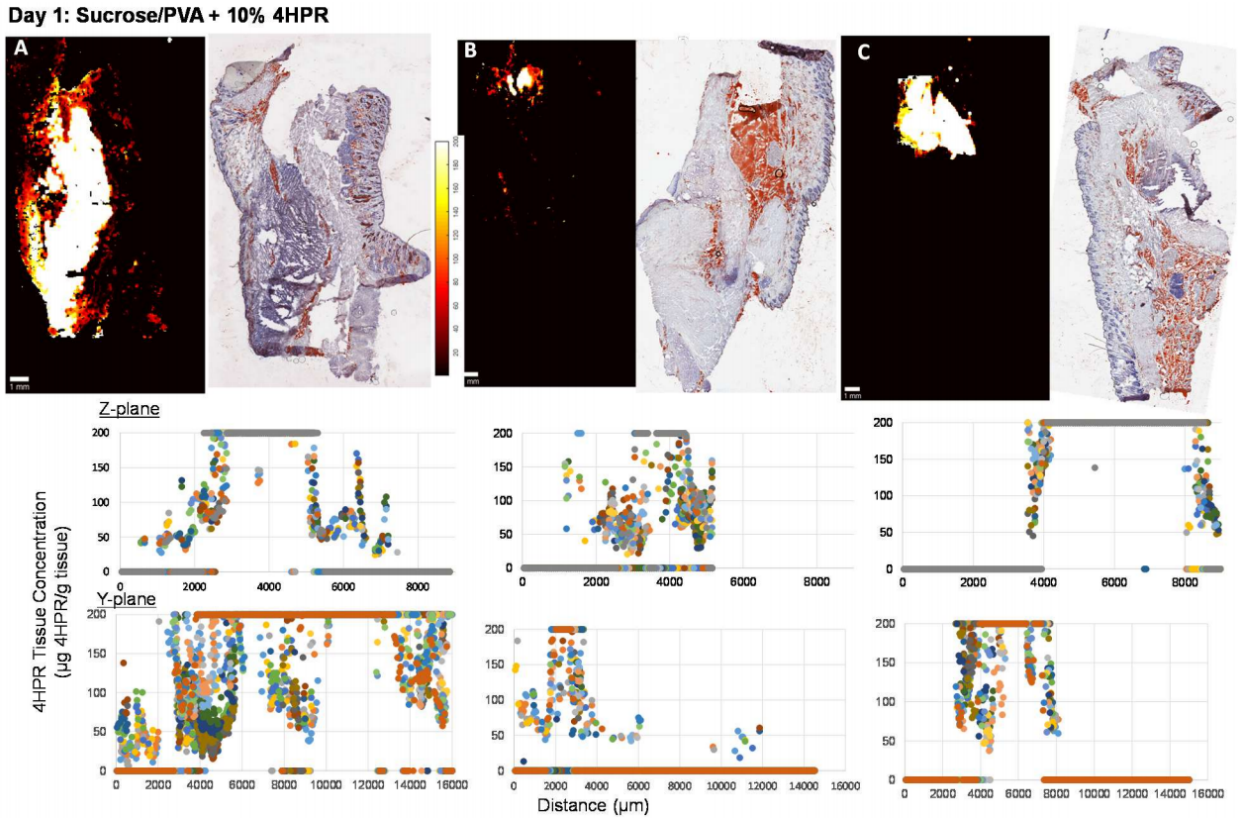


Figure S-33. Tissue distribution after 1 day from sucrose/PVA + 10% 4HPR.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

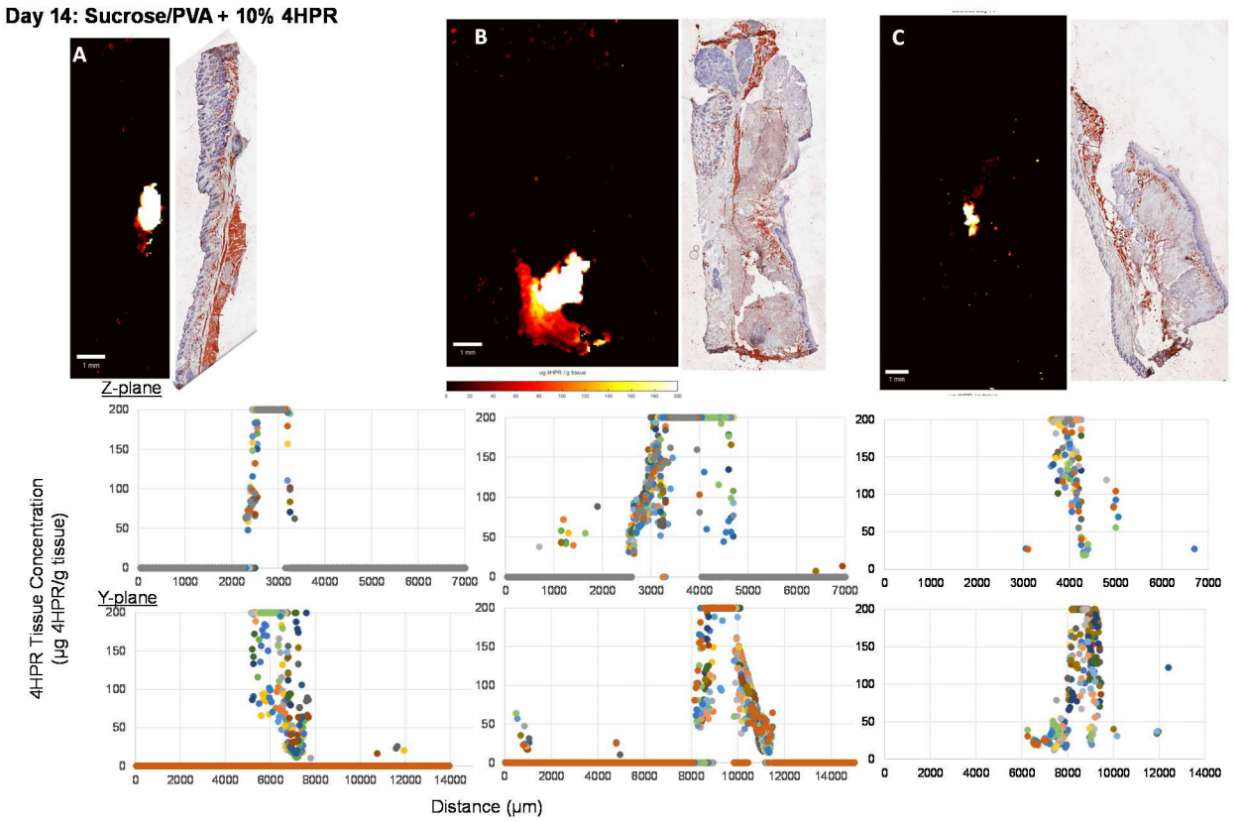


Figure S-34. Tissue distribution after 14 days from sucrose/PVA + 10% 4HPR.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

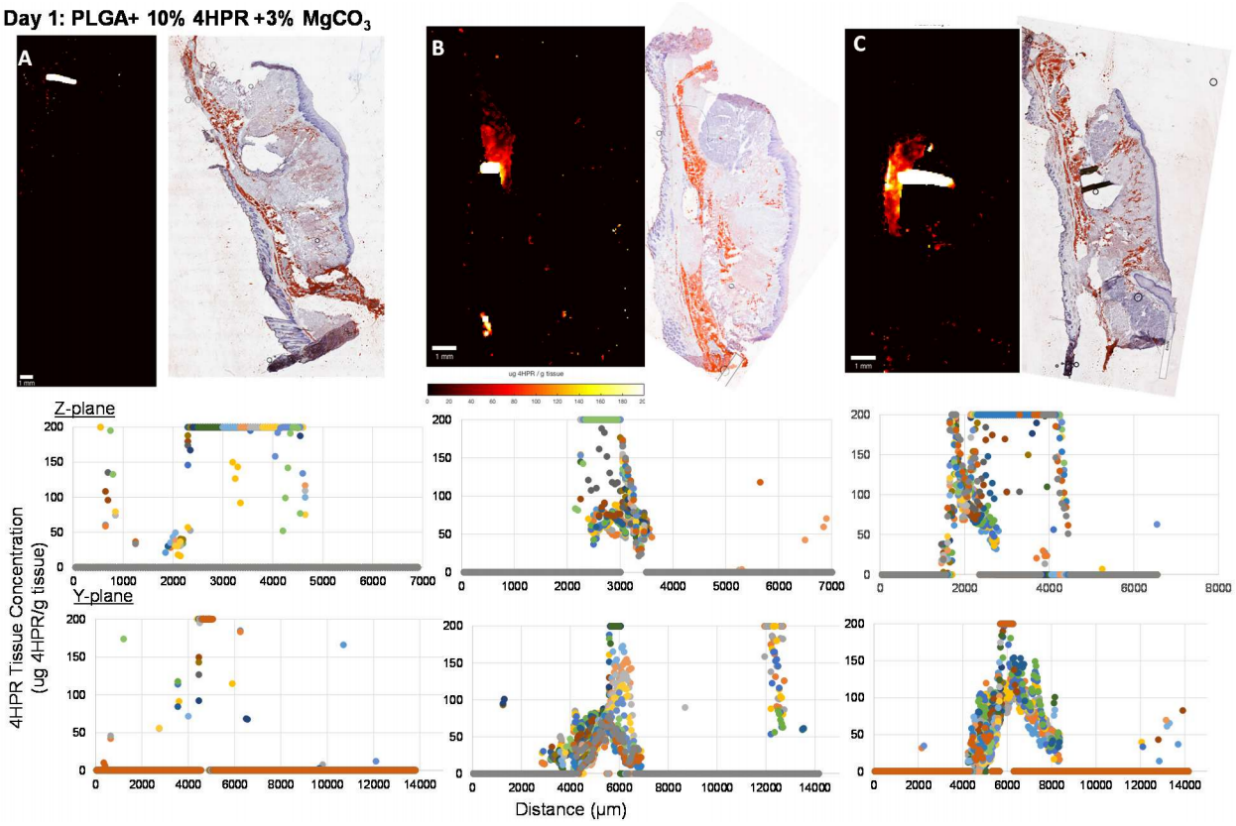


Figure S-35. Tissue distribution after 1 day from PLGA + 10% 4HPR + 3% $MgCO_3$.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

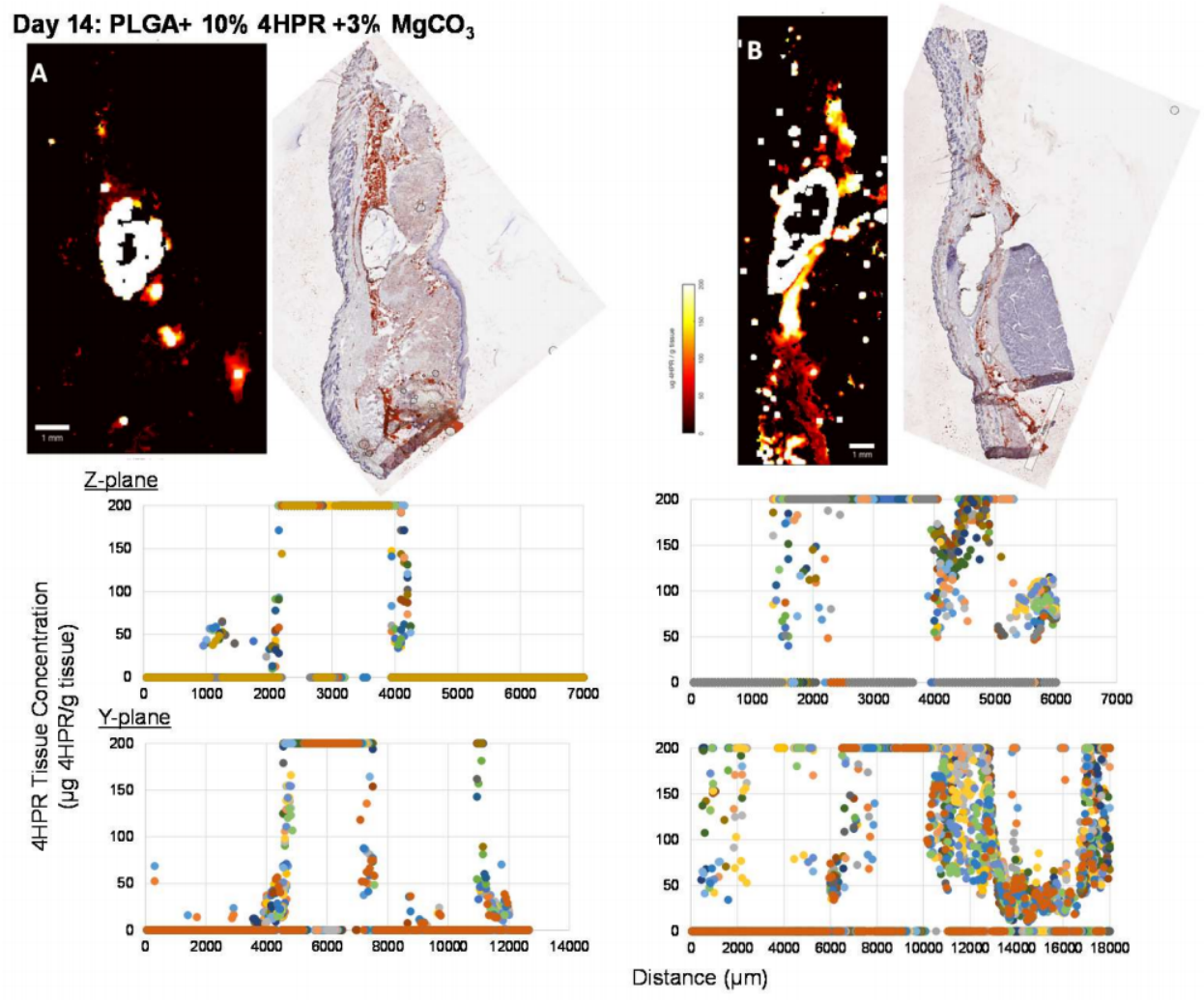


Figure S-36. Tissue distribution after 14 days from PLGA + 10% 4HPR + 3% $MgCO_3$.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

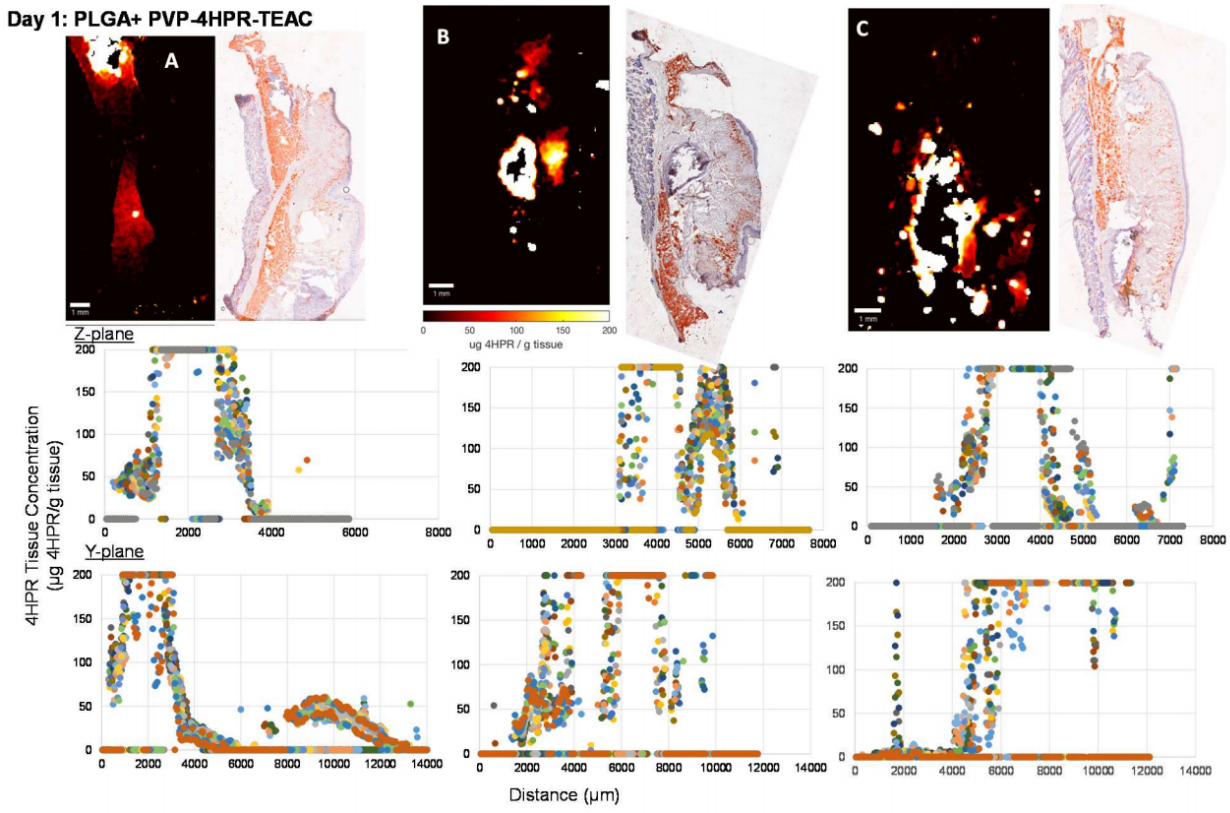


Figure S-37. Tissue distribution after 1 day from PLGA + PVP-4HPR-TEAC.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

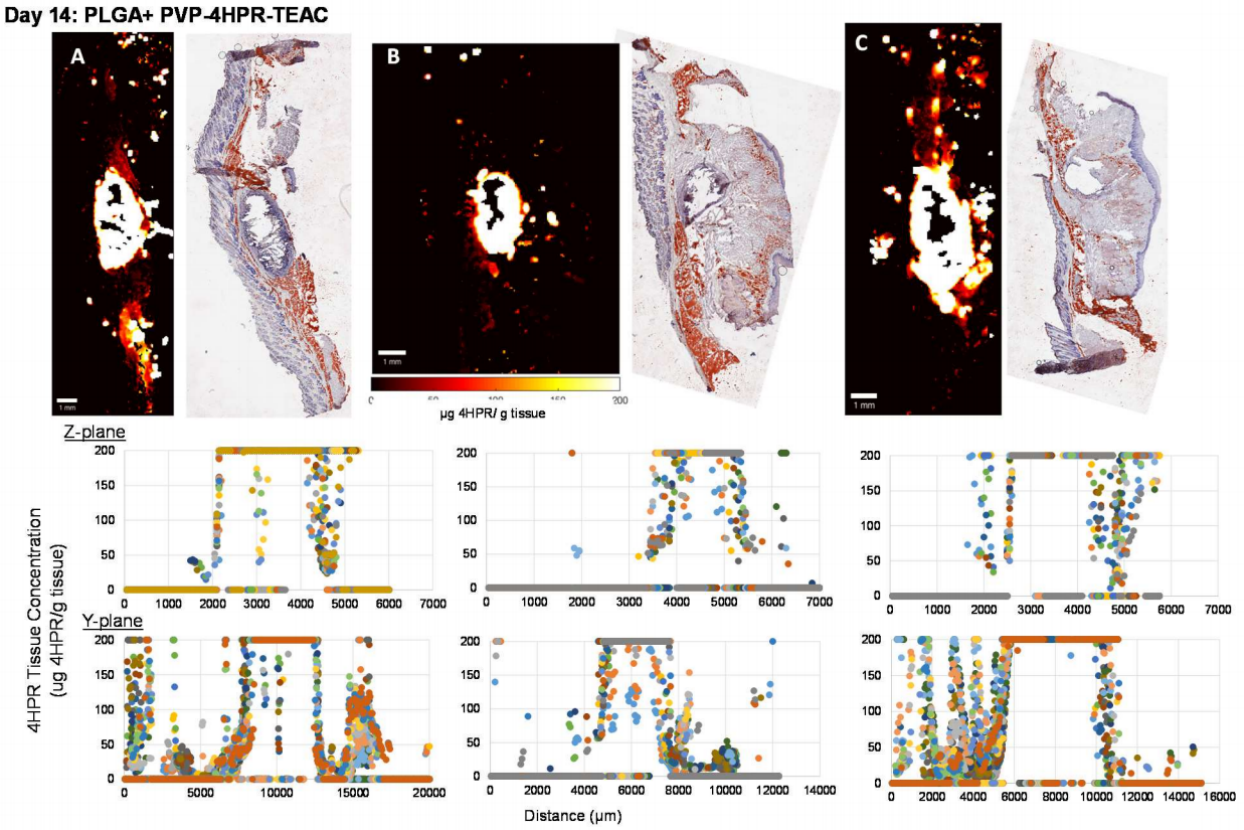


Figure S-38. Tissue distribution after 14 days from PLGA + PVP-4HPR-TEAC.

Raman images of 4HPR distribution in buccal epithelium section after release from long-acting implants after days 1 and 14, n=3 each.

E.3 Tables

Table S-8. Algorithmic error probability testing determined from test datasets.

Spectral components present in test dataset	Glass				
	Tissue				
		Blood			
			4HPR		
				Protein	
					Lipid
# Spectra acquired	9600	625	625	625	625
# 4HPR-positive	6	1		0	0
# 4HPR-negative			0		
% False (+)	0.063	0.160		0	0
% False (-)			0		