UNIVERSITY OF CALIFORNIA SAN DIEGO

Development of a Label-free Optical Analysis Platform for Cells and Tissues

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Bioengineering

by

Anthony A. Fung

Committee in charge:

Professor Lingyan Shi, Chair
Professor Gensheng Feng
Professor Geert W. Schmid-Schoenbein
Professor Peter Ying-Xiao Wang

2023

The Dissertation of Anthony A. Fung is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

Dedication

For my parents.

Epigraph

If it has legs, run with it.

Acknowledgments

Foremost, I thank Professor Lingyan Shi for her support, guidance, and inspiration both in the lab, and in my personal life. I've deeply enjoyed being mentored by, and growing with, Dr. Shi through our insightful conversations and research struggles. Without her reminders to focus on our abilities and persevere, none of this would be possible. She has continuously searched for, and created opportunities for me, and I am supremely confident that much of my future would not be possible without her as well. I would also like to thank my parents, loved ones, and pets for their unwavering support and contagious pride throughout my time at UC San Diego.

I am indebted to the department of Bioengineering and record a special note of appreciation for the members of my committee, Professors Geert W. Schmid-Schoenbein, Peter Ying-Xiao Wang, and Gensheng Feng, for their valuable comments, experience, and for being wonderful role models. I am also grateful to the Dr. Dennis Carson Collaboratory and its members for giving me the confidence and skills necessary to enhance my research in the biological arena.

Finally, I would like to thank my fellow lab members for their warm company, intellectual contributions, and for occasionally giving me their time slots for the microscope.

Chapter 2, in full, is a reprint of the material as it appears in Frontiers in Oncology Volume 12, 2022. Fung, Anthony A.; Hoang, Khang; Zha, Honghao; Chen, Derek; Zhang, Wenxu; Shi, Lingyan, Frontiers, 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication of the material as it may appear in Nature Communications, 2023, Anthony A. Fung, Wenxu Zhang, Yajuan Li, Zhi Li, Hongje Jang, Honghao Zha, Xiaoping Chen, Fangyuan Gao, Jane Y. Wu, Huaxin Sheng, Junjie Yao,

Dorota Skowronska-Krawczyk, Sanjay Jain & Lingyan Shi. The dissertation author was the primary investigator and author of this paper. Additionally, we thank Drs. K. Zhang, C. Metallo, F. Liu, and G. Schmid-Schoenbein for helpful discussions. We acknowledge support from UCSD Startup funds, NIH 5R01NS111039, NIH R21NS125395, NIHU54DK134301, NIHU54 HL165443, NIH U54CA132378, and Hellman Fellow Award. We are grateful for the support of the Washington University Kidney Translational Research Center (KTRC) for kidney samples and the HuBMAP grant U54HL145608. We thank Dr. E. Bigio and Dr. M-M. Mesulam from Mesulam Center for Cognitive Neurology and Alzheimer's Disease (MCCNAD) for providing the de-identified autopsy brain samples; and MCCNAD is supported by NIH P30 AG013854.

Chapter 4, in full is currently being prepared for submission for publication of the material. Anthony A. Fung, Zhi Li, Craig Boote, Petar Markov, Sanjay Jain, Lingyan Shi. The dissertation author was the primary researcher and author of this material.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AL	Antennal Lobe
ANOVA	Analysis of variance
APP	Amyloid Precursor Protein
ArPLS	Asymmetrically Reweighted Penalized Least Squares
AUC	Area Under Curve
CCD	Charge Coupled Device
CCT	CTP:phosphocholinecytidylyl transferase
CD	Carbon-Deuterium
CER	Ceramide
CH	Carbon-Hydrogen
CHO	Chinese Hamster Ovary
CKD	Chronic Kidney Disease
CL	Cardiolipin
CO	Oil Condenser
CO2	Carbon Dioxide
CPU	Central Processing Unit
D2O	Deuterium Oxide
DFT	Discrete Fourier Transform
dH2O	deionised water
DIY	Do It Yourself
DKD	Diabetic Kidney Disease
DKL	Kullback-Leibler Divergence
DM	Dichroic Mirror
DMEM	Dulbecco's Modified Essential growth Media
DN	Diabetic Nephritis
DNA	Deoxyribonucleic acid
DO-SRS	Deuterium Oxide Stimulated Raman Scattering
ECM	Extracellular Matrix
EOM	Electro-optical modulator
ER	Endoplasmic Reticulum
FAD	Flavin Adenine Dinucleotide

FFPE	Formalin Fixed Paraffin Embedded
FFT	Fast Fourier Transform
FLIM	Flourescence Lifetime Imaging
FMN	Flavin Mononucleotide
FMO	Flavin Monooxygenase
FPGA	Field Programmable Gate Array
FTIR	Fourier Transformed Infrared
FTU	Functional Tissue Unit
GAN	Generative Adversarial Network
GBM	Glomerular Basement Membrane
GCMS	Gas Chromatography Mass Spectrometry
GPU	Graphical Processing Unit
GSH	Glutathione (reduced)
GSSG	Glutathione (oxidized)
H&E	Hematoxylin and Eosin
HDL	High Density Lipid
HEK	Human Embryonic Kidney
HSI	Hyperspectral Imag(e)(ing)
HSV	Hue Saturation Value
IF	Immunofluorescence
IR	Infrared
L*a*b*	International Commission on Illumination Lstar Astar Bstar
LD	Lipid Droplet
LDL	Low Density Lipid
LM	Light Microscopy
LUT	Look-up Table
MALDI-TOF	Matrix Assisted Laser Desorptiob Ionization Time of Flight
MCR-ALS	Multivariate Curve Resolution Alternating Least Squares
MIP	Maximum Intensity Projection
MLP	Multilayer Perceptron
MS	Mass Spectrometry
mTOR	Mammalian Target of Rapamycin
NA	Numerical Aperture

NAD[P]H	Nicotinamide Adenine Dinucleotide [phosphate] Reduced
NAO	Nonyl Acridine Orange
NMR	Nuclear Magnetic Resonance
NOX	Nitric Oxide
NS	Nephrotic Syndrome
OD	Optical Density
OB	Objective Lens
OCT	Optimal Cutting Temperature
ORR	Optical Redox Ratio
PAS	Period Acid Schiff
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
PGS1	phosphatidylglycerophosphate synthase 1
PFA	Paraformaldehyde
PGP	Phosphatidylglycerophosphate
PINV	Pseudo Inverse
PPP	Pentose Phosphate Pathway
PPSID	Periodicity Plus Smooth Image Decomposition
PRM	Penalized Reference Matching
PSI	Point Sample Intercept
PTPMT1	phosphatidylglycerophosphate phosphatase
RAM	random Access Memory
ReLU	Rectified Linear Unit
ROI	Region of Interest
SAM	S-Adenosyl Methionine
SEM	Scanning Electron Microscopy
SHG	Second Harmonic Generation
shPGS1	short hairpin PGS1
SIRT3	Sirtuin 3
SNR	Signal to Noise Ratio
SOD	Superoxide Dismutase
SRG	Stimulated Raman Gain
SRH	Stimulated Raman Histology
SRL	Stimulated Raman Loss

SRS	Stimulated Raman Scattering
SVD	Singular Value Decomposition
t-SNE	t-distributed Stochastic Nearest Neighbor Embedding
TAG	Triacyl Glyceride
TEM	Transmission Electron Microscopy
TNBC	Triple Negative Breast Cancer
TPF	Two Photon Fluorescence
UMAP	Uniform Manifold Approximation Projection
VCPA	Variable Combination Population Analysis
Vv(Mes/Glom)	Volume Fraction of Mesangium to Glomerulus
WT	Wild Type

Vita

2018 Bachelor of Science in Bioengineering, University of California San Diego
2023 Doctor of Philosophy in Bioengineering, University of California San Diego

Publications

1. Jang, Hongje, Yajuan Li, **Anthony A. Fung**, Pegah Bagheri, Khang Hoang, Dorota Skowronska-Krawczyk, Xiaoping Chen, Jane Y. Wu, Bogdan Bintu, and Lingyan Shi. "Super-Resolution SRS Microscopy with A-PoD." *Nature Methods* 20, no. 3 (March 2023): 448–58. https://doi.org/10.1038/s41592-023-01779-1.
2. **Fung, Anthony A.,** Zhang, Wenxu, Yajuan Li, Zhi Li, Hongje Jang, Honghao Zha, Xiaoping Chen, et al. "Multi-Molecular Hyperspectral PRM-SRS Imaging." bioRxiv, July 25, 2022. https://doi.org/10.1101/2022.07.25.501472.
3. Zhang, Zunming, Rui Tang, Xinyu Chen, Lauren Waller, Alston Kau, **Anthony A. Fung**, Bien Gutierrez, et al. "A High-Throughput Technique to Map Cell Images to Cell Positions Using a 3D Imaging Flow Cytometer." *Proceedings of the National Academy of Sciences* 119, no. 8 (February 22, 2022): e2118068119. https://doi.org/10.1073/pnas.2118068119.
4. **Fung, Anthony A**., Khang Hoang, Honghao Zha, Derek Chen, Wenxu Zhang, and Lingyan Shi. "Imaging Sub-Cellular Methionine and Insulin Interplay in Triple Negative Breast Cancer Lipid Droplet Metabolism." *Frontiers in Oncology* 12 (2022). https://www.frontiersin.org/articles/10.3389/fonc.2022.858017.
5. Li, Yajuan, Pegah Bagheri, Phyllis Chang, Audrey Zeng, Jie Hao, **Anthony Fung**, Jane Y. Wu, and Lingyan Shi. "Direct Imaging of Lipid Metabolic Changes in Drosophila Ovary During Aging Using DO-SRS Microscopy." *Frontiers in Aging* 2 (2022). https://www.frontiersin.org/articles/10.3389/fragi.2021.819903.
6. Li, Yajuan, Wenxu Zhang, **Anthony A. Fung**, and Lingyan Shi. "DO-SRS Imaging of Diet Regulated Metabolic Activities in Drosophila during Aging Processes." *Aging Cell* 21, no. 4 (2022): e13586. https://doi.org/10.1111/acel.13586.
7. Bagheri, Pegah, Khang Hoang, **Anthony A. Fung**, Sahran Hussain, and Lingyan Shi. "Visualizing Cancer Cell Metabolic Dynamics Regulated With Aromatic Amino Acids Using DO-SRS and 2PEF Microscopy." *Frontiers in Molecular Biosciences* 8 (2021). https://www.frontiersin.org/articles/10.3389/fmolb.2021.779702.
8. Li, Yajuan, Wenxu Zhang, Anthony A. Fung, and Lingyan Shi. "DO-SRS Imaging of Metabolic Dynamics in Aging *Drosophila*." *The Analyst* 146, no. 24 (2021): 7510–19. https://doi.org/10.1039/D1AN01638E.
9. **Fung, Anthony A.**, Andy Zhou, Jennifer K. Vanos, and Geert W. Schmid-Schönbein. "Enhanced Intestinal Permeability and Intestinal Co-Morbidities in Heat Strain: A Review and Case for Autodigestion." *Temperature* 8, no. 3 (July 3, 2021): 223–44. https://doi.org/10.1080/23328940.2021.1922261.
10. Wei, Weichen, Xuejiao Wang, Yajuan Li, Yong Cheng, **Anthony A. Fung**, Xinyi Yang, and Lingyan Shi. "Chapter Five - Advances in Optical Imaging of Drug Delivery across the Blood-Brain Barrier." In *Progress in Optics*, edited by Taco D. Visser, 66:171–253. Elsevier, 2021. https://doi.org/10.1016/bs.po.2021.01.002.
11. **Fung, Anthony A.**, and Lingyan Shi. "Mammalian Cell and Tissue Imaging Using Raman and Coherent Raman Microscopy." *WIREs Systems Biology and Medicine* 12, no. 6 (November 2020). https://doi.org/10.1002/wsbm.1501.
12. **Fung, Anthony A.**, Shi, Lingyan, and Zhou, Andy. "Advances in Stimulated Raman Scattering Imaging for Tissues and Animals." *Quantitative Imaging in Medicine and Surgery* 11, no. 3 (March 2020): 1078–1101. https://doi.org/10.21037/qims-20-712.

Abstract of the Dissertation

Development of a Label-free Optical Analysis Platform for Cells and Tissues

by

Anthony A. Fung

Doctor of Philosophy in Bioengineering

University of California San Diego, 2023

Professor Lingyan Shi, Chair

In instances where tissue availability is limited or surgical risks to the patient are significant, the application of multiplexed biomedical imaging can be a challenge due to limitations in spectral bandwidth of fluorophores and histological stain chroma. This dissertation presents a label-free multimodal optical imaging platform that combines stimulated Raman scattering (SRS), second harmonic generation (SHG), and multiphoton fluorescence (MPF) in a single microscopy setup. Using this home built multimodal microscopy, the dissertation investigates subcellular organelle morphology, their molecular composition, and metabolic dynamics in triple negative breast cancer cells under tandem nutrient modulation,

as well as assesses various biomarkers such as 3D mesangial expansion, collagen fiber thickening, oxidative stress, and lipid dysregulation in diabetic kidney disease. From subcellular to tissue levels, the platform provides spatial, biomolecular, and metabolic insight across disease models without relying on exogenous labels or serial sectioning. Developing and employing analytical techniques such as chemometric spectro-microscopy such as relative entropy, penalized reference matching, and stimulated Raman histology, this work advances quantitative bioimaging while preserving the flexibility of traditional methods through its non-destructive properties. Ultimately, this custom platform optimizes informatics while minimizing sample utilization and disruption, unlocking novel biological insights that were previously reliant on more intricate sample preparation, expensive reagents, and abundant tissue availability.

Chapter 1: An Introduction to the Label-free Optical Platform and its Development

This chapter first introduces the relevant concepts and technologies, highlighting recent developments in the literature that champion biomedical applications. The following section presents the development of tools and methodologies employed by the label-free optical platform in subsequent chapters. Experimental approaches and considerations related to platform development are then presented, and the chapter concludes with a discussion of the dissertation aims.

1.1 Label-free Imaging Concepts and Technologies

Spontaneous Raman scattering is a non-destructive process that involves the inelastic scattering of light by molecules, resulting in a shift in the frequency of the scattered photons. The shift in frequency corresponds to the vibrational energy of the molecule, providing information about the chemical composition and molecular structure of the sample. Raman scattering was first demonstrated by Indian physicist C.V. Raman, who in 1930, received the Nobel Prize in Physics for his discovery. As a spectroscopic technique, it has been used extensively in various fields including chemistry, physics, and biology and has undergone significant development in instrumentation and analytical methods. In recent years, Raman spectroscopy and stimulated Raman scattering microscopy has gained attention in the biomedical field due to its non-invasive and label-free nature. The technique has been used for a range of applications, including cell characterization [1, 2], disease diagnosis [3, 4], drug development [5, 6], and tissue imaging [7].

Raman scattering is a natural phenomenon and is referred to as “spontaneous” because the scattered light is emitted spontaneously, without the need for an external stimulation or excitation. When incident radiation in the form of a monochromatic laser with known wavelength interacts with the molecules of interest, there is a transfer of energy due to the electric field of the

incident laser photons. This transfer can either decrease or increase the scattered photon energy, and is referred to “Stokes” (red-shifted Raman spectrum) and “anti-Stokes” (blue-shifted Raman spectrum), respectively, and depends on whether the molecules were in a ground or excited state initially. This dissertation exploits Stokes Raman scattering by measuring the scattering frequencies excited from a ground state.

A simple diatomic model is often used to describe energy levels. In this model, the atoms and electrostatic forces are represented as bodies with mass *m* and position *u*, connected by a massless spring of constant *k*. The photons electronic field induces a temporary change in polarizability by inducing an internuclear displacement *Q*. In reality, biomolecules of interest are more complex and contain many bonds, but the harmonic oscillation of the model system can be described by classical equations 1.1-1.3 below, assuming Maxwell-Hookean mechanics. The set of forces acting at all points of the system is a function of the stiffness *k*, and flexibility constants *c* in the direction of the applied forces. The total work of the system *W* is stored as energy, and is both positive and definite.

$$\boldsymbol{F_i^{(e)} = m_i \ddot{u}_{ii} + \sum_{j=1}^{n} k_{ij}\, u_j} \qquad \textbf{1.1}$$

$$\boldsymbol{Q_i = k_{in} c_{in} Q_n} \qquad \textbf{1.2}$$

$$\boldsymbol{W = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}\, Q_i Q_j} \qquad \textbf{1.3}$$

Here we can clearly see the work done on the system is proportional to the square of the symmetric internuclear distance *Q*. It is important to note that because energy is quantized and the exact position of an electron is unknown at any time, energy is represented not as a continuous function, but instead as probabilistic quantities. Figure 1.1 below shows an energy diagram of Stokes Raman scattering of an anharmonic oscillator model of this system. A perfect harmonic oscillator is shown by the dashed lines, but is not representative of an actual diatomic

molecule because bond breakage occurs beyond a certain internuclear distance, unlike the Maxwell-Hooke models used to describe it.

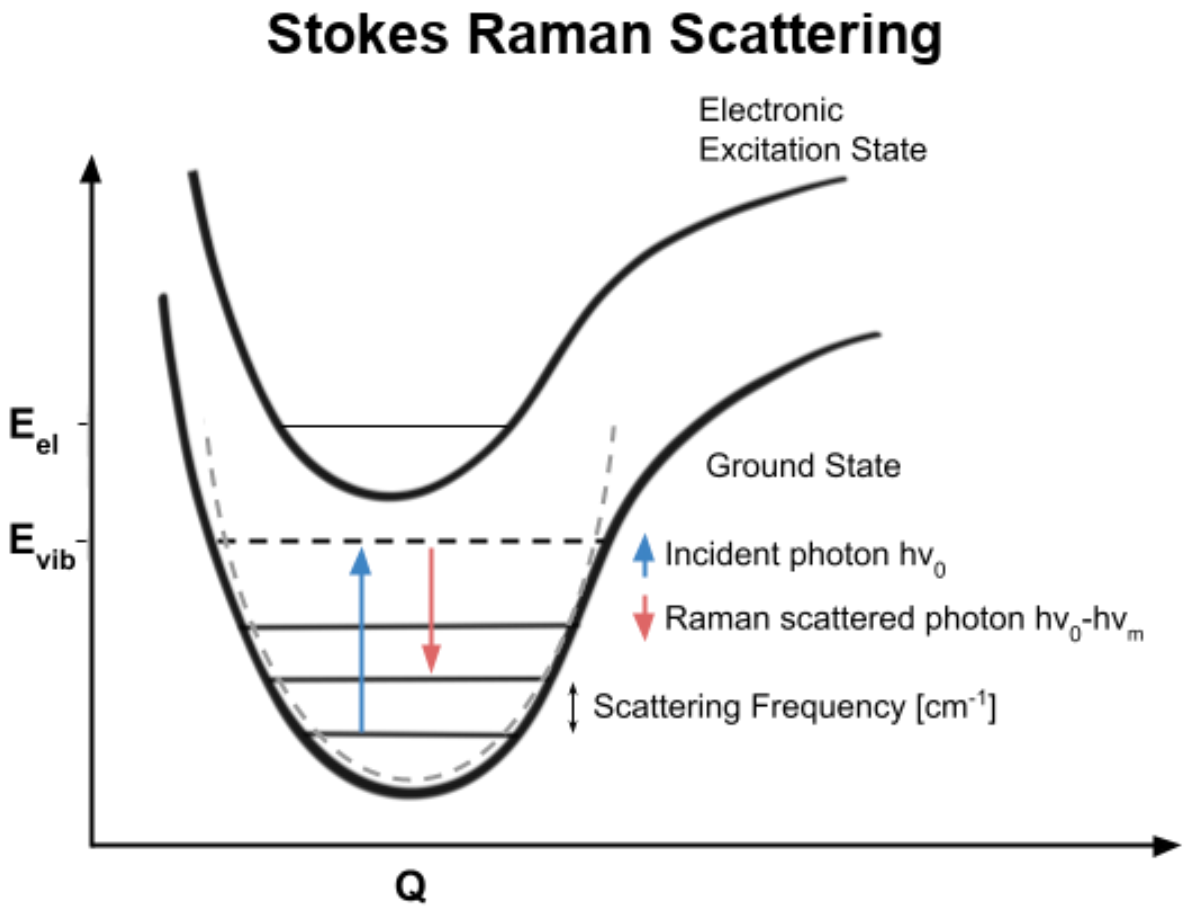


Figure 1.1 Energy diagram of Stokes Raman scattering
The work stored as strain energy is quadratically proportional to the internuclear distance, Q, of the diatomic molecule model. The perfect harmonic oscillator case is shown as a dashed parabola, but solid lines are represented to include the potential of bond breakage beyond a certain internuclear distance. The incident photon has a much higher energy than the scattered photon, and excites the molecule to a virtual energy state, E_{vib}. The scattering frequency is the difference in energy between the energy levels, and the Raman shift, expressed in wavenumbers is proportional to the change in frequency v_0-v_m divided by the speed of light.

The amount of energy transferred depends on the molecules of interest and their rho-vibrational modes. These modes refer to the polarizability of the molecules and degrees of freedom in their structures. A more polarizable molecule will have greater Raman scattering capability because incident radiation will cause a change in the induced electric field by exciting electrons to a virtual energy state. These states are very temporary and do not correspond to actual energy levels of a molecule that emit new photons during relaxation. There are many ways a molecule can vibrate such as symmetric or asymmetric stretching, scissoring, bending, rocking, and twisting. In contrast to infrared (IR) spectroscopy, which also probes these vibrational modes, Raman scattering focuses on scattering and not absorption. Additionally, Raman spectroscopy is more sensitive to polarizability changes, while IR spectroscopy is more sensitive to changes in dipole moment, allowing for better detection of certain molecules. In either case, the result is a spectrum in which the abscissa is the difference between the excitation wavelength and the scattered wavelength, called a Raman shift, and is represented by wavenumbers [cm^{-1}]. The ordinate is proportional to the number of scattering molecules *N*, the scattering cross section (the polarizability, α and vibrational amplitude, *Q*), the laser frequency, and incident laser power. This is summarized in equation 1.4 below.

$$I_R \propto \nu^4 I_0 N \left(\frac{\partial \alpha}{\partial Q}\right)^2 \qquad \textbf{1.4}$$

These spectra can be used to identify the molecular composition of an analyte. However, this process is slow, relative to fluorescence imaging, due to the weak Raman scattering cross section. That is, the probability for a molecule to inelastically scatter light is low, which results in longer acquisition times and higher laser powers. Much progress has been made in the effort to improve Raman signals, including advances in detectors, resonance enhancement using nanoparticles and substrates, and of course non-linear stimulation.

1.1.2 Stimulated Raman Scattering Microscopy

Unlike spontaneous Raman scattering, a Stimulated Raman Scattering (SRS) system uses two lasers, a pump beam and a Stokes beam. These beams are overlaid in space and time, and the pump beam excites the specific vibrational frequency while the Stokes beam stimulates the scattered photons in a coherent fashion. Since both beams of different energy are involved in the vibrational transition of the analyte molecule, the signal intensity is greater than the spontaneous process alone. This affords faster scanning times of about three orders of magnitude [8]. For this reason, SRS is more appropriate for acquiring image data than spontaneous Raman scattering. SRS can be achieved either by measuring stimulated Raman gain (SRG) or stimulated Raman loss (SRL). In this document, only SRL, which measures the attenuation of pump beam energy and filters out the Stokes beam, was used for SRS imaging. While both SRG and SRL are considered “background-free” due to their coherent nature, SRG may require a reference measurement of the pump beam energy. Both spontaneous and stimulated Raman scattering are chemometric techniques that can generate Raman spectra. Spontaneous Raman scattering captures a full Raman spectrum at a single user-defined point on a sample, while SRS captures a single Raman shift at many spatial points. A hyperspectral image (HSI), a three-dimensional array in which each pixel of the image is a spectrum of intensities, can be acquired with either SRS or spontaneous Raman scattering, but there is typically a trade-off between spectral resolution, spatial resolution, range, and imaging speed. In this document, HSI’s are acquired using SRS due to the faster imaging speeds, and encompass the CH stretching region with a spectral resolution of about 10-20 cm^{-1}. Full Raman spectra are acquired using a commercial spontaneous Raman scattering microscope (XPloRa plus, HORIBA) with a spectral resolution of 2-5 cm^{-1}.

SRS serves multiple functions in this optical platform including the analysis of metabolic activity, pseudo-colored digital histology, and label-free multimolecular detection. As such, several steps were necessary in the development of the platform.

1.1.3 Two-Photon Fluorescence Microscopy

Two-photon fluorescence (TPF), like SRS, is a second-order non-linear optical process and involves the simultaneous absorption of two photons, each with half the energy of the corresponding single-photon absorption event. The emission intensity is proportional to the square of the excitation intensity, which allows for weaker incident laser powers and thus mitigates photodamage and photobleaching effects [9]. These harmful effects can permanently alter structural information through sample burning, or lose biomolecular information through conformational changes in fluorescent molecules that prevent them from fluorescing. Penetration depth generally increases when using longer wavelengths which are scattered less efficiently, and spatial resolution improves because both photons must be intersected at the same point in space and time to excite the fluorophore.

These advantages are integral for label-free imaging because endogenous compounds like NAD[P]H and Flavins have a weaker fluorescence intensity than many commercial probes. This phenomenon is known as autofluorescence, a property shared among all cells, and upon which this dissertation capitalizes.

1.1.4 Second Harmonic Generation Microscopy

Another second order non-linear optical process is second harmonic generation (SHG). Here, two photons of equal energy interact with non-centrosymmetric material through an oscillating electromagnetic field and emits a new photon with twice the energy of the incident photons. In this document, all SHG imaging uses a two photon 1031 nm laser and a 515 nm filter to probe type I – III collagens, which are highly ordered structures that lack centro-symmetry.

Collagen can be an indicator of several diseases such as renal fibrosis, fatty liver disease, glaucoma, and melanoma. Due to the wide variety of applications, this optical platform incorporates methods to characterize collagen through SHG microscopy such as anisotropy analysis [10] and fiber thickness analysis [11–13].

1.2 Developing a Label-free Optical Imaging Platform

1.2.1 Relative Entropy Development for Spontaneous Raman Spectroscopy

The output of the spontaneous Raman scattering microscopy employed in this document takes the form of a spectrum. This spectrum describes the molecular composition at a point in the sample selected by the user, and can be used to characterize experimental groups through changes in Raman intensity in various regions of the spectrum. Experiments with a large number of experimental groups and many independent variables may be sensitive to multivariate effects. As a result, plotting all spectral groups and comparing them in a fully-connected pair-wise manner can be cumbersome and less quantitative.

In developing this optical platform, multivariate problems demand objective solutions. At the time of writing, the current literature often highlights spectroscopic differences between experimental groups arbitrarily. This also entails that every spectrum or group of spectra be displayed juxtaposed, which can be cumbersome in experiments with many groups and independent variables. Put simply, differences in Raman peak intensities between the mean spectra of two experimental groups do not imply significance. In multiplexed experiments with many independent variables, discerning which Raman features contribute most to the discrimination of experimental groups is even more challenging. Techniques such as t-distributed stochastic neighbor embedding (t-SNE), principal component analysis (PCA), and uniform manifold approximation projections (UMAP) are non-linear dimensionality reduction tools to visualize how distinct groups of data, like Raman spectra, can be separated. However, they do

little to explain which variables within the data contribute to these separations. To quickly rank and visualize all the wavenumber variables that may have been influenced by a particular treatment, the Kullback-Leibler divergence (D_{KL}), a simple metric for the distance between two distributions for classification problems [14], at each Raman shift can be plotted for each independent variable. This method is also known as relative entropy.

For binary classification systems, the amount of data lost in classifying data B as data A is described by the one-dimensional cross entropy equation 1.5 below.

$$\boldsymbol{H(\mathbf{A},\mathbf{B}) = -\sum_{i}^{n}(p_{\mathbf{A}}[\![(v]\!]_{\mathbf{i}})\log p_{\mathbf{B}}(v_{\mathbf{i}}))} \quad \textbf{1.5}$$

where, *H*(A,B) is the cross-entropy, $p_{\mathrm{A}}(v_{\mathrm{i}})$ and $p_{\mathrm{B}}(v_{\mathrm{i}})$ are probability vectors from the distributions of intensity values a wavenumber variable v_{i}. Probability vectors can be derived from various distributions, but only Gaussian (or, normal) distributions were used in this study. The D_{KL} is related to cross-entropy as it is the additional entropy beyond the entropy of the data A. Since both distributions are already labeled and we are not interested in generating probability vectors, but rather supply them, the D_{KL} is described by Equation 1.6 below.

$$\boldsymbol{D_{KL}(\mathbf{A}||\mathbf{B}) = \sum_{i}^{n}\left(p_{\mathbf{A}}(v_{\mathbf{i}})\log\left(\frac{p_{\mathbf{A}}(v_{\mathbf{i}})}{p_{\mathbf{B}}(v_{\mathbf{i}})}\right)\right)} \quad \textbf{1.6}$$

This divergence is made symmetric by equation 1.7 below. For this analysis probability vectors are of length 10^7

$$\boldsymbol{D_{KL} = \frac{D_{KL}(\mathbf{A}||\mathbf{B}) + D_{KL}(\mathbf{B}||\mathbf{A})}{2}} \quad \textbf{1.7}$$

For multiclass situations in which the number of groups exceeds two, the average divergence is calculated following equation 1.8 [15, 16].

$$D(P_1 \dots P_k) = \frac{1}{k(k-1)} \sum_i^k \sum_j^k D_{KL}(P_i || P_j) \quad \textbf{1.8}$$

Relative entropy is an excellent method for quantifying the relative importance of a wavenumber in the discrimination of spectral datasets because it is computationally fast and capitalizes on the tremendous chemometric potential afforded by the spectral resolution of modern spontaneous Raman systems. Below shows an example in which two experiments each compares four groups of Drosophila antennal lobe tissues along two dimensions, sex and age, and sex and genetic strain, respectively. In the first experiment, t-SNE plots, initialized with PCA and preserved global structure, are able to display the separation of groups with the age dimension having greatest separation between clusters (**Fig. 1.2A**). However, even with this dimensionality reduction, it is still not clear which Raman features contribute to this separation until the relative entropy is plotted along these dimensions (**Fig. 1.2C**). By regrouping the spectra along these dimensions and comparing them using relative entropy, the amide I, amide III, and saccharides Raman bands appear to have great discriminatory power, and warrant further investigation. Likewise, when the amyloid precursor protein (APP) overexpressed Drosophila tissues are grouped along the sex and strain dimensions, the relative entropy shows slightly higher discriminatory power in the amide I and amide III regions as well. However, as reflected in the t-SNE plot (**Fig. 1.2B**), the overall relative entropy is lower in this experiment (**Fig. 1.2.D**).

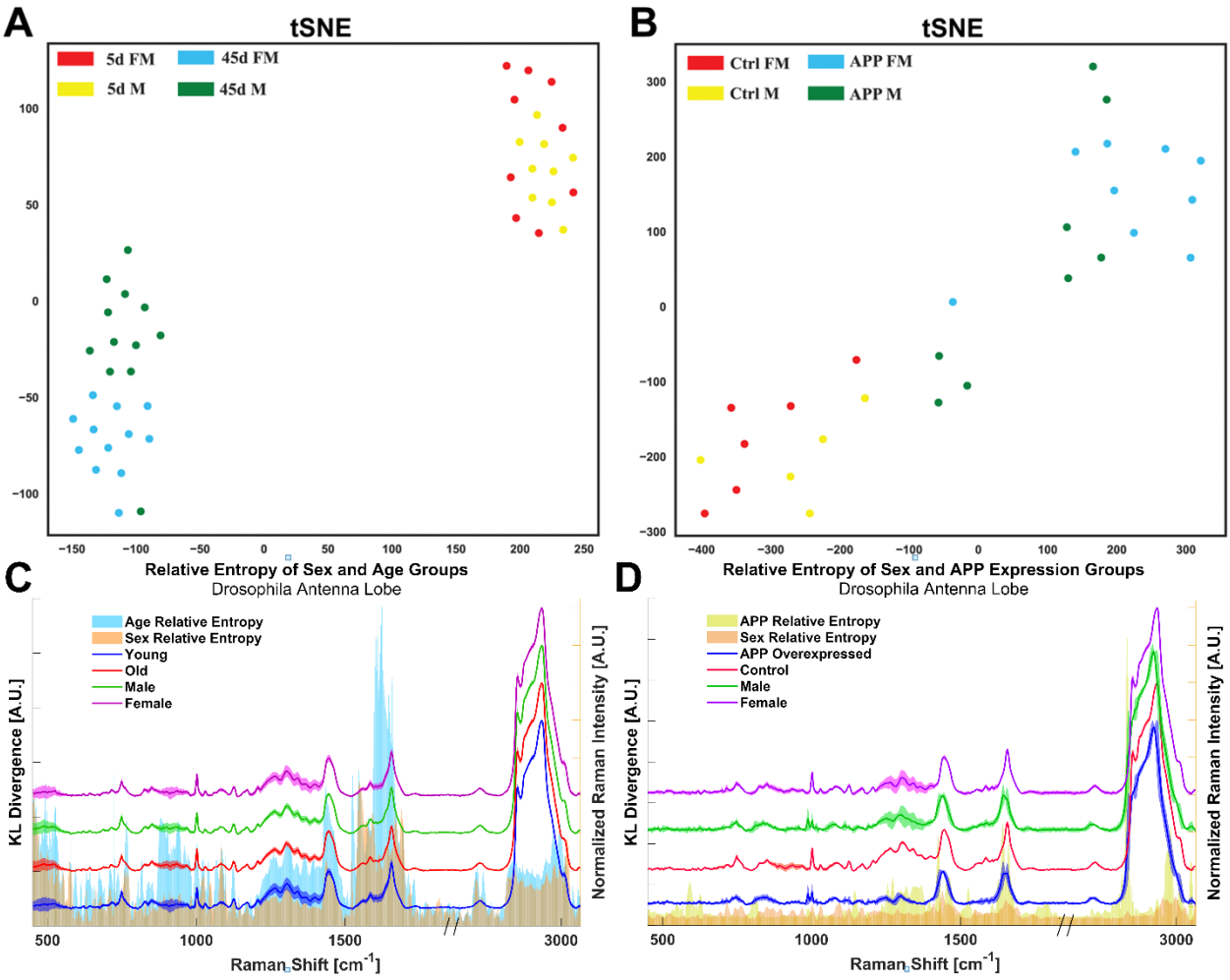


Figure 1.2 Relative Entropy in Raman Spectroscopy
(A) t-SNE analysis was conducted on the AL Raman spectra from 5-day (5d) and 45-day (45d) female and male flies, respectively. PCA initialization was performed, and global structure was preserved. **(B)** t-SNE analysis was conducted on the AL Raman spectra from WT and APP overexpression female and male flies. M, male; FM, female. PCA initialization was performed, and global structure was preserved. **(C)** Relative entropy plots of age and sex dimensions with overlaid plots of grouped and normalized spectra with 1-sigma error bands. The Amide I region has the highest entropy among the spectral features, and an overall higher entropy along the age dimension. **(D)** Relative entropy plots of APP expression and sex dimensions with overlaid plots of grouped and normalized spectra with 1-sigma error bands. The APP expression dimension has an overall higher relative entropy than the sex dimension, with the highest Raman feature being the Amide I band.

1.2.2 Multimodal Metabolic Feature Selection

To detect metabolic activity, deuterium oxide (D_2O), also known as heavy water, was used as a bio-orthogonal probe. Previous studies have shown that sparse metabolic incorporation of deuterium (D) in macromolecules such as lipids, proteins, and DNA allows for their visualization

in the Raman cell silent region [17]. This is because carbon-deuterium (C-D) bonds have a unique vibrational mode around 2150 cm^{-1} to 2200 cm^{-1}, allowing for the visualization of newly synthesized macromolecules. Along with this metabolic feature, other Raman features were selected as described in figure 1.3 below, and are elaborated in Chapter 2.

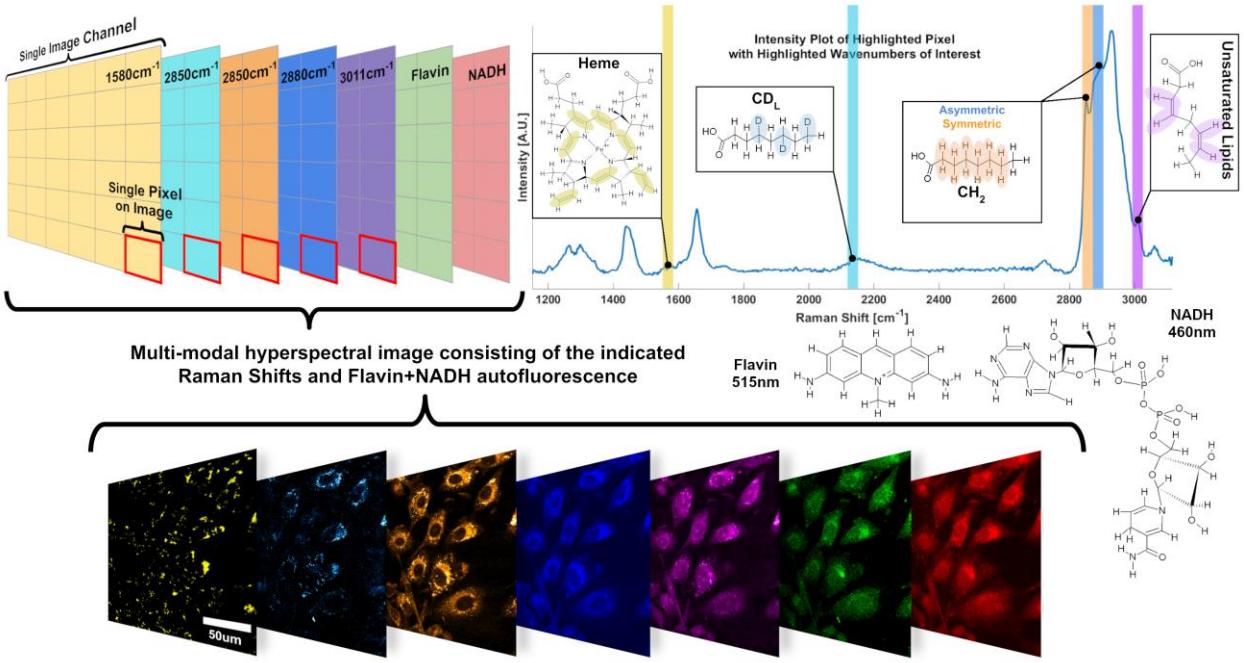


Figure 1.3 Multimodal Imaging Feature Selection
Hyperspectral Image (HSI) format is shown conceptually to convey the multi-modal approach to quantitative optical analysis. Vibrational modes are color coded, with an example image of a HSI of MCF10A control cells.

1.2.3 Stimulated Raman Histology

One of the most widely used pathological detection methods is histology. Cancers, infections, neurodegeneration, and arthritis are among the many diseases that can be visually detected using stained biopsies. Despite the years of training required of histopathological diagnoses, there is a drawback. Limited biomolecular insight. Even with more recent immunofluorescence probes, there is a limit to the number of simultaneous markers due to the overlap of stain hues and chroma, and fluorescent probe excitation and emission wavelengths.

Modern workflows can incorporate sequential wash steps to bind stains and probes individually, but these multiplexing methods can be costly in time and money. Additionally, they have certain drawbacks such as altered molecular distribution due to the size and hydrophobicity of the probe itself, or the reagents and washing processes themselves. It is common practice to serially slice tissue and independently stain each slice, but the morphologies are not the same, and may be subject to deformations that can make co-registration of these slices difficult.

To address this issue, stainless staining can be employed. Stainless staining is a form of digital histology that pseudo-colors the sample image based on spectrally acquired data. This can be done using fluorescence, FT-IR imaging, SRS, and others. With respect to SRS, it is referred to as stimulated Raman histology (SRH) [18]. Using this method, various views of the same sample can be generated without any physical staining. Since a major tenet of this optical platform is that it maintains the flexibility of traditional methods, and does not preclude these methods, SRH is a critical part of this platform development.

Two common histological stains, hematoxylin and eosin (H&E) and periodic acid–Schiff (PAS), were tested. First, SRS HSIs of the CH stretching region were acquired, followed by physical staining of the samples (**Fig. 1.4 A, D**). Images were carefully co-registered by hand, but future experiments may employ useful tools such as the ImageJ plugin; stack-reg (Biomedical Imaging Group - EPFL). For the H&E stain, which is comprised of two main colors, a blue-purple nuclei stain (Hematoxylin), and a pink interstitial stain (Eosin). These colors can be easily segmented in the L*a*b* color space to generate ground truth masks (**Fig. 1.4 B**). From these masks, the pixel spectra of the nuclei and cytoplasm can be visualized (**Fig. 1.4 C**). Of note, three Raman peaks of interest comprise the majority of the variance: 2850 cm^{-1}, 2880 cm^{-1}, and 2935 cm^{-1}. Eosin pixels contained a much high ratio of 2850 cm^{-1} to 2935 cm^{-1}, which is what is typically employed in the literature. Therefore, the H&E SRH modality is constructed from unmixing these Raman shifts, applying custom lookup tables (LUTs), blending them, and normalizing them. First, the protein (2935 cm^{-1}) and lipid (2850 cm^{-1}) SRS images are unmixed using previously published

methods [17]. Nuclei were segmented from the ratiometric result by thresholding pixels with the highest lipid to protein ratio. The nucleic mask is then de-speckled, closed, and hole-filled using standard ImageJ binary functions. Custom LUTs are applied to the unmixed protein and lipid images and blended together in the RGB space in a similar fashion to other studies [18, 19]. To ensure the nuclei and background have enough contrast, the pixels within these masks have their L* decreased and increased, respectively, and their b* chromacity slightly decreased. Finally, RGB histogram normalization may be applied to ensure clarity and consistency [20].

For the PAS SRH implementation, the stained ground truth images were preprocessed by enhancing contrast (raising the lower threshold by 10% and decreasing the upper threshold by 10%), and increasing the blue channel's chroma and saturation by 10% each (**Fig. 1.4 D**). This is done to provide a greater separation between the nuclei counterstain and the PAS aldehyde chroma. Once the image is converted to the L*a*b* color space, the PAS can be readily analyzed (**Fig. 1.4 E**). Unlike the Eosin, which stains almost everything the same color, and can then be scaled in brightness proportionally with the Raman intensity, the hue and chroma of the PAS stain is more of a gradient. Period acid first oxidizes glycol groups found on saccharides in glycoproteins and mucin such as in basement membranes. The aldehyde product can then bind with Schiff's reagent, and upon the release of a sulfonic group in washing steps, a reddish hue develops. The chroma of this hue is more red where the concentration of Schiff's reagent is higher, however it unfortunately doesn't scale linearly with Raman intensity. In the L*a*b* space, the areas with higher PAS staining correlate logarithmically with the difference between a* and b* chroma, as indicated by the red arrow in **Fig. 1.4 E**. Several snapshots of the thresholded PAS image along this red arrow are also shown for reference. Plotting the pixel spectra from various groups of a*-b* values shows that the PAS stain correlates logarithmically with the ratios of 2850 cm^{-1} and 2880 cm^{-1} to 2935 cm^{-1}, as well as with the AUC of the Raman spectra.

It should be noted that other color spaces such as Hue Saturation Value (HSV) may also be used, however the L*a*b* space was implemented because Cartesian coordinates are easily

interpretable with Raman intensity, and the histological stains employed are well suited to the a* and b* chromaticity representation. These axes measure green-magenta, and blue-yellow opponent colors, respectively. The chroma becomes redder with positive increasing a*, and bluer with negative decreasing b*. For images with another color gamut, a different segmentation method may be necessary. Unlike the more common RGB and CMYK color-spaces, the L*a*b* color-space is device independent. This offers advantages in color normalization methods to standardize datasets and establish uniform stain-augmentation in the future – a critical consideration for co-registration of spatial molecular imaging modalities.

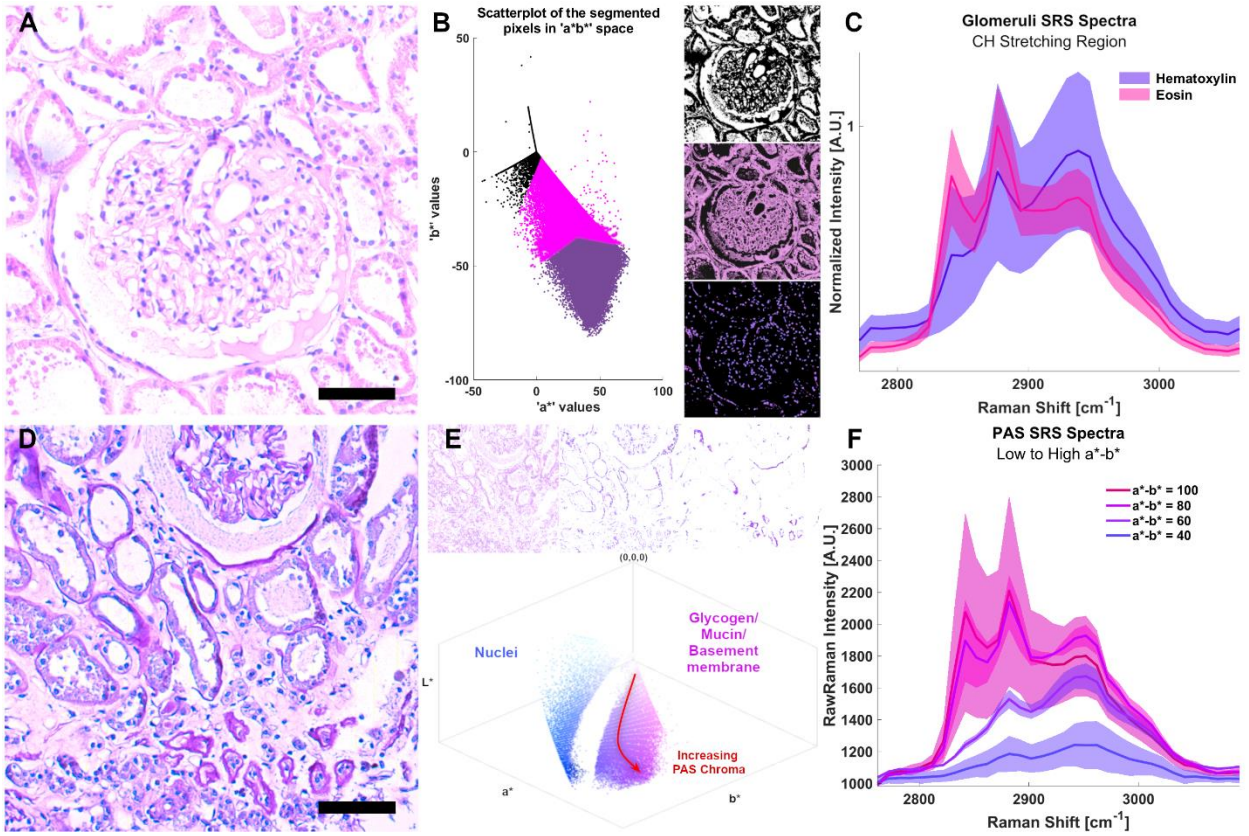


Figure 1.4 Stimulated Raman Histology Platform
(A) After a HSI is acquired, the same sample is stained using Mayer's H&E (Thorlabs). Scale bar 50 microns. **(B)** The H&E stained image serves as the ground truth, and is converted to the L*a*b* color space to segment background, eosin, and hematoxylin masks. **(C)** The masks are used to acquire pixel spectra from the SRS HSI, which are plotted with 1-sigma error bands. **(D)** After a HSI is acquired, the same sample is stained with PAS stain. Scale bar 75 microns. **(E)** The stained image undergoes [10%-90%] contrast enhancement, and the blue channel saturation and chroma is increased by 10% to make the nuclei more distinct. This image serves as a ground truth and is then converted to the L*a*b* color space. The mucin stain becomes a gradient, with logarithmically increasing mucin chroma along the a*-b* axis. Snapshots of the L*a*b* thresholded image along the red arrow are shown above. **(F)** Several groups of pixels with indicated a*-b* quantities are plotted with 1-sigma error bands.

Nuclei may also be segmented in the phasor space, which transforms a spectrum into a real and imaginary component that correspond to amplitude and phase, respectively. This technique has been demonstrated in figure 1.5 below using methods published in the literature [21].

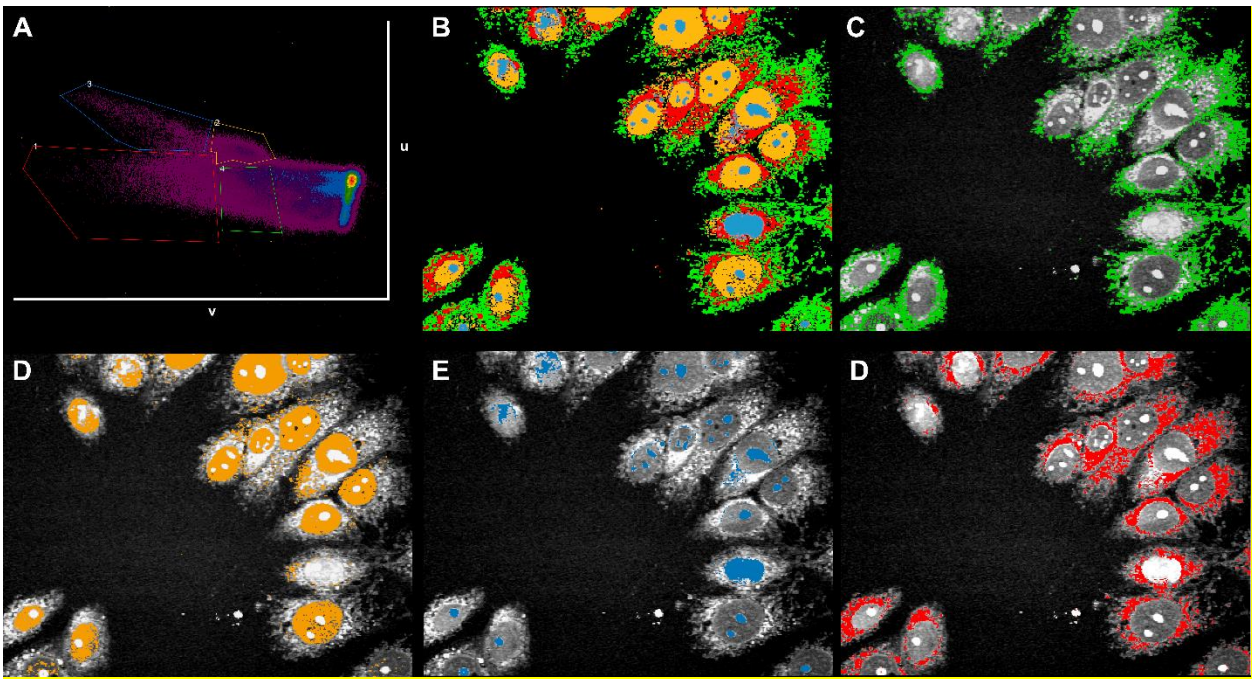


Figure 1.5 Phasor segmentation can isolate nuclei and other components
(A) Phasor plot of pixel spectra with segmented regions drawn. **(B)** An overlay of the segmented pixel masks from (A). **(C)** Cytoplasmic component segmented via phasor only. **(D)** Nuclei component segmented via phasor only. **(E)** Nucleoli component segmented via phasor only. **(F)** Lipid droplet component segmented via phasor only. Scale bar 10 µm.

1.2.4 Penalized Reference Matching

Mapping spatial distributions of biomolecules remains a persistent challenge for any optical platform, especially for label-free platforms. Spatial molecular visualizations can be achieved through destructive methods such as Matrix Assisted Laser Adsorption Deposition Ionization (MALDI), gas chromatography mass spectrometry (GCMS), or digestive assays. These methods allow the user to directly measure the presence of a given standard analyte within the sample, at the cost of permanently altering the sample. Recent work has demonstrated label-free

unmixing of molecules using multivariate curve resolution alternating least squares (MCR-ALS),which allows the user to unmix a SRS-HSI into probabilistic or proportional component spectra to visualize relative concentrations [22]. Albeit label-free, there is no guarantee that the unmixed component spectra correspond to an analyte of interest. Furthermore, the processing time can take 30 minutes to process a 512 x 512-pixel hyperspectral image and presupposes the number of chemical species in a sample. The result displays a pixel's identity by its relative proportional composition of reference species. However, this is not always feasible in a complex biological sample. Singular Value Decomposition (SVD) can estimate the number of components, but analytical results may be sensitive to slight deviations from the exact number of components. Therefore, there is sorely needed solution that allows for direct visualization of a selected analyte like MALDI-TOF, while retaining the label-free non-destructive abilities of MCR-ALS. Finally, the solution should be robust enough to account for differences in spectral position that may vary across instruments, such as between the spontaneous Raman spectrometer and SRS microscope.

To address these issues, spectral angle mapping, also called reference matching or cosine similarity, was modified with a penalty term to allow, but penalize, small differences in spectral position. This method allows for the selection of a reference spectrum, such as that of an analytical standard or a specific lipid subtype, acquired by a spontaneous Raman spectrometer, and visualize relative concentrations within a HSI acquired by a SRS microscope.

SRS HSI pixel spectra and reference spectra were linearly interpolated such that the spectral interval is 1 wavenumber. This ensures that the inner product, which requires vector dimensions to be the same, is possible. After all spectra were adjusted to the same interpolated resolution, they were simplex normalized using equation 1.9,

$$\boldsymbol{I_1 = \frac{I - I_{min}}{I_{max} - I_{min}}} \qquad \textbf{1.9}$$

where I_{min} is the minimum value and Imax is the maximum value in the pixel spectrum. This normalization is done prior to reference matching so that the process generates results that are solely based on spectral shape without being affected by any intensity fluctuation. The normalized pixel spectra were then divided by their Euclidean norm as shown in equation 1.10

$$\boldsymbol{I_2 = \frac{I_1}{||I_1||_2}} \quad \textbf{1.10}$$

where I_2 is the interpolated signal of the pixel spectrum. Reference spectra from spontaneous Raman acquisitions follow the same pre-processing steps as the HSI pixel spectra, as shown in equations 1.11 and 1.12 below.

$$\boldsymbol{I_4 = \frac{I_3 - I_{3,min}}{I_{3,max} - I_{3,min}}} \quad \textbf{1.11}$$

$$\boldsymbol{I_5 = \frac{I_4}{||I_4||_2}} \quad \textbf{1.12}$$

where I_3 denotes spontaneous Raman spectra, and I_5 is the interpolated signal of the reference spectrum. Due to the nature of Raman spectral intensity, similarity scores between each pixel spectrum and the reference spectrum were calculated using the dot product of I_2 and I_5.

These spectroscopic methods have been deployed for several decades, but due to high false positive rates, direct label-free characterization of multiple lipid subtypes in cells and tissues has not been achieved by using optical imaging. To address this, we added a penalty term to the canonical cosine similarity algorithm, which decreases the false positive rates by proportionally reducing the similarity score with the positional discrepancy to the best spectral match (Fig. 2A-C). This process is summarized as

$$\boldsymbol{score = \max_{i \in [1,N]} (u_i \cdot v - \alpha \Delta {x_i}^2)} \quad \textbf{1.13}$$

where $\boldsymbol{u}$ represents the interpolated signal of a pixel's shifted spectrum at various positions; $\boldsymbol{v}$ represents the interpolated signal of the reference spectrum; $\boldsymbol{\alpha}$ is the penalty coefficient, with a unit of [cm^2]; $\boldsymbol{\Delta x_i}$ is the deviation in position of the spectrum in $\boldsymbol{u}_i$ from the initial observed position; and $\boldsymbol{N}$ is the number of interpolated signals, which depends on the spectral resolution of the HSI.

The penalty term $\alpha\Delta x^2$ inherently addresses the slight positional deviations due to the diverse chemical environment, as well as the variations in instrumentation (such as thermoelectric noise, lensing, and other interference). Without this term, even if the spectral shape of a pixel matches the reference spectrum, the final similarity score may still be low when the positions of the peaks differ greatly. An example is shown below, in which two spectra are compared with varying relative positions (**Fig. 1.6 A-B**). With the penalty term, all pixel spectra are evaluated as they occur at multiple Raman shifts, and the highest similarity score is returned in a pixel-by-pixel manner. In this case, the maximum score was 0.9866.

By leveraging positional information in addition to peak shape, the breadth of similarity score is increased or decreased, akin to a change in contrast. This ensures that pixels with similar shapes and positions are scored accordingly.

Most images collected in this study were taken from the Raman CH stretching region (2700 cm^{-1} to 3150 cm^{-1}) with 75 total Raman shifts (a spectral distance of 6 cm^{-1} between images). The position deviation Δx was the shift of peaks in the spectrum. We assessed several values for the penalty coefficient and chose $\alpha = 1 \times 10^{-4}$ empirically. This is because if the penalty is too low, the pixel and reference spectra are free to shift themselves relative to each other until the highest similarity score is returned, no matter how far that shift may be from the original position. For full justification and examples, see Chapter 3.

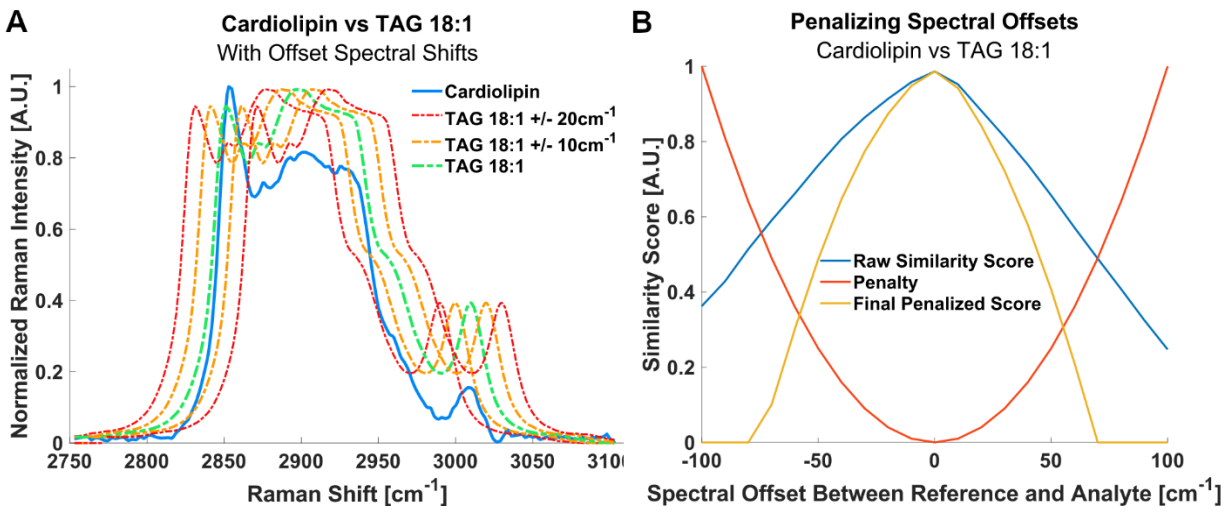


Figure 1.6 Penalized Reference Matching Principle
(A) Two spectra of lipid standards, Cardiolipin (Blue) and TAG 18:1 (Green), are plotted, along with spectrally shifted relative positions along the wavenumber axis to simulate positional differences between microscopy setups. **(B)** The similarity score is calculated using the above equations at every spectral shift in (A), and plotted (Blue). The penalty term is proportional to the square of the spectral shift, with an amplitude of 10^{-4}. The penalty term is plotted in red. The final penalized similarity score is calculated by subtracting the penalty term from the raw similarity score, and is plotted in yellow. Since the similarity score cannot be negative, all negative scores are automatically zero.

Since Raman spectra contain molecular bond information that correlates with concentration, similarity scores may be used to estimate the relative levels of different molecules, such as different lipid subtypes. When the Raman spectra of a sample exhibit a high degree of similarity to that of a reference standard, it will suggest a higher concentration of that reference molecule in the sample. Different biomolecules may have the same types of chemical bonds, and the cumulative mixture of various molecules may result in a spectrum that displays the same spectral shape as an unrelated molecule. From a macromolecular perspective in biological samples, however, we find that factors such as the diversity of the analyte composition do not necessarily void the correlation between relative ratios and similarity scores. Validation trials using lipid standards purchased from Sigma Aldrich show that various ratios of lipid standards are reflected in the similarity score under ideal conditions with two lipid standards (**Fig. 1.7 A-B**), with high background contamination (**Fig. 1.7 C-D**), and with three lipid standards (**Fig. 1.7 E-F**). Importantly, these measurements correlate with relative, not absolute, concentration, and are thus more visually intuitive when normalized (**Fig. 1.7 F**).

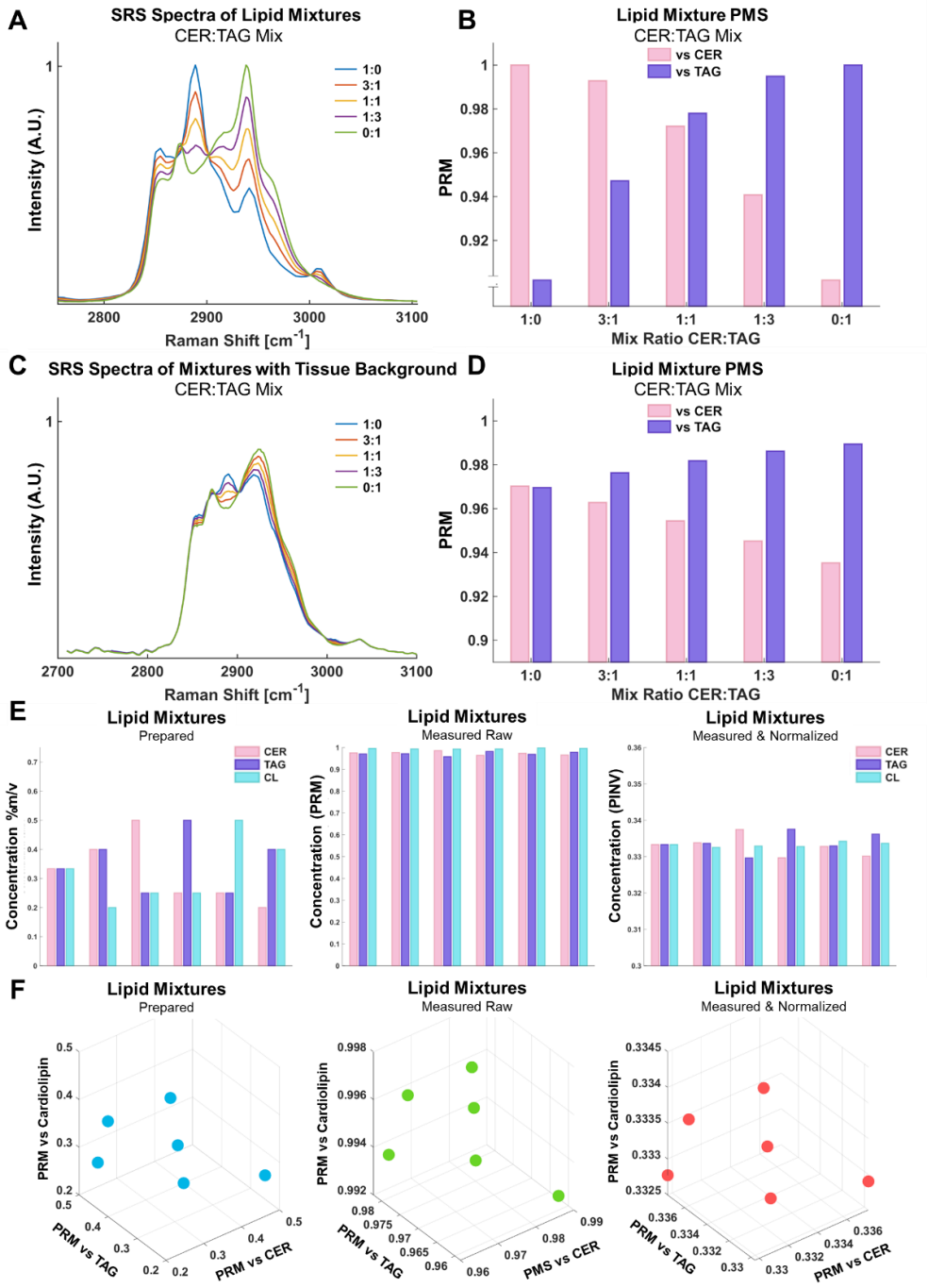


Figure 1.7 PRM Measures Relative Lipid Concentrations

(A-B) Spontaneous Raman spectra and corresponding similarity scores of mixtures of ceramide and TAG (C2-C10) (Sigma 17810-1AMP-S) at the indicated ratios. **(C-D)** The experiment repeated with the addition of equal volume of methanol-washed tissue lysates, showing a deterioration of similarity scores but a preservation of trend. **(E)** Experiments with three-lipid mixes (CER, TAG and cardiolipin) at various ratios were also conducted. Raw similarity scores (middle panel) and relative similarity scores normalized to the first group with equal ratios (right panel) are shown. **(F)** Plotting both raw and normalized (relative) similarity scores using each component as an axis, show similar patterns to the actual concentrations. All spectra were acquired on the same instrument (HORIBA XploRa plus). The highest similarity scores were achieved with a spectral shift of 0.

We further compared PRM with pseudo-inverse matrix (PINV) multiplication, which is another technique with similar functionality that is commonly used in spectra platforms. We found that both PRM-based similarity scores and PINV-derived coefficients correlated with relative ratios of lipids, with the PINV coefficients having a more linear relationship (**Fig 1.8 A-B**). However, the PINV calculation using the documented MATLAB function was much slower (**Fig 1.8 C**) PINV-derived coefficients can have negative values and have unbounded ranges (**Fig. 1.8 D-E**).

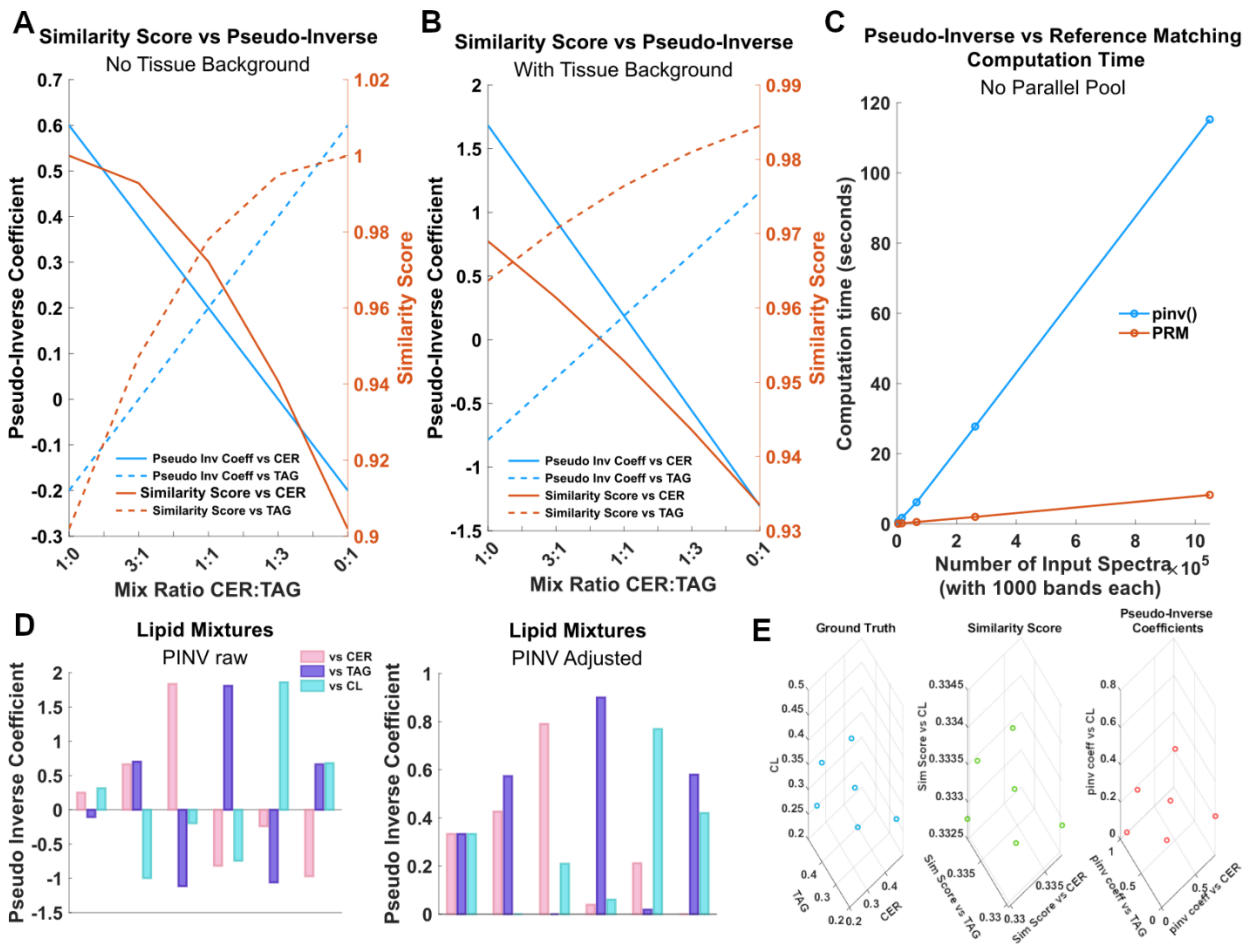


Figure 1.8 Comparison of PRM and pseudo-inverse matrix multiplication
(A) The data shown in Fig. 1.5 A-B were analyzed using pseudo-inverse (PINV) matrix multiplication and the resulting coefficients were plotted in blue. The similarity scores obtain by PRM were in orange. Pseudo-Inverse matrix coefficients have a linear relationship with relative ratios of the components but have negative coefficients and a theoretically unbounded range. **(B)** The data shown in Fig. 1.5C-D were analyzed using PINV and the resulting coefficients were plotted (blue) against those using PRM (orange). PINV shows a better linear relationship than PRM similarity scores but retains the negative and unbounded range. (C) Comparison of computation time for PRM and PINV with single processing unit and no spectral shifts shows that PINV takes much longer time to compute compared with PRM. PINV is also a one-liner that is difficult to split to parallel processing units, and does not allow for the input matrix to be split into more manageable sizes because the rank sums would not be equal. **(D)** The data shown in Fig 1.5E-F were re-analyzed using PINV, and the linear relationship is also observed, but additional rescaling steps removing negative values were required to obtain concentration-corelated results. **(E)** The three-lipid mixtures from (D) are plotted in 3D, showing that both PRM and PINV can obtain results similar to the ground truth.

To speed up the PINV calculation, one can avoid the built-in checks for invertibility through manual inverse multiplication, but the speed is still slower that inner product calculations with PRM (**Fig 1.9 A**). Additionally, the results correlate almost linearly between PRM and PINV (**Fig.**

1.9 B). Therefore it is natural to prefer PRM over PINV given typical HSI sizes can range from 512x512 pixels to 1024x1024 pixels and span hundreds of Raman shifts. Although similarity scores for Raman spectra in the CH-stretching region are typically close to 1, the variance within pixels of a pure sample is much lower (see chapter 1.3.4). This result suggests that SRS HSI are suitable for PRM analysis. Furthermore, since HSI have various pixel sizes, the spectral dataset is subject to corresponding changes to PINV coefficients due to the assumptions of linear combination. Therefore, changes in the spectral dataset such as a pixel value or removal in a HSI can alter the PINV result (**Fig. 1.9 C-F**), while the PRM result is more robust. Finally, the nature of inner product multiplication allows for the spectral dataset to be easily split into smaller subsets for parallel computing.

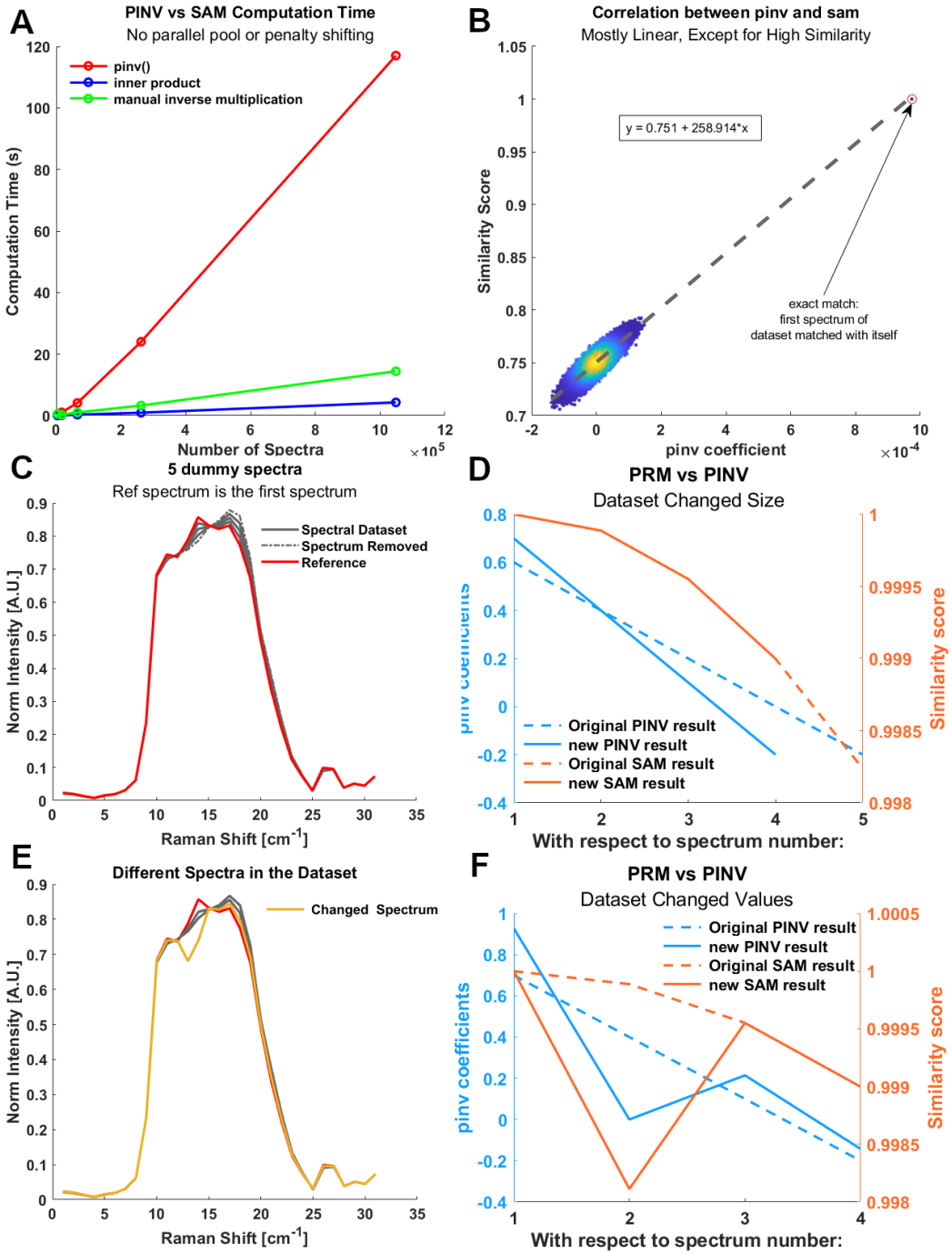


Figure 1.9 PRM vs PINV Correlation and Robustness

(A) Comparison of calculation times with the addition of manual inverse multiplication to bypass internal invertibility checks in the standard documented pinv() MATLAB function. **(B)** Density scatterplot of PINV and PRM results on a randomly generated array of spectra. **(C)** A dummy set of 5 spectra meant to mimic a Raman spectral feature. **(D)** PRM and PINV scores calculated for the dummy spectra. After removal of one of the dummy spectra in the dataset, the PINV result changes for every spectrum in the dummy set with respect to the first spectrum. **(E)** Further altering a single value in the second spectrum. **(F)** Further alterations observed in the PINV result, while the PRM result only changes with respect to the intended spectrum.

1.2.5 Optical Redox Ratio

Oxidative stress is a precursor to many disease states including fibrosis and cancer. Detecting this biomarker would enhance early diagnosis accuracy, but measuring the redox status of a sample can be expensive. Superoxide dismutase (SOD) or Glutathione (GSH) antibody labels can cost thousands of US dollars per milligram, and may preclude other imaging modalities due to the sensitive nature of these proteins. Certain tissue processing methods such as DMSO washes used in CODEX imaging would generate free radicals that would alter the native levels of these proteins. Total antioxidant capacity (TAC) assays are destructive and sacrifice spatial information. A label-free optical technique called the normalized optical redox ratio capitalizes on the endogenous NAD[P]H and Flavin autofluorescence to calculate the redox status of a sample (**Fig. 1.10A**). NAD[P]H are co-factors in cellular respiration via glycolysis and TCA cycle, and redox maintenance through the pentose phosphate pathway (PPP) and its reduction of glutathione. Flavins are a class of proteins involved in many processes, including cellular respiration and lipid peroxidation. Under oxidative stress, flavin enzymes may target poly-unsaturated lipids, which act as buffers to oxidative stress. The autofluorescence of Flavin and NADH are measured with an excitation/emission pair of 340nm/460nm and 430nm/515nm, respectively (**Fig. 1.10B**). This is done so there is minimal overlap of quantum yield between these biomarkers due to the negligible absorption of NAD[P]H at 430nm [23]. Together, they culminate in the normalized optical redox ratio described in equation 1.13 below.

$$\boldsymbol{ORR} = \frac{\boldsymbol{I_{Flavin}}}{\boldsymbol{I_{Flavin}+I_{NAD[P]H}}} \qquad \textbf{1.14}$$

In developing this optical platform, a crucial aspect of the protocol is to capture the ORR first, before the SRS and SHG modalities. This is because endogenous fluorophores such as NAD[P]H and Flavins are sensitive to photobleaching effects – the gradual and sometimes permanent alterations in fluorescence intensity. **Figure 1.10C** highlights this effect on MCF7

human HER2- breast cancer cells. SRS, however, is less sensitive to longer imaging times. Therefore, all ORR images are acquired first, prior to SHG and SRS imaging.

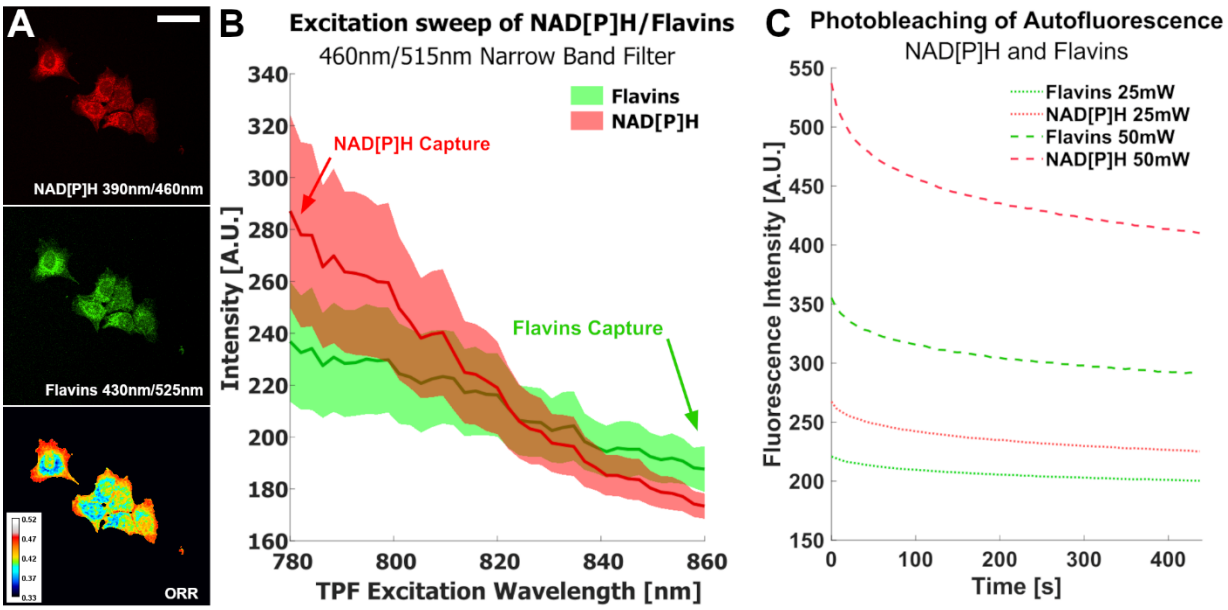


Figure 1.10 Optical Redox Ratio Platform Development
(A) MCF7 breast cancer cells in vitro imaged using 50mW two-photon laser and narrow band RNDDs. Scalebar 10 microns. **(B)** Average intensity of cells with 1-sigma error bands collected from a multichannel spectral sweep with NAD[P]H and Flavin autofluorescence captured simultaneously using the indicated two-photon excitation wavelengths. **(C)** Diminishing fluorescence intensity of a single region of interest imaged over time using 25mW and 50mW incident power. New regions of interest were used for each laser power.

A more accurate representation of oxidative stress can be developed by incorporating Fluorescence liftetime imaging (FLIM) which can distinguish between NADH and NADPH pools [24]. Modular and open-source development steps have also been developed, making this more attainable [25].

1.2.6 Collagen Anisotropy Analysis

Collagen anisotropy is a spatial metric of relative orientation. Fibrotic tissues may exhibit a more isotropic fiber orientation, as well as fiber density. Therefore, it is critical for an optical

platform to be able to capture this orientation at multiple scales. To that end, a FFT/PPSID software was created and successfully applied in the literature to diseases such as glaucoma and to characterize structures in the optic nerve [10, 26]. A more directional and ordered collagen signal will result in a higher anisotropy value. By adjusting the bin size of pixels used to generate the polar vector angle, the effect of the different amounts of regional averaging also varies, and consequently various structures can be isolated. **Figure 1.11A** below shows a slice of a 3D SHG image of human kidney tissue, and the corresponding polar vector map with colors scaled to the anisotropy value using a 16x16 pixel bin size. This capability can be extended to 3D views as well (**Fig 1.11 C-D**), showing promise in deep tissue imaging analysis. It is important to note that the main polar vector angles shown in the map representations are simply the longest vector, and are mainly used when there is a definite uniaxial orientation. This angle can also be influenced by excessive noise. Future work incorporating fast single-frame deconvolution algorithms may be instrumental for more sophisticated angle detection. This may also aid in the computation of the degree of fiber recruitment around a single preferred direction, which needs robust optimization between different tissues.

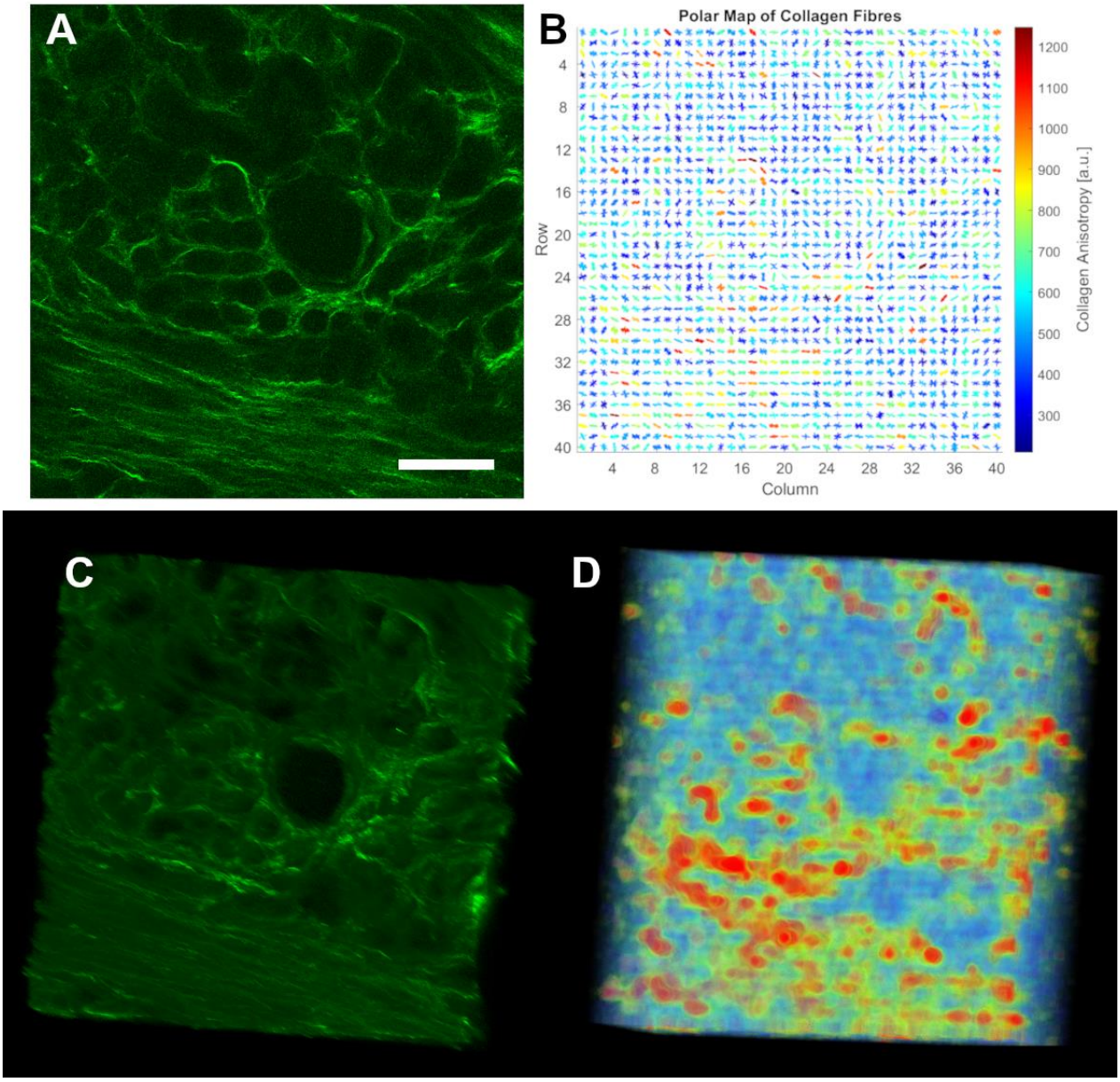


Figure 1.11 Collagen Anisotropy of 3D Human Kidney Images
(A) A 2D plane of the human kidney SHG image, showing the collagen 1-3 fibers surrounding the glomerulus and tubulointerstitial space, as well as a medullary ray. Scalebar 200 microns. **(B)** The corresponding polar vector map for (A) showing the polar vector angles in the 16x16 pixel bins. The color of the angles corresponds to the anisotropy value. **(C)** 3D view of A. **(D)** 3D interpolated view of (B) with only the bin color, and not the angle shown.

1.2.7 Collagen Fiber Thickness Analysis

Collagen fiber thickness is correlated with SHG intensity [11, 12], but can also be measured spatially. This is important because it has also been demonstrated that different types

of collagen exhibit different SHG intensity [27]. For highly disarrayed structures like in the kidney (**Fig. 1.12 A**), generating a distribution of all fiber measurements is a challenge. Several methods can be employed to address this issue. First, a sweep through various bin sizes can be applied to the SHG image to obtain the polar vector map (**Fig. 1.12 B**). From this map, a perpendicular axis can be drawn across each bin, and the intensity profile plotted (**Fig. 1.12 C**). This way, it ensures that measurements such as the full width half maximum (FWHM) distance of the SHG intensity plot is the perpendicular distance. Unfortunately, this method requires some supervision since the ultimate choice of bin size is subjective, especially in samples with diverse fiber diameters. A more direct method to estimate the fiber thickness is to transform the SHG to a distance map. To do this, the SHG image is thresholded such that there is a defined background with a zero value. Then each pixel in the image is converted to the shortest Euclidean distance between that pixel and the background (**Fig 1.12 D**). Pixels in the middle of the fiber will have higher values because they are furthest from the background. Isolating these values by finding regional maxima (**Fig 1.12 E**) provides a list of half diameters. These can be used to estimate the distribution of all fiber thicknesses (**Fig. 1.12 F**). In some instances, with highly varied distances, a morphological top-hat filter is applied to avoid distortion of the thickness distribution. This is because the regional maxima algorithm may double-count areas with high thicknesses, and neglect areas with low thicknesses.

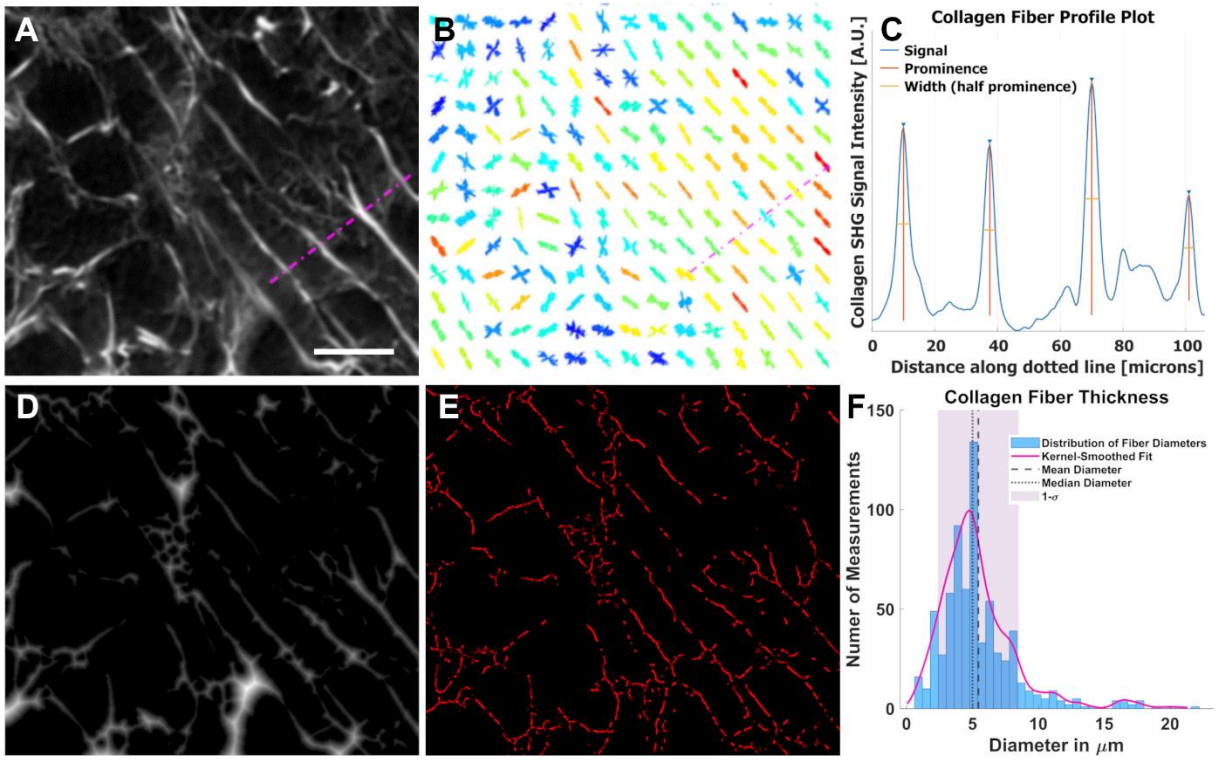


Figure 1.12 Collagen Fiber Diameter Measurements
(A) A 2D plane of a human kidney SHG image, showing the collagen 1-3 fibers. Scalebar 200 microns. **(B)** Polar vector map of the same region. Main angles are color coded from low to high anisotropy following the jet LUT in MATLAB. A dotted magenta line perpendicular to the average fiber angle of the lower right quadrant is overlaid. **(C)** The SHG intensity profile is plotted along the dotted magenta line in (A) and (B). The FWHM is calculated using the findpeaks() function in MATLAB, and shows a fiber thickness of about 5-10 microns. **(D)** The distance map transformation of (A) in which pixel intensity corresponds to the distance of SHG signal to the nearest (Euclidean) background pixel. **(E)** Local maxima were found using imregionalmax() function in MATLAB after a 1-pixel top-hat morphological filter. **(F)** A nonparametric kernel-smoothed fitted histogram with density evaluated at 100 equally spaced points throughout the range of the data shows the mean fiber thickness to be 5.4 microns, with a standard deviation of 3.1 microns. Scale bar 50 µm.

1.3 Experimental Procedures and Considerations

1.3.1 Spontaneous Raman Spectral Acquisition

A Raman spectrometer (XploRA PLUS, Horiba) attached to a confocal Raman microscope was used to measure all tissue samples. A 100x objective lens (MPLN ×100, Olympus) and a 532 nm diode line focusing laser (~40 mW) were used. Before the start of all experiments, the laser power was tested to avoid any damage to the samples. The signal was detected by a cooled

charge-coupled device (CCD) detector mounted on a 2400-slot/mm grating spectrometer. Spectra were collected at 30 s acquisitions with an accumulation of 2. Background spectra, acquired at the same focal plane for each AL region, were subtracted from the original spectra immediately after each data acquisition was completed. The phenylalanine peak at 1001 cm^{-1} was used to normalize the corresponding Amide peaks. The instrument calibration was verified by using the silicon line at 520 cm^{-1}.

1.3.2 Spectral Pre-Processing

Cosmic Ray removal is automatically applied during acquisition steps using the LabSpec6 software (HORIBA).

Background subtraction is performed by capturing a spontaneous Raman spectrum in the same focus plane as the sample, but in a region without sample and only PBS, and then subtracting the background spectrum. Rarely, small adjustments are required to ensure the background spectra is lower in intensity at all wavenumbers than the sample spectrum.

Baseline subtraction is performed using ArPLS [28], unless otherwise indicated, to strip the peaks from broad fluorescence effects. To ensure accurate baseline estimation, spectra were first linearly interpolated at every wavenumber.

Total intensity normalization (AUC) where the integral of the function is calculated using trapezoidal numerical integration. Then each raw intensity value is divided by the integral value calculated, giving a new set of values that present the relative proportion of the intensity at each data point compared to the overall intensity of the spectrum. These values are then transformed to be on a common scale, 0 to 1, between all spectra by subtracting the minimum value and dividing by the range.

Vector normalization is performed where indicated, such as for PRM-SRS, using the standard vecnorm function in MATLAB.

1.3.3 Stimulated Raman Scattering Imaging Acquisition

An upright laser-scanning microscope (DIY multiphoton, Olympus) with a 25x water objective (XLPLN, WMP2, 1.05 NA, Olympus) was applied for near-IR throughput. Synchronized pulsed pump beam (tunable 720–990 nm wavelength, 5–6 ps pulse width, and 80 MHz repetition rate) and Stokes (wavelength at 1032nm, 6 ps pulse width, and 80MHz repetition rate) were supplied by a picoEmerald system (Applied Physics & Electronics) and coupled into the microscope. The pump and Stokes beams were collected in transmission by a high NA oil condenser (1.4 NA). A high O.D. shortpass filter (950nm, Thorlabs) was used that would completely block the Stokes beam and transmit the pump beam only onto a Si photodiode for detecting the stimulated Raman loss signal. The output current from the Si PIN photodiode was terminated, filtered, and demodulated in X with a zero phaseshift by a lock-in amplifier (HF2LI, Zurich Instruments) at 20MHz. The demodulated signal was digitized and fed into the FV3000 software module FV-OSR (Olympus) to form the image during laser scanning. All 3D lipid droplet images were obtained with a pixel dwell time 40 µs with 3-frame averaging for a total imaging speed of ~10-15 min per image stack. Laser power incident on the sample is approximately 40mW.

1.3.4 Lipid Subtype Measurements

Lipid standards such as cardiolipin 14:0 (63988-20-5, Sigma), TAG 18:1 (122-32-7, Sigma), and Ceramide 18:1;2/17:0 (67492-16-4, Sigma) were dissolved in 1:1 Chloroform and Methanol at indicated proportions and then allowed to dry. The precipitate was analyzed using the Spontaneous Raman Spectrometer (XploRA plus, HORIBA) with a 40mW 532 nm laser, using 1800 slits/mm grating, an acquisition time of 15s, and accumulation of 2. Spectra were interpolated at every wavenumber and baseline corrected using ArPLS. GC/MS lipid measurement is described in Chapter 3.4. Measured reference spectra are shown in figure 1.13 and figure 1.14 below.

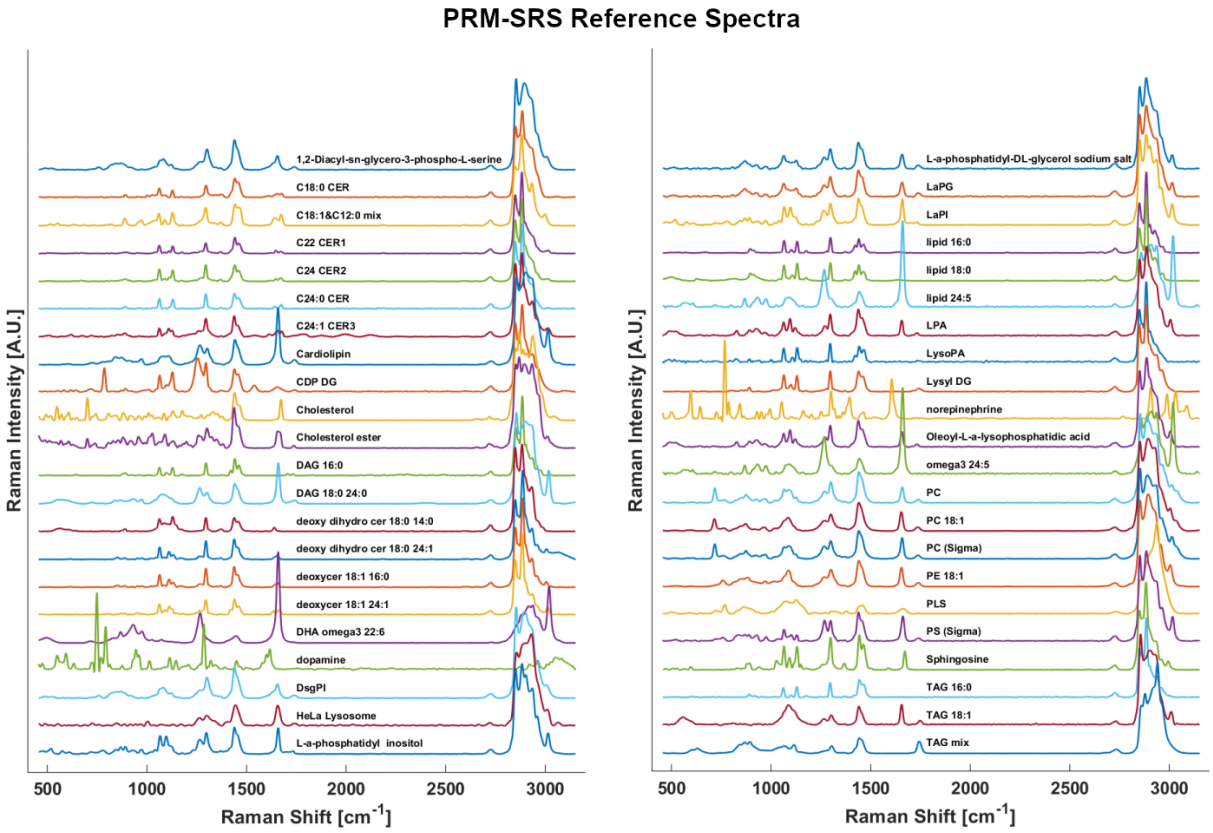


Figure 1.13 Reference Spectra
List of reference spectra acquired for use in PRM-SRS. Spectra are baseline corrected according to section 1.3.2 and normalized from 0 to 1.

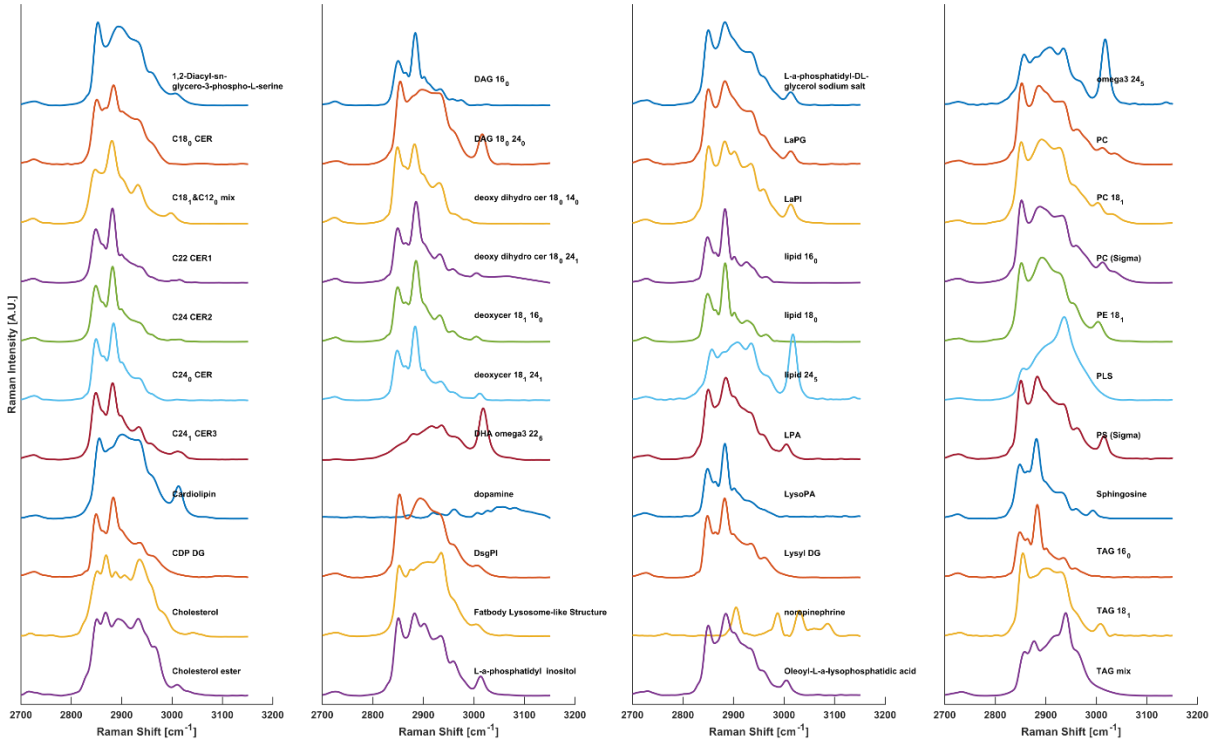


Figure 1.14 Reference spectra enlarged CH stretching region
CH stretching region from Fig.1.9 is enlarged to show more detail.

1.3.5 Characterizing Noise in PRM-SRS

Noise is a concern for SRS imaging because it can affect the ability to identify the molecular signature of the analyte and in turn influence the results of PRM-SRS and other methods. An example of a calibration standard is shown below in figure 1.15. Pure ethanol was imaged transmissively with the SRS microscope. After conducting PRM-SRS on the HSI, the distribution of the similarity score results is plotted (**Fig.1.15 B**), showing extremely high similarity scores with a negligible variance between pixels.

Another related concept is bit depth. To fully utilize the 16-bit depth output of our SRS signal, the scale and offset parameters used when transforming and digitizing the signal need to be considered. Figure 1.16 below shows the relationship between these parameters and the image contrast. Since PRM-SRS is a relative estimation of concentration, ratiometric image processing is frequently used. If the contrast is not great enough, the ratiometric result will suffer

unless an even greater bit depth is used, which can quickly exhaust computer memory. This is because if an image of certain bit depth is unable to assign a value to enough decimal places, then it is rounded, potentially making it indistinguishable from another pixel which would otherwise have a very different result. Adjusting the scale through modulating the gain and resistance of the incoming electrical signal from the photodiode also enhances noise. However, we have found that using simple frame averaging is enough to combat this issue.

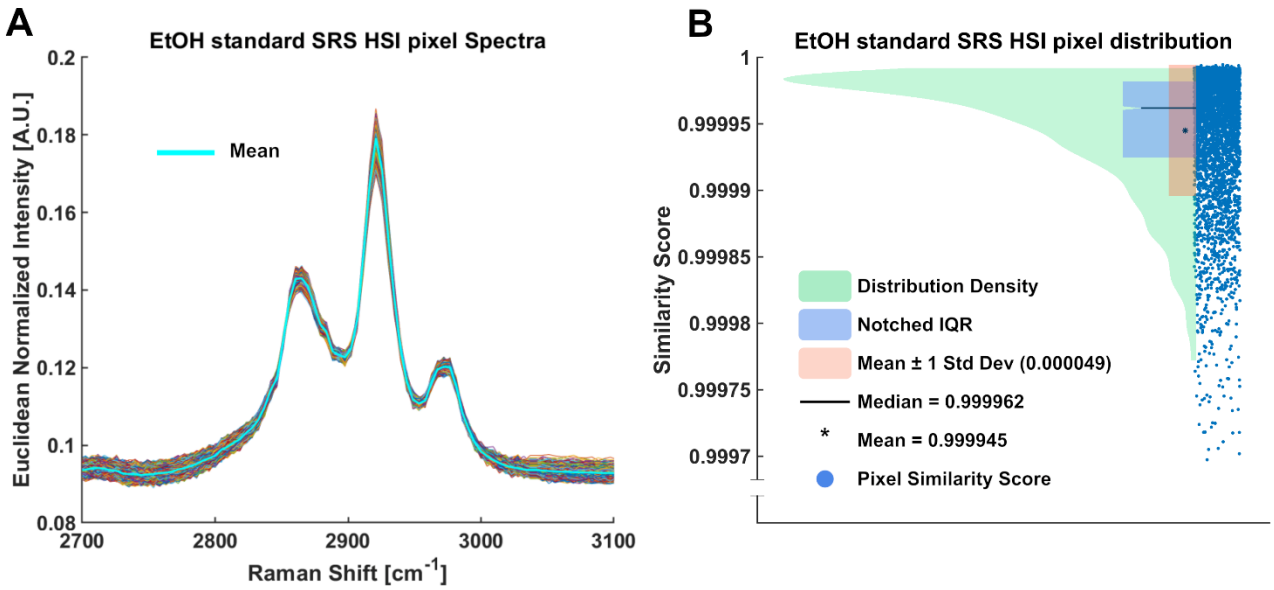


Figure 1.15 Low SRS HSI Pixel Spectral Variance
(A) All pixel spectra in a 64x64 test SRS HSI of pure ethanol shows very little pixel-wise "noise". All pixels in the image have a relatively consistent spectral profile, as expected. **(B)** This homogeneity is translated into similarity scores, and confirms that the image quality of the SRS HSI is sufficient, considering the vast majority of pixels all have nearly perfect similarity scores with respect to a spontaneous Raman spectrum of the ethanol standard.

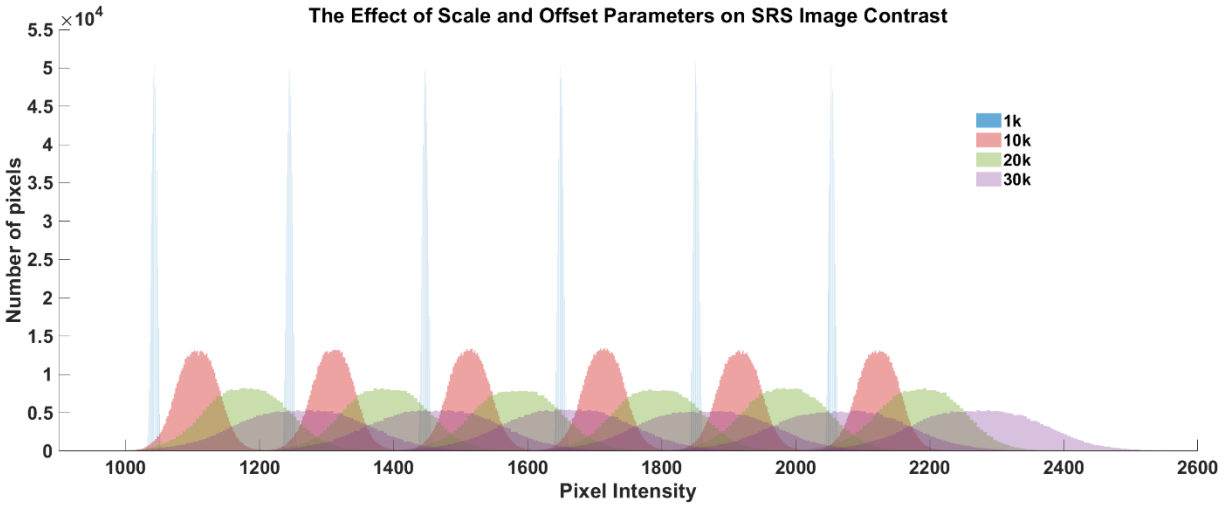


Figure 1.16 The effect of scale and offset parameters on SRS image contrast
Four different scales: 1k, 10k, 20k, and 30k, and 6 different offsets (0, -0.1, -0.2, -0.3, -0.4, -0.5) were used. Offset linearly shifts the entire pixel intensity distribution of an image, while the scale parameter has a multiplicative effect.

1.3.6 Heavy Water Cell Culture

Normal-like breast epithelial cell line (MCF10A) were cultured in Dulbecco's modified Eagles' medium (DMEM), supplemented with 5% heat-inactivated fetal bovine serum (FBS) (Hyclone) and 1% penicillin/streptomycin (Fisher Scientific, Waltham, MA), and incubated with 5% CO_2 at 37°C. Cell cycles were synchronized using double thymidine block [29]. After passaging at 80% confluence, cells were seeded at a concentration of 2×10^5/mL atop 70% ethanol-soaked cover glass in 24-well plates and incubated for 8 hours. Then the growth media was changed to 50% heavy water (D_2O) and treatment media as follows. Cells were incubated for 48 hours, which corresponds to a deuterium-retarded cell cycle. Cyclin dependent kinase 1 (CDK1) inhibitor (RO-3066, Sigma) was added with 8 hours remaining to arrest growth before mitosis. Cells were gently rinsed with 1x PBS with Calcium and Magnesium ions at 4°C (Fisher Scientific, 14040216), and finally fixed in 4% methanol-free PFA solution (VWR, 15713-S) for 15 minutes. The cover glass was mounted on 1mm thick glass microscope slides with 120 µm spacers filled with 1x PBS without calcium and magnesium ions. These samples are stored at 4°C submerged in PBS when not in use.

1.3.7 Drosophila Tissue Dissection and Preparation

The Drosophila lines incorporated into this research were obtained from the Bloomington Drosophila Stock Center (BDSC), barring any exceptions otherwise noted. For successive generations, these lines have been nurtured on a standard diet consisting of Nutri-Fly medium (Catalog Number: 66-113, provided by Genesee Scientific Corporation). The wild-type model deployed was w1118 (designated as stock #5905). The expression of UAS-APP (#33797) was driven by the mifepristone-inducible expression pan-neuronally ElavGS (#43642).

Dissection of the drosophila central brain was performed in phosphate buffered saline (PBS), followed by fixation using a 4% formaldehyde solution for a duration of 30 minutes at ambient temperature. Subsequent to fixation, tissues were subjected to four washes in PBS. The

dissected tissue samples were placed on slides along with in PBS, then sealed under number 1 thickness cover glass to prevent the sample from drying out.

1.3.8 Human Kidney Sample Preparation

Human kidney samples were obtained from The Kidney Translational Research Center (KTRC) at the Washington University School of Medicine in St. Louis under an approved IRB protocol. Samples were fixed using 4% PFA and stored in 1x PBS at -20deg C when not in use. Samples were sliced using a sliding microtome (HM 450, Epredia) at 200 µm and cleared using 8M urea for 48 hours at room temperature. Samples were submerged in 1ug/mL Hoechst 33342 (Thermofisher) for 15 minutes to stain nuclei for co-localization verification. Samples were imaged between 1mm thick glass slides and number 1.5 thickness cover glass (Erie Scientific). A tunable 780nm-810nm pump laser with 50mW incident power was modulated with a 1031nm stokes laser of 30mW incident power (PicoEmerald, APE) to illuminate the sample.

For FFPE samples, a deparaffinization step was conducted to remove strong CH stretching contamination. Briefly, samples were warmed to 55degC for 5 min using a slide warmer (Cole-Palmer) and submerged in pure xylene for 10 min with gentle rocking. Slides then underwent sequential washes in graded concentrations of ethanol in xylene from 20% to 100% and then rehydrated in decreasing concentrations of ethanol in water from 100% to 20%. OCT embedded samples were soaked in 1x PBS without calcium and magnesium ions overnight with gentle rocking, and quickly rinsed in dH_2O for 1 min. Samples are stored in PBS without calcium and magnesium ions at 4degC when not in use.

Relevant tissue preparation reagents are also measured (Figure 1.17 below)

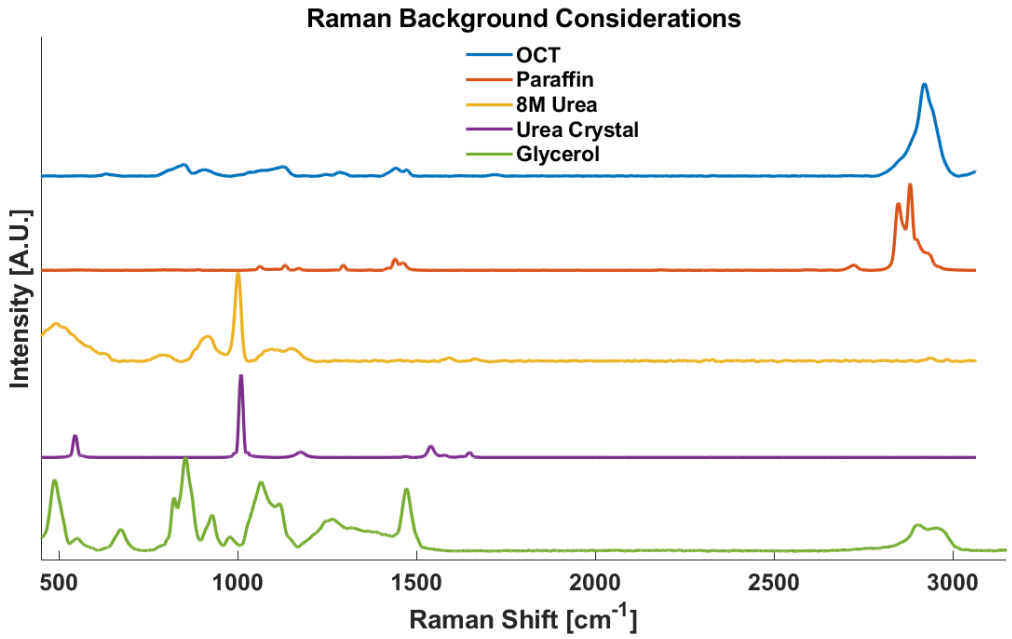


Figure 1.17 Raman spectra of common tissue sample preparation reagents

1.4 Dissertation Aims

In driving the adoption of any platform, two main approaches exist: technological developments and novel applications. This dissertation aims to combine various label-free optical imaging techniques into a cohesive and powerful platform contained in a single microscopy setup, and determine the extent to which it is able to reveal novel insights at both the cellular and tissue levels. To this end, development and fine-tuning of methodologies, as well as their applications to novel biomedical questions were conducted. Specifically:

- **Chapter 2:** Triple negative breast cancer (TNBC) subtypes were analyzed in 1D (spontaneous Raman spectroscopy), 2D, and 3D *in situ* with heavy-water under various levels of insulin and methionine concentration. Lipid droplet (LD) spectra reveal that unlike normal breast epithelial cells, TNBC express significant interdependence of insulin and methionine in tandem. Biochemically, excess insulin and methionine up-regulate *de novo* lipogenesis, particularly of saturated fatty acids. Morphologically, the newly synthesized

LDs appear larger and more compartmentalized away from Flavin signal and mitochondria.

- **Chapter 3:** Going beyond lipid saturation, this dissertation aims to characterize the relative levels of any lipid in a sample by leveraging SRS HIS. To do this, PRM-SRS was applied to multiple cellular and tissue samples and shows excellent correlation with fluorescence and GC/MS results. Results indicate that this method is much faster than any label-free spectral method published to date, cross compatible with various Raman instruments, and preserves the ability to later perform GC/MS, IF staining, and histological staining after being imaged with this platform.
- **Chapter 4:** Combining the 2D and 3D multimodal methods employed in Chapter 1 with the hyperspectral analysis methods employed in Chapter 2, this platform was then applied to *ex vivo* human kidney biopsies. Multiple biomarkers such as mesangial expansion, collagen fiber thickening, oxidative stress, and lipid saturation were revealed, all without the use of labels.

In summary several novel experiments have been conducted at multiple levels, from subcellular organelles to entire tissue units, and at multiple dimensions, from 1D to 3D. Throughout this dissertation, an emphasis is placed on the unique ability to visualize both structural and biomolecular phenotypes and produce novel insights and reveal biomarkers without the need for any labeling.

Chapter 2: Single Cell Analysis of Tandem Nutrient Control in Human Breast Cancer

2.1 Introduction

Breast cancer is the most reported form of cancer in biological women, but its pathophysiology is rife with subtypes that have material consequences on clinical outcomes. Triple negative breast cancer (TNBC) is a particularly aggressive cancer subtype that accounts for approximately 15% of all breast cancer cases and its epidemiology reveals a discriminating predilection for non-Hispanic African women [30, 31] (**Supplementary Fig. 2.11**). Although the genomes and proteomes of these breast cancer subtypes are distinguishable, little is known about their metabolic phenotypes and the consequential prognoses they manifest.

Recently, lipid metabolism has emerged as a major indicator of cellular stress, phenotypic state, and disease status in biological research and medicine. Dysregulation of lipid metabolism and heightened lipid synthesis are hallmarks of cancer, as varying demands of lipids for energy maintenance, metastasis, and angiogenesis warrant transcriptional changes that contribute to the metabolic phenotype [32–34]. The quantity and diversity of lipids and their functions have been instrumental in profiling cancers as well. For example, membrane lipid compositions of cholesterol, phosphatidylcholine (PC), and phosphatidylethanolamine (PE) are essential to cell membrane fluidity, which has become a target for cancer treatments [35–38]. Additionally, the degree of saturation of lipid content in a cell may provide further insight into its state of stress, as breast cancer cells may produce more saturated and monounsaturated membrane lipids to guard against oxidative stress [39–41]. To interrogate lipid metabolism, lipid droplets (LD) were the primary organelle of interest since their ubiquitous structures not only serve as energy stores, but are also involved in protein folding and trafficking, signaling pathways, and have diverse spatial and chemical information that may reflect oxidative stress, metabolic flux, and disease status [39–47]. However, direct visualization of LD metabolism manipulated by tandem nutritional

interventions at a subcellular level has not yet been reported in TNBC cells, which is partially due to a lack of spatial resolution in conventional lipidomic modalities. Optical techniques such as spontaneous Raman spectroscopy and SRS imaging microscopy are well suited to both the chemometric and spatial dimensions for imaging LD metabolism; they can analyze not only the size, number, and distribution of LDs, but also their protein and lipid diversity and metabolism at subcellular resolution.

Despite the many mysteries of TNBC, a documented hallmark is its hyperactivity of mammalian target of Rapamycin (mTOR) pathways, which play important roles in glucose, protein, and lipid metabolism [48–52]. Insulin and L-methionine (an essential amino acid involved in protein translation, genetic/epigenetic control, nutrient sensing, and redox homeostasis) [53] are both involved in mTOR pathways but have not been directly studied in tandem to date [54–59]. This is due, in part, to previous studies that observed MDA-MB-231 cells to be insulin insensitive to mitogenic effects, despite having many receptors that bind insulin [60]. Other studies observe insulin effects in the same cell line, and there is currently no consensus on the independent effects of insulin. With respect to TNBC, insulin and methionine both independently drive cancer proliferation [61–64] and affect lipid metabolism [54, 63, 65–69], and separate studies indicate insulin metabolism directly affects the uptake of amino acids in yeast and dogs [70, 71]. Given the well-documented relationships between insulin, methionine, and mTOR, it is possible that TNBC's mTOR hyperactivity exhibits a unique lipidomic response to insulin and methionine manipulation. The conceptual pathway detailing macroscopic mTOR-mediated lipid response to insulin and methionine (**Fig. 2.1**) highlights the points discussed in this paper. Lipid peroxidation, de novo synthesis, and chemical diversity can all be investigated using optical techniques that provide subcellular spatial and chemical information. Given that TNBC has been an archetype for methionine dependence [64], and that PI3K/AKT/mTOR is a key driver of the aggressive biology of TNBC [52], the interplay between methionine and insulin, coupled with the perspective of lipid biology, may illuminate promising directions for future therapeutic research.

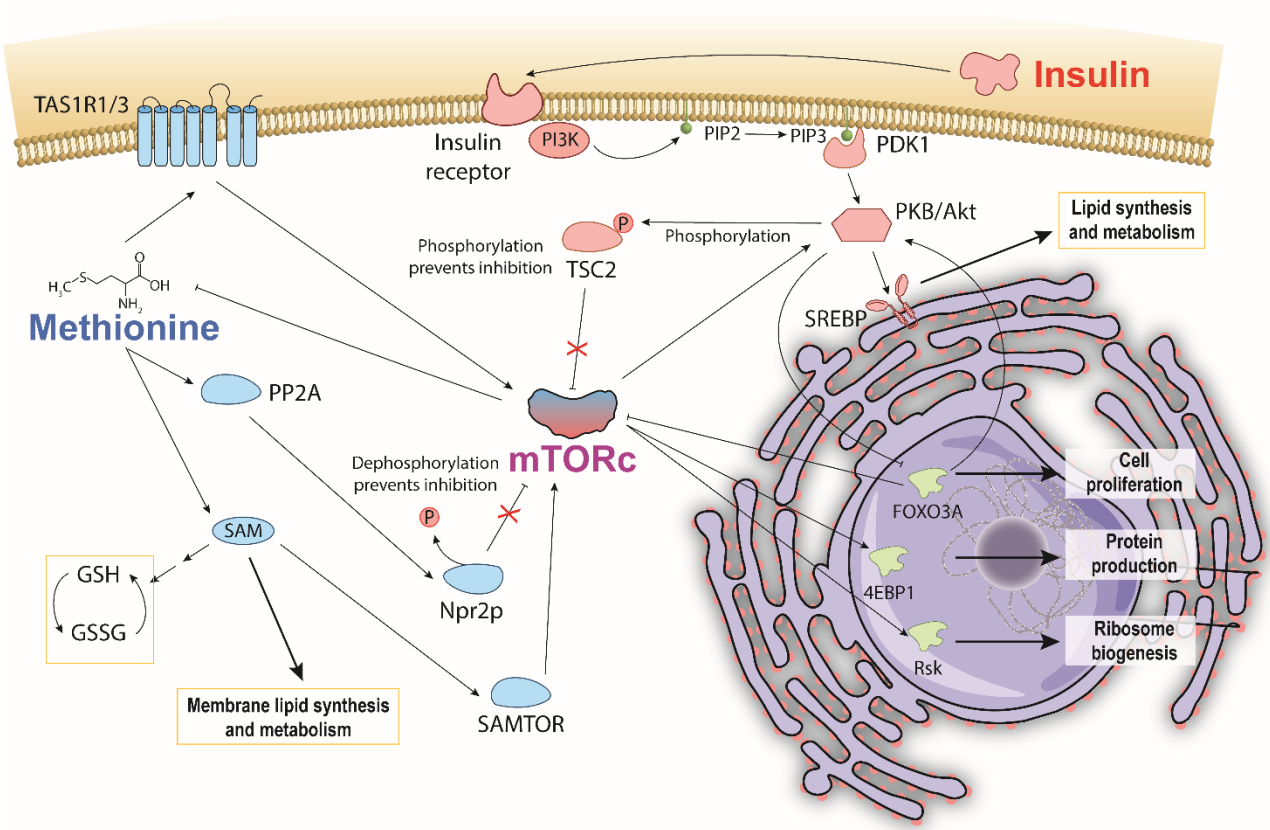


Figure 2.1 Hypothesized Insulin and Methionine Pathway Interaction Mediated Through mTOR
Hypothesized pathway illustrating a potential methionine and insulin interaction mediated through mTOR. Bi-directional control of methionine and mTORc1 depicts general mechanisms by which methionine is sensed by and activates mTORc1. Insulin also activates mTORc1 by phosphorylating TSC2, and consequently affecting mTORc1 regulation of methionine. Insulin stimulates SREBP mediated lipid synthesis and metabolism. Methionine stimulates SAM PC and PE membrane lipid synthesis. Increased production of reduced glutathione via SAM is thought to reduce the extent of lipid peroxidation.

Non-linear optical techniques such as coherent Raman scattering microscopy and two-photon fluorescence (TPF) microscopy have been used to profile breast cancer metabolism by revealing correlations between cancer metastasis and cellular redox state, and lipid metabolism [72]. Recent studies have identified several metabolites implicated in tumorigenesis and lipid metabolism in cancer, such as glutamine [73–75] and serine [76] dependence. Raman spectroscopy/microscopy coupled with D_2O probing allows for direct visualization of metabolic dynamics of a variety of biomolecules including lipids, protein, and DNA in cells, C. elegans, zebrafish, and rodents by highlighting the newly synthesized macromolecules [17]. In this study we first employed spontaneous Raman spectroscopy to differentiate molecular signatures within

LDs between TNBC and normal cells. Using D_2O probing and SRS (DO-SRS) imaging we then examined the impacts of methionine and insulin on lipid metabolism in cancer cells. The effects of methionine and insulin on cellular respiration and lipid peroxidation were also examined by using TPF microscopy. To analyze the rich chemometric dataset and inspire targeted image analyses, we applied a relative entropy approach to Raman spectra for the first time. This method can quickly highlight distinct or tandem effects of independent variables in any Raman spectroscopy study.

2.2 Results

2.2.1 Lipid Droplet Metabolism

We first examined the effects of methionine on LD metabolism in TNBC cells (MDA-MB-231), luminal A breast cancer cells (MCF-7), and normal breast epithelial cells (MCF10A, as a control) by adding excess (20x) methionine to the growth media supplemented with 50% D_2O. Cells were scanned by using a spontaneous Raman spectroscopy, and revealed that TNBC cells most starkly contrasted MCF10A cells with respect to overall lipid content (CH_2 stretching at 2850cm^{-1}). This attenuated lipid:protein ratio difference between excess and physiological methionine concentrations is shown in **Fig. 2.2 A**. This absence of marked differences is also personified by poorer ReLu neural network classification between TNBC cells with and without excess methionine (**Supplementary Fig. 2.12, 2.13**). Despite the absence of insulin in TNBC cell culture growth media recipes [77], we then added various concentrations of insulin (1mg/L, 10mg/L, and 20mg/L, correspondingly, 0.1x, 1x and 2x) to the media and evaluated its interaction with methionine in both cell lines. In this second part, insulin concentration in growth media was modulated in tandem with methionine, and augmented effects in several Raman spectral regions were observed, including the C-H stretching region, which illustrates the relative contents of CH_2 (lipid) and CH_3 (protein) (**Fig. 2.2 B**). TNBC contrasted MCF10A cells which exhibited decreased

lipid:protein ratios in the presence of excess methionine at all insulin concentrations. Importantly, it was found that the difference in lipid:protein ratio between excess and physiological methionine increased with the addition of insulin in TNBC. **Fig. 2.2 C** highlights this effect, marked by orange arrows in **Fig. 2.2 B**, and supports potential interactions between insulin and methionine. A significant interaction term was confirmed by 2-way ANOVA (**Supplementary Table 2.1**) in TNBC.

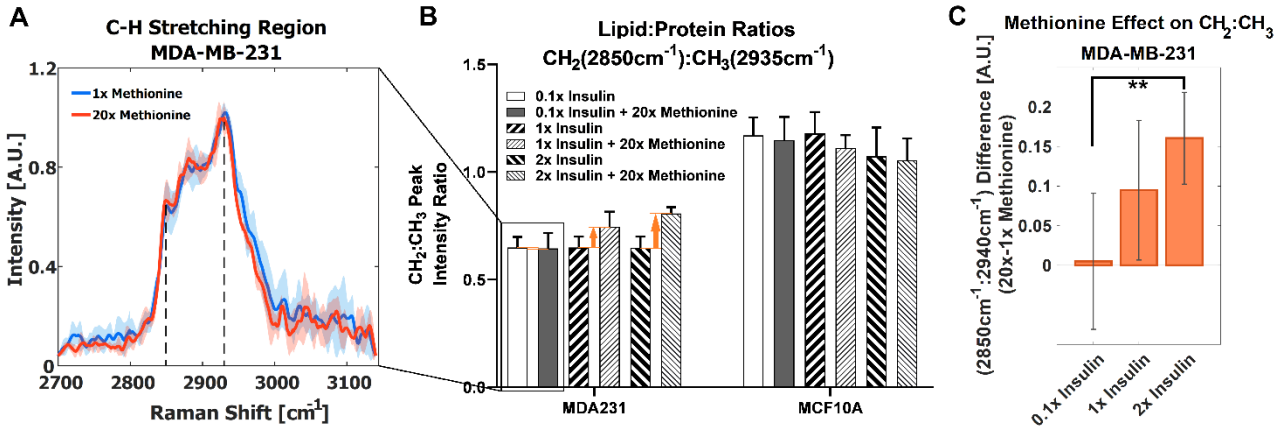
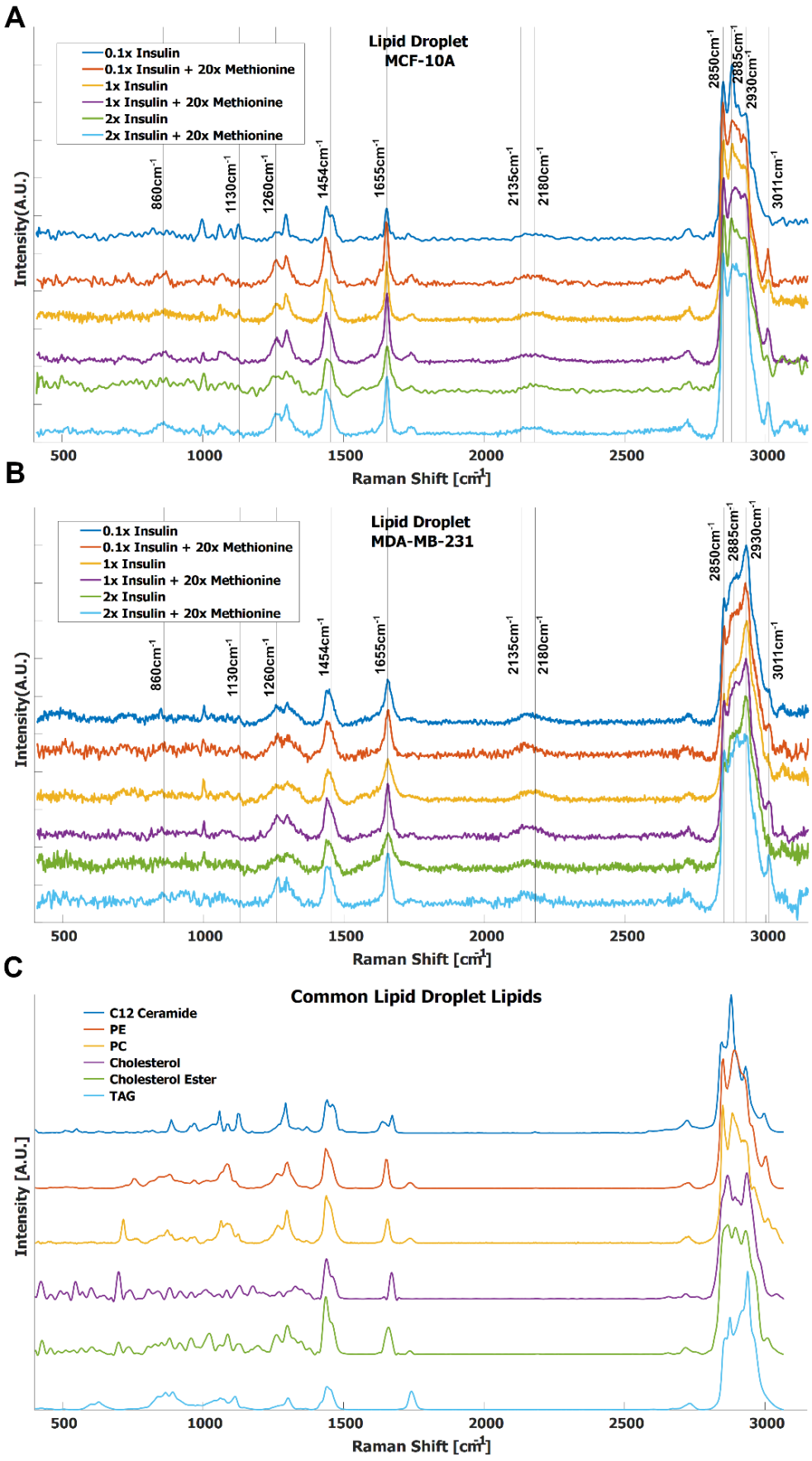


Figure 2.2 Cell type and Condition Dependent Responses to Excess Methionine
(A) Average CH stretching region spectra for TNBC with highlighted CH_2 and CH_3 ($2850cm^{-1}$ and $2935cm^{-1}$) levels ascribed to total lipid and protein content, respectively. Regular and excess methionine groups refer to 0.03g/L and 0.6g/L respectively. Of note, TNBC did not exhibit significant relative lipid and protein changes in the presence of excess methionine. One standard deviation is indicated by shaded areas surrounding the lines. **(B)** With tandem insulin control CH_2:CH_3 peak ratios at each insulin and methionine concentration group are shown. 2-way balanced ANOVA results for TNBC cells highlights significance of methionine and insulin-methionine interaction term in lipid:protein ratios. Error bars indicate one standard deviation. **(C)** The difference in CH_2:CH_3 ratios for the 15x methionine and 1x methionine groups of the MDA-MB-231 subtype is negligible at the lower insulin concentration but is increased ten-fold in the 2x insulin case. Error bars are propagated in quadrature from (B).

In TNBC cells, the ratio of total lipid-to-protein did not change with the increase of methionine concentration alone, but slightly decreased in normal cells (**Fig. 2.2 B**). With the addition of insulin, this difference was augmented with higher concentrations of insulin (**Fig. 2.2 C**). At this point, it is still unclear whether de novo lipid synthesis increased alone, or if protein synthesis decreased, or some combination of both. Perhaps de novo lipid synthesis decreased, but not as much as protein synthesis. This clarity entails DO-SRS, which provides insight into de novo synthesis. As cells incorporate deuterium from heavy water into macromolecules such as

lipids and proteins, the C-D bonds in the newly synthesized molecules become visible in the cell silent region around 2150 cm^{-1}. Even though lipids and proteins are the main biomolecular constituents of cells, the CH_2 and CH_3 peaks may only paint part of the picture. **Fig. 2.3 A, B** shows average Raman spectra of both cell lines treated with D_2O and different concentrations of methionine and insulin. These spectra are consistent with previous LD studies using Raman spectroscopy, which display minute protein peaks in the fingerprint region such as the phenylalanine peak at 1000cm^{-1} and amide I-III peaks at 1660cm^{-1}, 1450cm^{-1}, and 1200-1300cm^{-1}, respectively, as well as elevated CH_2 stretch at 2850cm^{-1}, saturated CH_2 stretch at 2880cm^{-1} (typical of cholesterol and other saturated lipids) [78], and H-C= stretch at 3010cm^{-1} (typical of unsaturated lipids) [79]. Common lipid components of LD are shown in **Fig. 2.3 C** in descending order of prevalence. The structure of LDs is such that a phospholipid monolayer surrounds a core of neutral lipids such as cholesterols and TAGs. Less prevalent lipid species such as ceramides, sphingolipids, and their metabolites only account for a small percentage of LD composition, but have gained increasing significance in LD physiology and diseases [80]. Furthermore, there are hundreds of apo-lipoproteins on or near the surface of LDs, which may contribute to the observed Raman spectra of LDs. The presence of the C-D peak in the spectra confirm de novo synthesis. Some Raman shifts of interest are shown, but minute differences may be difficult to discern by raw visual inspection alone.

Figure 2.3 Average LD Spectra of Treated MCF10A and MDA-MB-231
(A) MCF-10A and **(B)** MDA-MB-231 cells under various methionine and insulin concentrations. Manual identification of potential Raman peak targets is highlighted with vertical lines and labels. For example, the CD_{Lipid} peak in the cell silent region at 2135cm^{-1} has a noticeable increase relative to the $CD_{Protein}$ peak at 2180cm^{-1} in the excess methionine groups. This could indicate preference for de novo lipogenesis in excess methionine environments. **(C)** Raman spectra of common lipid species in LD, in descending order of prevalence.


A
Lipid Droplet
MCF-10A
0.1x Insulin
0.1x Insulin + 20x Methionine
1x Insulin
1x Insulin + 20x Methionine
2x Insulin
2x Insulin + 20x Methionine
860cm-1
1130cm-1
1260cm-1
1454cm-1
1655cm-1
2135cm-1
2180cm-1
2850cm-1
2885cm-1
2930cm-1
3011cm-1
Intensity(A.U.)
500
1000
1500
2000
2500
3000
Raman Shift [cm-1]
B
Lipid Droplet
MDA-MB-231
0.1x Insulin
0.1x Insulin + 20x Methionine
1x Insulin
1x Insulin + 20x Methionine
2x Insulin
2x Insulin + 20x Methionine
860cm-1
1130cm-1
1260cm-1
1454cm-1
1655cm-1
2135cm-1
2180cm-1
2850cm-1
2885cm-1
2930cm-1
3011cm-1
Intensity(A.U.)
500
1000
1500
2000
2500
3000
Raman Shift [cm-1]
C
Common Lipid Droplet Lipids
C12 Ceramide
PE
PC
Cholesterol
Cholesterol Ester
TAG
Intensity [A.U.]
500
1000
1500
2000
2500
3000
Raman Shift [cm-1]

Although the delineated Raman shifts in **Fig. 2.3 A, B** highlight several aspects of lipid and protein metabolism, there are others ascribed to lipids and other important molecules as well. Principal component analysis (PCA) shows that 12 principal components (PCs) account for nearly all the variance in the 6 groups of MDA-MB-231 LD spectra. To visualize this while avoiding over-fitting, a t-SNE diagram of the top 10 PCs is shown in **Fig. 2.4 A, B**. There exists at least one dimension that discriminates insulin effects and methionine effects on Raman spectra of TNBC LD. Importantly, this demonstrates that LD, alone, contain sufficient chemometric data to discriminate cell phenotypes. This confirms the ability of LDs to reflect cellular state. To date, label-free chemometric demonstrations of this ability are sparse. PCA initialization can be a robust step to reduce dimensions, denoise data, and preserve global structure in t-SNE visualizations, but even though PCA can vectorize these values, the PCs themselves do not take the form of Raman peaks suitable for direct assignment of methionine and insulin effects individually.

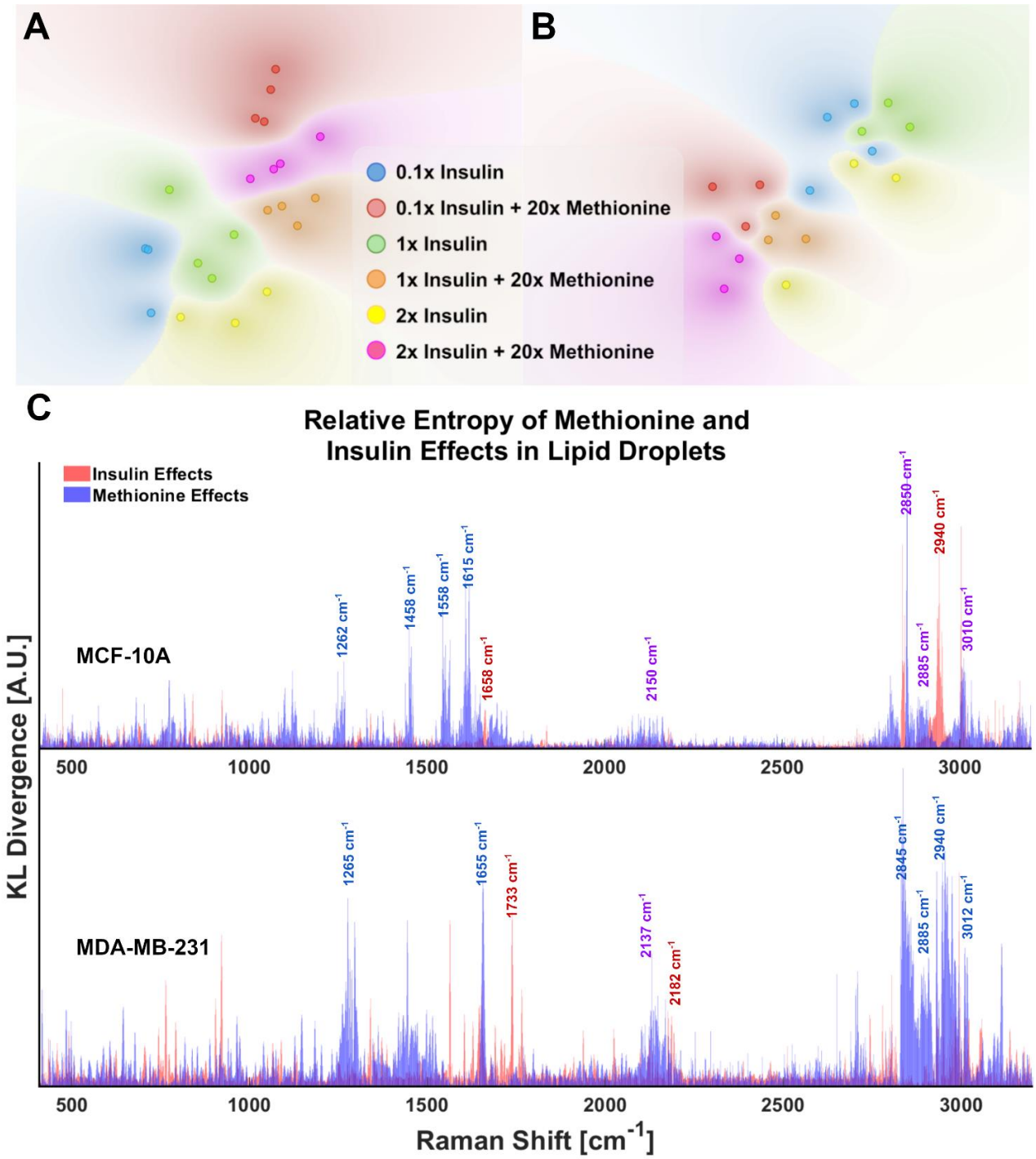


Figure 2.4 LD Spectra Discriminates Cell Phenotype with t-SNE and Relative Entropy
(A) tSNE separates experimental groups for MCF-10A and **(B)** MDA-MB-231 tSNE plots of the top 10 PCs from PCA of LD spectra. Global structure is preserved, and no exaggeration was applied. Each point represents the average of 5 LD spectra taken from a single cell of the corresponding sample. **(C)** Relative entropy provides a metric for ranking features (Raman peaks) by their ability to classify the spectra as belonging to 20x methionine or 1x methionine groups, as well as 0.1x, 1x, 2x insulin groups. Raman peaks that appear to be influential in both classification schema are denoted in purple text labels for clarity. Note: These cell subtype plots are aligned manually for clarity. Subplots are not generated in such a fashion automatically.

Statistical quantification of independent variable effects at every Raman shift entails a new measure in which the separation of insulin and methionine effects, as well as relative significance in class attribution is shown. To quickly rank and visualize all the wavenumber variables that may have been influenced by a particular treatment, the Kullback-Leibler divergence (DKL), a metric for the distance between two distributions for classification problems [14], at each Raman shift is plotted for each metabolite manipulation (**Fig. 2.4 C**). This method is also known as relative entropy.

KL divergences of Raman spectra were plotted on the same axes for MCF10A and MDA-MB-231 with selected wavenumbers labeled for clarity (**Fig. 2.4 C**). From **Fig. 2.4 C**, it is apparent the lipid peak of MCF10A cells at 2850cm^{-1} was heavily influenced by both insulin and methionine concentrations, while the protein peak at 2935cm^{-1} seems to be more heavily influenced by insulin concentration. This contrasts with the MDA-MB-231 TNBC cells in which divergences at most wavenumbers were dominated by the delineation of methionine concentration. Although these representations are not perfect, this is especially useful when simultaneous treatment groups have both compound and independent effects. For example, MCF10A spectra (**Fig. 2.4 A**) exhibit changes in the unsaturated lipid peak (3010cm^{-1}) under either insulin or methionine manipulation, while the TNBC spectra (**Fig. 2.4 B**) exhibit changes here (3010cm^{-1}) mainly under methionine manipulation. This can be easily seen though the relative entropy at that Raman shift in **Fig. 2.4 C**, in which MCF10A has high relative entropy at 3010cm^{-1} when examined along either the insulin or methionine dimension, while TNBC shows a higher relative entropy when examined along the methionine dimension.

While excess methionine appears to decrease the lipid-to-protein ratio in MCF10A cells and increase the ratio in MDA-MB-231 cells, the results do not necessarily indicate discrepant rates of de novo lipogenesis since these values are affected by both synthesis and degradation of lipid and protein. For instance, the decreased lipid-to-protein ratio might be due to enhanced

lipid utilization. To explore how much lipid and protein were synthesized, we quantitatively examined the carbon-deuterium (CD) peaks at 2135cm^{-1} (de novo synthesized lipids, CD_L) relative to 2180 cm^{-1} (de novo synthesized proteins, CD_P), and 2850cm^{-1} (total lipids, CH_2) for each treatment group [81] (**Fig. 2.5 A, B**). **Fig. 2.5 C** shows that excess methionine stimulates de novo lipogenesis in TNBC. Together, **Fig. 2.5** illustrates both direct and relative de novo lipid and protein synthesis and metabolism, and informs the potential reasons for the discrepant lipid:protein effects of excess methionine on TNBC and normal-like breast cells.

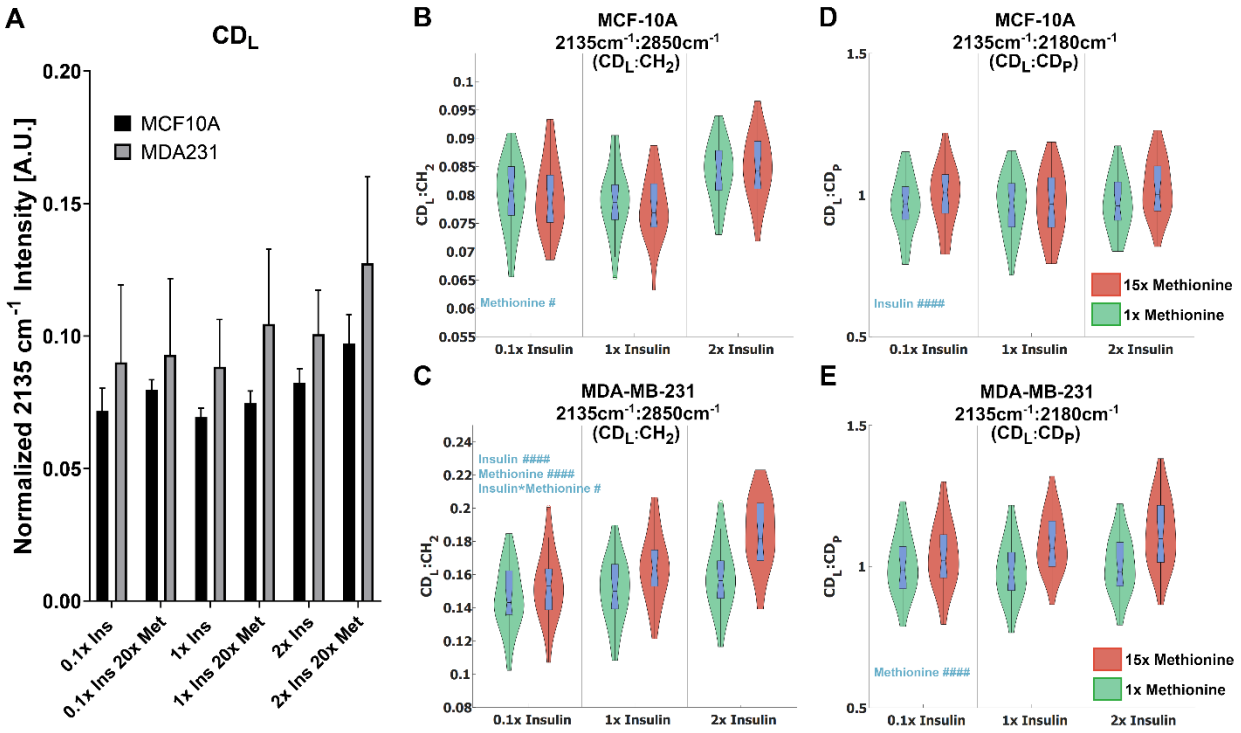


Figure 2.5 Quantitative *de novo* lipid synthesis under methionine and insulin control
(A) Normalized CD_L intensities show excess methionine stimulates *de novo* lipogenesis. **(B, C)** CD_L Ratios show violin box-plots of *de novo* lipid synthesis CD_L ratios for MCF10A and MDA-MB-231, respectively. CD_L : CH_2 illustrates the relative *de novo* lipid synthesized to total ascribable lipid content. Balanced 2-way ANOVA with constrained sum of squares results of CD_L ratios shows methionine concentration significantly influenced the CD_L: CD_P ratio in both MCF10A and MDA-MB-231 lipid droplet spectra with rejection levels of $^{\#}P < 0.05$ and $^{\#\#\#\#}P < 0.0001$, respectively. **(D, E)** CD_L: CD_P illustrates the relative *de novo* lipid and protein synthesized biomolecules for MCF10A and MDA-MB-231, respectively. Values were taken from spectra of lipid droplets only. while. There was no significant evidence of interactions between these two independent variables for these ratios. Balanced 2-way ANOVA with constrained sum of squares results of CD_L ratios indicate insulin significantly influenced the CD_L: CH_2 ratio in MCF-10A lipid droplet spectra with a rejection level of $^{\#\#\#\#}P < 0.0001$, but no significant evidence of interactions between these two independent variables. However, in TNBC insulin, methionine, and the interaction term significantly influenced the CD_L : CH_2 ratio in MDA-MB-231 lipid droplet spectra with a rejection level of $^{\#\#\#\#}P < 0.0001$ and $^{\#}P < 0.05$ for the individual and interaction terms, respectively.

Two-way ANOVA (**Supplementary Table 2.2**) confirms a significant interaction term for methionine and insulin concentrations in TNBC for the de novo synthesized lipids relative to the total lipids (**Fig. 2.5 A, B** Right). Contrarily, only the insulin independent variable was significant for the MCF10A in the de novo synthesized lipids relative to the total lipids (**Fig. 2.5 A, B** Right), but no interaction term, or even a significant methionine term. Only in TNBC did the methionine term have a significant impact on this ratio, which lead us to believe the discrepant effects on CH_2:CH_3 ratios we observed (**Fig. 2.2 B**) might arise from differential de novo lipogenesis, rather than protein synthesis and metabolism. Since the excess methionine stimulated de novo lipogenesis (**Fig. 2.5 C**) and was a significant term in the ratio of de novo synthesized lipids to proteins for both cell lines (**Fig. 2.5 A, B**), methionine is likely to preferentially stimulate lipid production more than protein production. Despite these findings, the relative proportion of lipids to proteins in MCF10A still decreases under excess methionine (**Fig. 2.2 B**). Therefore, either the pool of proteins must be getting larger, or the lipid utilization must increase. Excess methionine did not stimulate lipid utilization because CD_L:CH_2 did not significantly increase (**Fig. 2.5 A** Right). This leads us to believe that the protein signal must increase excess methionine. However, excess methionine did not stimulate protein production faster to a greater extent than lipid production since the CD_L:CD_P slightly increased under excess methionine (**Fig. 2.5 A** Left). If there was no relative increase in protein nor decrease in lipids, then MCF10A may not breakdown proteins as much in the presence of excess methionine, or uptake and retain the excess methionine itself more efficiently than TNBC. The hydrophobic amino acid can interact with the acyl chains of the fatty acids in lipid droplets, and since the excess methionine supplied was not deuterated, this protein would not appear in the cell silent region. This way, the excess methionine can affect the CH_3 peak without affecting the CD_P peak, and explain the behavior observed in **Fig. 2.2** and **Fig. 2.5**. Excess methionine can also incite endoplasmic reticulum stress due to complex interactions with cysteine pathways since both are sulfur containing and are critical in protein

folding due to disulfide bonds. These misfolded proteins may be sequestered by LDs differently across subtypes.

2.2.2 Morphological Changes in Lipid Droplet

3D SRS images were taken for each individual cell at 2850cm^{-1} (**Fig. 2.6 A-D**) to assess the size and number of LDs more accurately. LDs were computationally segmented using MATLAB (**Fig. 2.6 E-H**) to acquire individual LD volume and number of LDs per cell. The addition of excess methionine produced the most noticeable changes in lipid droplet morphology – a decrease in lipid droplet number but increase in volume. This effect was observed in both MCF10A and TNBC cells (**Fig. 2.6 I, J**). Of note is the insulin restricted case in TNBC cells, which had no discernible change to lipid droplet number or size. Qualitatively, the lipid droplets also appeared more clustered in excess methionine cases. Lipid droplet volume was also observed to slightly increase from restricted insulin (0.1x) to physiological and excess insulin (1x and 2x) in TNBC under excess methionine conditions. This corroborates the potential interplay between insulin and methionine in TNBC.

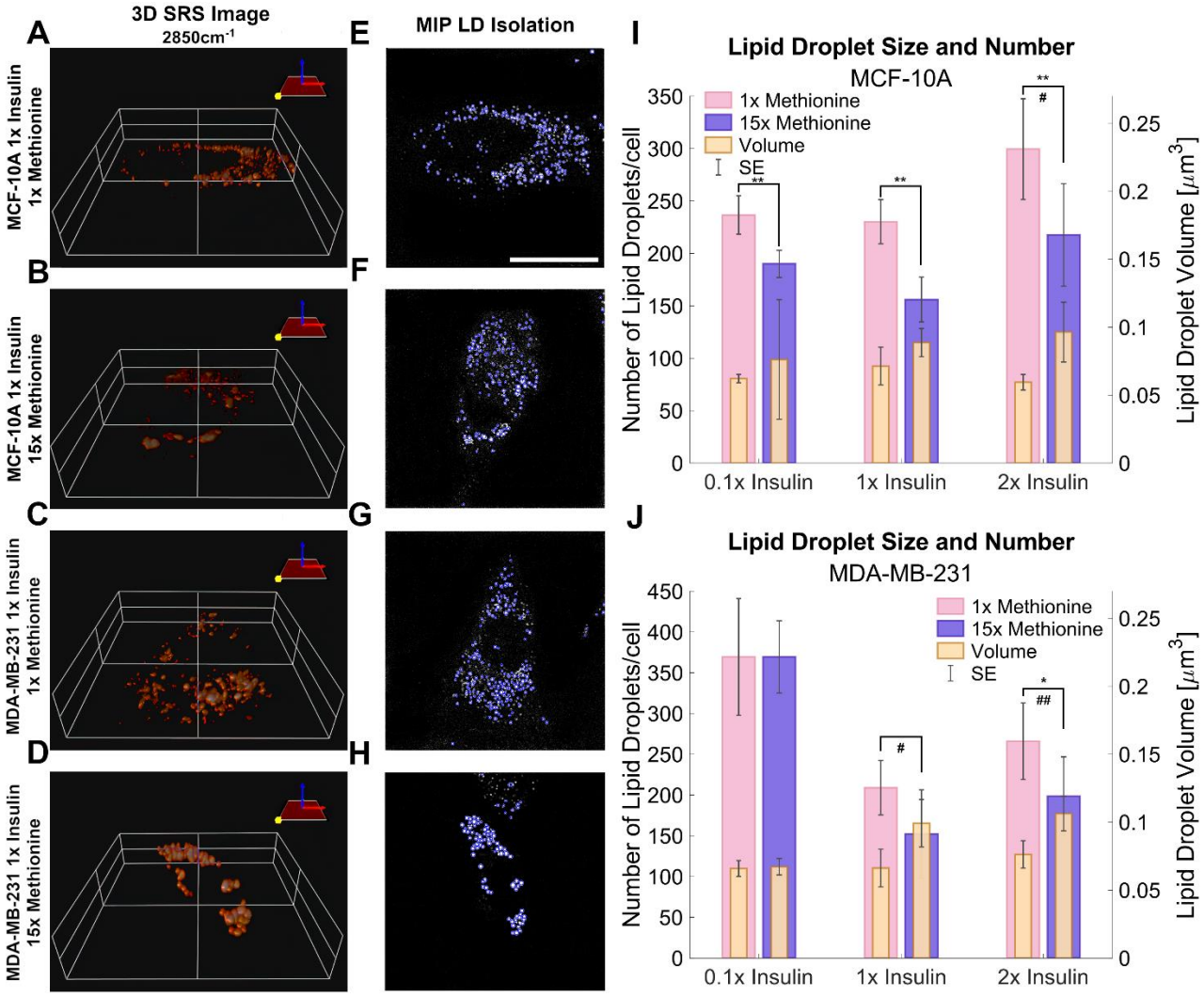


Figure 2.6 3D SRS image lipid droplet analysis
(**A-D**) 3D isosurface reconstructions of the single cell SRS images taken at 2850cm^{-1}. **(E-F)** LD segmentation shows representative maximum intensity projections of SRS image stacks shown in (A-D) with lipid droplets highlighted in blue outlines. **(I-J)** Quantitative LD structure summary shows average lipid droplet number and volume for each experimental group. Excess methionine groups display decreased lipid droplet number and increased size. Lipid droplets also appear qualitatively more clustered in excess methionine as well. Scale bar is 20 μm. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for LD number: *P<0.05, **P<0.01. Octothorps '#' correspond to the following p-values for LD volume: #P<0.05, ##P<0.01. Scale bar is 20 μm.

Lipid droplet distribution can be a major indicator of cell cycle status, nutrient availability, and ER stress [82]. LD size may influence the degree to which beta oxidation occurs in cells and be affected by mitochondrial recruitment during LD expansion in nutrient rich environments. Regardless, the physical contact between these organelles is thought to mediate their proper

function [83, 84]. A label-free approach to identifying mitochondrial presence near lipid droplets may be the spectral presence of cytochrome C (cytC), which is found in the intermembrane space of mitochondria. Some peaks canonically representative of cytC are the heme backbone at 1558cm^{-1} and the side chains of tryptophan, tyrosine, and phenylalanine in alpha structures at 1610cm^{-1} [85] which were weakly present near the fingerprint region of the spectra in MCF10A and TNBC cells. It was found in **Fig. 2.7 A** and **Fig. 2.8 A** that excess methionine cases displayed a diminished spectral presence of unmixed cytC peaks. This suggests LDs in excess methionine may cluster near other organelles such as lysosomes, or even with other LDs for fusion events, instead of co-localizing with mitochondria for energy. Examples of Gaussian-Lorentzian peak unmixing for MCF10A and TNBC (**Fig. 2.7 B and Fig. 2.8 B**), respectively, with quantitative summaries in the form of bar graphs (**Fig. 2.7 D and Fig. 2.8 C, D**). The number of unmixed peaks was optimized such that the overall fit is accurate, while the unmixed peaks are easily ascribed to canonical protein and lipid deformations. The Amide II' region contains various CH_2 and CH_3 deformations such as wagging, stretching, scissoring, and twisting [86, 87]. The Amide I region contains secondary structure information and has been used to study proteins such as collagen [88]. Between these peaks lies the C-C bond of the heme backbone. MCF10A and TNBC exhibited distinct peak shapes in all areas of this region. In MCF10A, the Amide II' peak had a narrower shoulder at 1458cm^{-1} under excess methionine (**Fig. 2.7 C**), while TNBC had a narrower Amide I peak under excess methionine (**Fig. 2.8 E-G**). The Amide I and II' regions also contain protein and lipid information and have various assignments in the literature. **Fig. 2.8 G** quantifies the width and prominence of the Amide I peak in TNBC with and without excess methionine. Results indicate altered protein folding, in which methionine plays crucial roles. Methionine is not only a protein translational initiator, but its metabolism is also involved in purine synthesis, epigenetic control, and secondary disulfide bond formation [53]. Misfolded proteins have tangible effects on ER stress and lipid droplet distribution and chemistry [42], as these proteins have been shown to accumulate in LDs destined for proteasomal breakdown [82]. In this manner, LDs may

serve as reservoirs and chaperones to mitigate lipid and protein toxicity. Although further investigations are required to confidently assign the phenomenon observed herein, the fact that consistent alterations in these areas were observed using label-free vibrational imaging techniques sets the stage for more in-depth studies of dietary methionine-controlled protein folding in breast cancer cells. **Fig. 2.7 C** quantifies the changes in the Amide II' peak of MCF10A and may indicate altered lipid and protein structure as well. Various bond deformations occur at slightly different wavenumbers, with CH_2 scissoring being red-shifted with respect to CH_2 stretching. Acyl chains of different length and saturation may influence the degrees to which each of these deformations take place. Further investigation into purified LD content with other techniques such as gas chromatography and mass spectrometry are warranted. Spectroscopic data are usually sensitive to baseline correction, background subtraction, and normalization methods, and are therefore better suited to relative observations, while chromatography and spectrometry offer absolute quantification and detailed chemical structure. Conjugating these techniques is beyond the scope of this label-free optical platform, but is promising and critical step in progressing this technology.

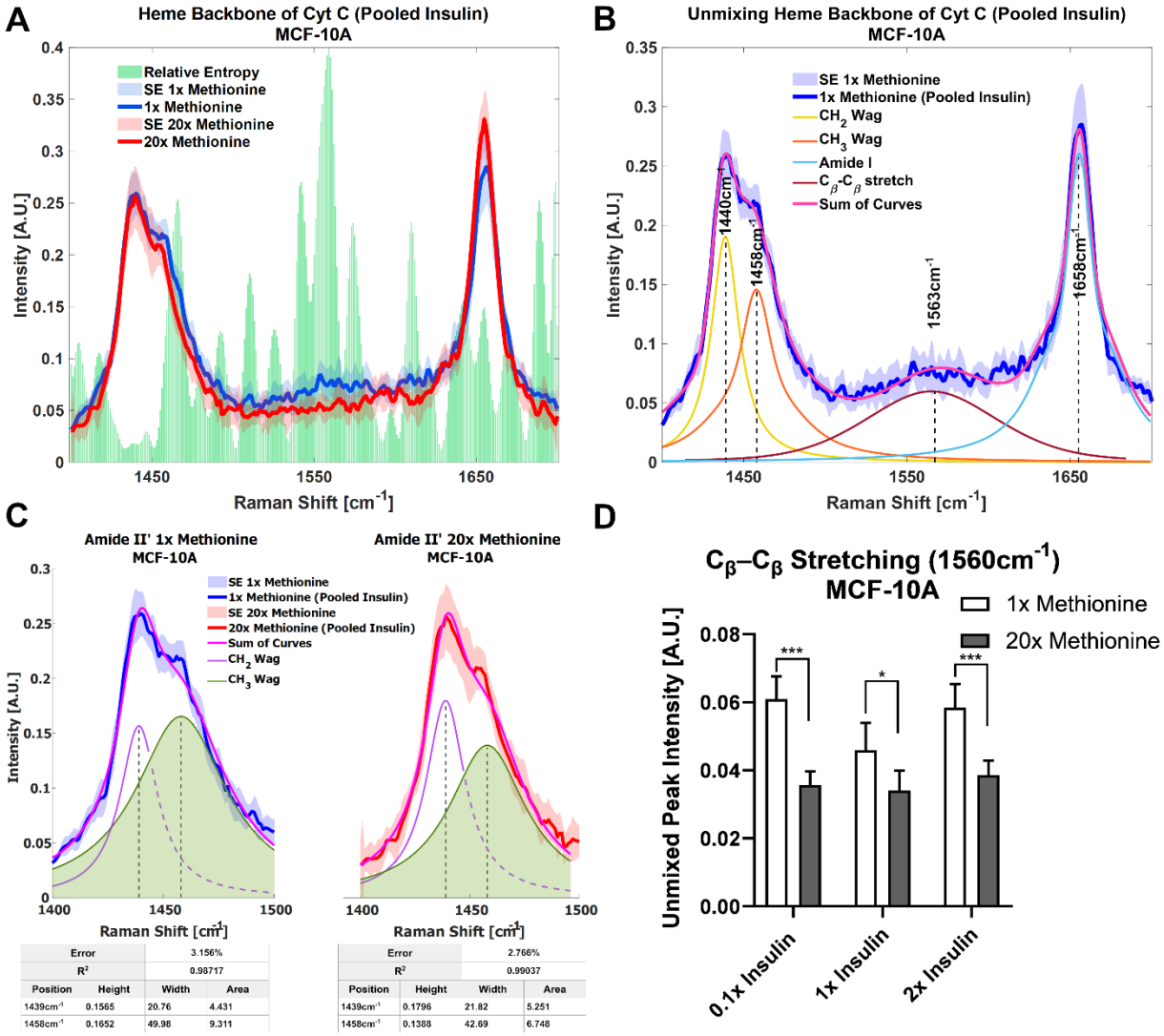
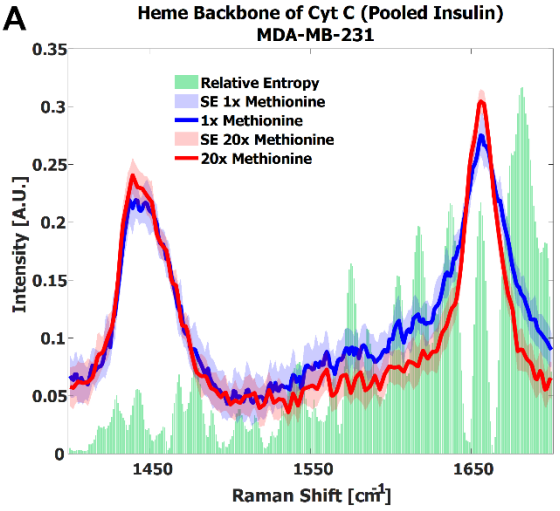
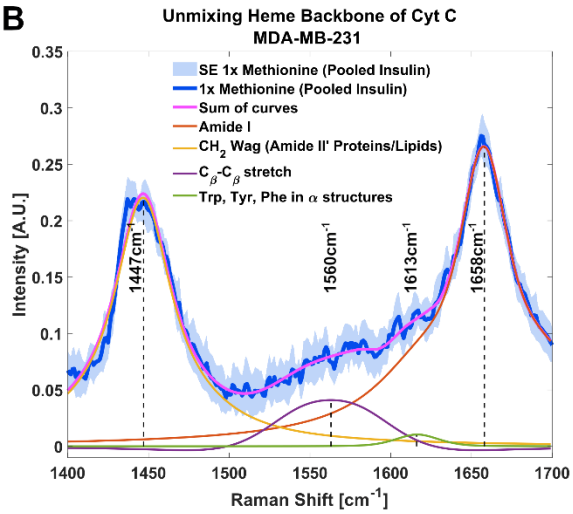
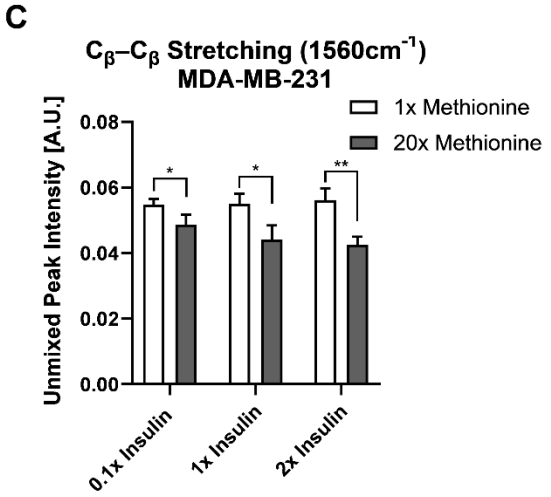


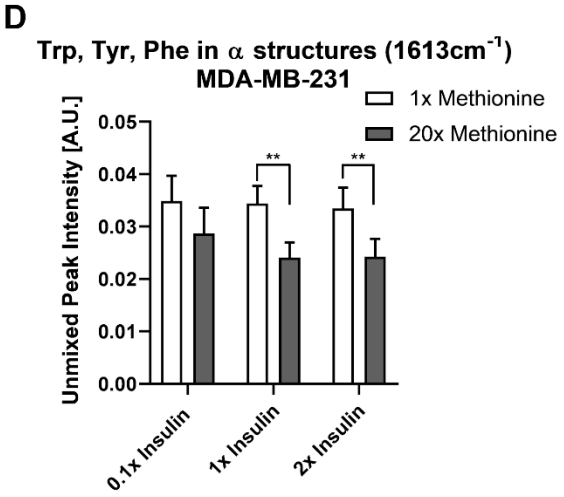
Figure 2.7 Spontaneous Raman Spectroscopy detects CytC presence and protein folding
(A) normal vs excess methionine Expanded view of lipid droplet spectra grouped by methionine concentration shows a high relative entropy in the 1550cm^{-1} region, ascribable to the heme backbone of cytochrome C. **(B)** Unmixing Peaks with four peaks using a Gaussian-Lorentzian blend yields an error of 2.367% and an R^2 of 0.98854. **(C)** Amide II' Peak Shoulder shows an **e**xpanded view of normal and excess methionine groups' Amide II' regions highlight a relatively narrowed shoulder at 1458cm^{-1}. Unmixed peaks follow the overall shape of the average Amide II' peaks, with the error and correlation coefficient reported in the table below. Width and area information is also summarized in the table to clearly communicate the disparate shoulder widths. **(D)** Quantitative summary of the heme backbone unmixed peak intensities for each experimental condition of MCF-10A cultures shows decreased spectral presence. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values: *P<0.05, **P<0.01.

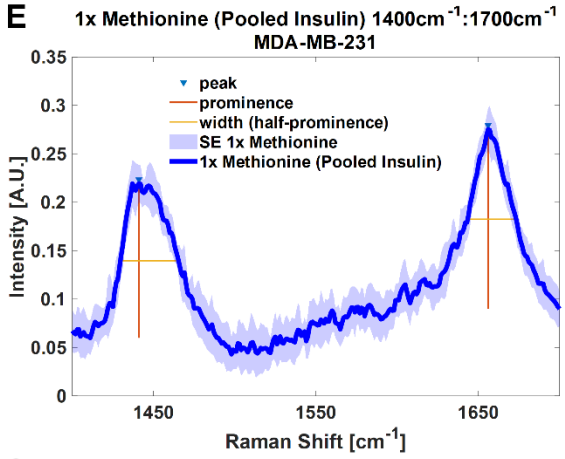
Figure 2.8 Raman Spectroscopy detects CytC presence and protein folding differences
(A) normal vs excess methionine shows expanded view of lipid droplet spectra grouped by methionine concentration shows a high relative entropy in the 1550cm^{-1} and 1650cm^{-1} regions, ascribable to the heme backbone of cytochrome C and side chains of tryptophan, tyrosine, and phenylalanine, respectively. (**B**) Unmixing peaks with four peaks using a Gaussian-Lorentzian blend yields an error of 2.620% and an R^2 of 0.98544. (**C-D**) Quantitative summary of the unmixed peak intensities for each experimental condition of MDA-MB-231 cultures shows decreased spectral presence of cytochrome C. **(E-F)** Amide I peak width shows an expanded view of lipid droplet from 1x Methionine (LEFT), and 20x Methionine (RIGHT) experimental conditions. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values: *P<0.05, **P<0.01. **(G)** Peak analysis shows that the peak prominence and peak width at half prominence is significantly narrower at the Amide I region in excess methionine lipid droplet spectra.

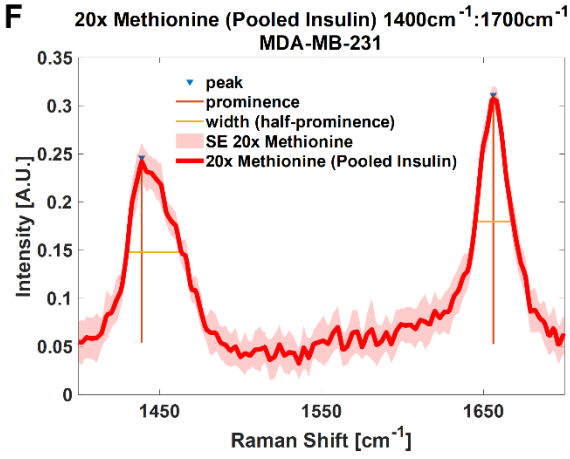












G

1x Methionine			20x Methionine	
Amide II'	Amide I		Amide II'	Amide I
0.236 ± 0.031	0.286 ± 0.028	peak height	0.264 ± 0.030	0.32319 ± 0.019
34.80 ± 1.96	**28.97 ± 2.83**	peak width	34.55 ± 3.11	**23.58 ± 2.31**
0.189 ± 0.017	**0.199 ± 0.024**	peak prominence	0.218 ± 0.023	**0.268 ± 0.027**

2.2.3 Lipid Peroxidation Status

Another global lipid response to excess methionine takes form in the lipid peroxidation status. Under oxidative stress, long chain unsaturated fatty acids can undergo a vicious cycle of lipid peroxidation [79]. Several Raman shifts have been used to describe the degree of unsaturation of fatty acids, including the one near 3010 cm^{-1} that corresponds to the H-C= stretching region [79]. Interrogating the relative entropy plot in **Fig. 2.4 C**, we find that the saturated lipid peak at 3010 cm^{-1} and the lipid peak at 2850 cm^{-1} both rank highly for both cell types, but TNBC is more heavily influenced by methionine concentration. That is, we can see from the spectroscopic data that MCF10A, whether L-methionine was normal or in excess, expressed relatively different levels of unsaturated lipids depending on the level of insulin. This suggests that de novo synthesis of branched chain fatty acids, or perhaps their accumulation in LDs was upregulated in excess insulin conditions. So, while insulin was critical in influencing de novo synthesis of lipids in TNBC, it may not influence lipid peroxidation as much as methionine does. **Fig. 2.9** shows the effects of excess methionine in TNBC using multi-modal optical techniques.

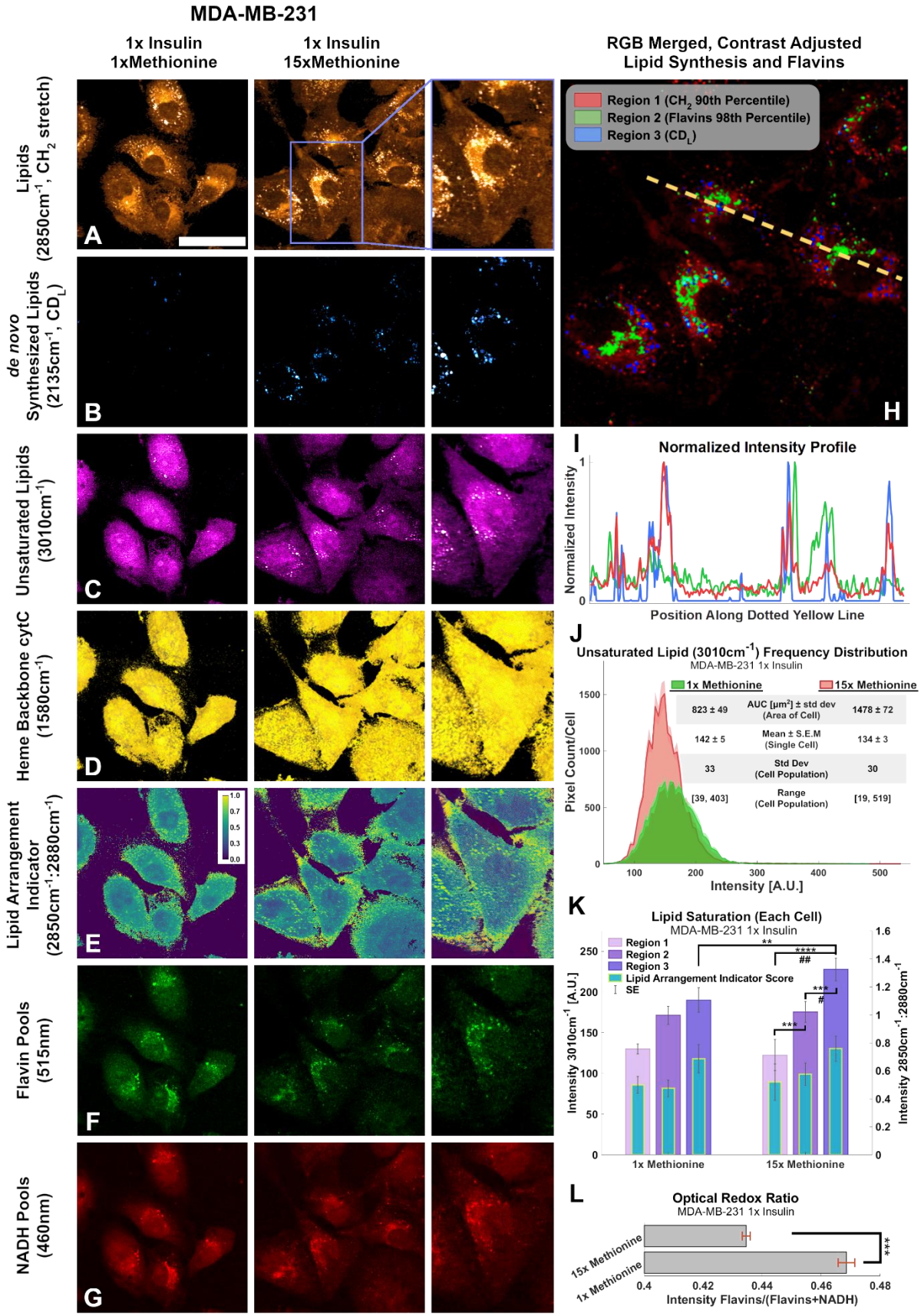
Conjugated SRS and TPF microscopy display spatial distributions of points of interest regarding excess methionine effects in TNBC cells (**Fig. 2.9 A-G**). These results also corroborate with the spectroscopic data. **Fig. 2.9 B** reveals that the cells undergo enhanced de novo lipogenesis under excess methionine with respect to control groups. Contrary to expectations, the unsaturated lipid signal in the excess methionine group was weaker than the control overall but was stronger near the large lipid droplets (**Fig. 2.9 C**). This information is lost in spectral acquisitions alone because spectra were obtained from lipid droplets only. Excess methionine treated cells exhibited larger cross-sectional area (**Fig. 2.9 J**) and may be due to the cells being more spread out and flatter. Due to the point spread function of the confocal laser scanning microscope, this spreading out of the cells may contribute to an apparent decrease in concentration of unsaturated fatty acids because the scattering cross section along the beam path

is smaller. Consistent contrast makes it difficult to discern the abundance of smaller lipid droplets in the control images without oversaturating the excess methionine images. There were no discernible differences in spatial distribution of heme groups at the 1580 cm^{-1} (**Fig. 2.9 D**), but co-localisation algorithms may help in future studies. The indicator of crystalline arrangement in lipids corresponding to the symmetric:antisymmetric CH_2 stretching ratio indicates that the excess methionine group may have less lipid saturation near the plasma membrane (**Fig. 2.9 E**). Higher ratios would indicate a lower concentration of 2880 cm^{-1} species, which has been ascribed to the Fermi resonance of CH methylene groups [78]. This ratio has been found to inversely correlate with thermodynamic stability, and when the ratio is larger, there may be less lattice order in the structure [89]. In the context of cell membranes, fluidity and saturation are critical functional properties, and the decreased lipid saturation score near the plasma membrane may also contribute to the observed “flattness” of the cells with excess methionine, as the cells may be able to spread out more easily.

Different areas of the cells provide niche microenvironments, in which lipid peroxidation may vary. Three subcellular regions of interest include where all lipids exist (**Fig. 2.9 H**, region 1), where flavins are more present (region 2), and where newly synthesized lipids are present (region 3). As shown in **Fig. 2.9 H, I**, these regions do not necessarily overlap. Flavins have been shown to report on oxidative stress, and certain flavin enzymes have been associated with lipid peroxidation as well. The quantitative image analyses of unsaturated lipids (3010cm^{-1}) and the lipid arrangement indicator ratio (2850cm^{-1}:2880cm^{-1}) are summarized in **Fig. 2.9 K**, in which distinct regions are separately quantified. **Fig. 2.9 J** highlights a larger cross-sectional area of the imaged cells, which may be afforded by a more fluid cell membrane. In certain cells, oxidative stress has been found to increase lipid saturation for protection. Furthermore, the presence of higher ratios near LDs suggests there is less synthesis of saturated lipid species as well. Finally, flavin autofluorescence decreased in the presence of excess methionine (**Fig. 2.9 F**), while NADH autofluorescence remained more consistent (**Fig. 2.9 G**). The flavin/(NADH + flavin) ratio has

been shown to be an indicator of oxidative stress and estimator for NAD+:NADH [90]. Results corroborate with previous studies in which this ratio was used to differentiate breast cancer cell lines [91], with the TNBC having relatively weaker flavin autofluorescence than the normal-like cell type. Under oxidative stress, this ratio has been shown to increase. A decrease here (**Fig. 2.9 L**) may demonstrate the antioxidant properties of methionine. Flavin autofluorescence data is summarized in **Supplementary Fig. 2.14**.

Figure 2.9 Multi-modal optical analysis of MDA-MB-231 methionine response with insulin
(A-G) Multichannel Images illustrate SRS and TPF image channels of interest for lipidomic responses to excess methionine. **(H)** Overlaid Composite Regions image of the 15x Methionine lipid (CH_2), flavins, and de novo synthesized lipids (CD_L). Channels were masked according to the indicated thresholds using ImageJ and contrast was adjusted for optimal clarity. **(I)** Intensity profile plot depicts the intensities of pixels along the dotted yellow line shown in (H) of each of the three composite channels. **(J)** Composite intensity histograms of the unsaturated lipid channel ($3010cm^{-1}$). Bolded distribution outlines represent the average frequency of pixel intensities among the cells in each group. Shaded areas around the bolded distribution outline represent the standard error of the mean of each bin of pixel intensities. Each distribution curve represents the pixel intensities of a cell sampled from the experimental condition. **(K)** Quantitative Lipid saturation summary depicts the $3010cm^{-1}$ pixel intensities (Left axis) in each of the three regions shown in (H) of a typical cell from the indicated experimental condition. Additionally, the lipid arrangement indicator ratio ($2850cm^{-1}$:$2880cm^{-1}$) for each of the regions in (H) is also depicted for the typical cell from each experimental condition. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for the unsaturated lipid peak ($3010cm^{-1}$) intensities: **P<0.01, ***P<0.001, ****P<0.0001. Octothorps '#' correspond to the following p-values for the Lipid Arrangement Indicator ($2850cm^{-1}$: $2880cm^{-1}$) intensities: #P<0.05, ##P<0.01. **(L)** ORR for the typical cell from each experimental condition. Results corroborate and extend spectral data findings, as well as previous third-party studies. Scale bar is 50um. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for the unsaturated lipid peak ($3010cm^{-1}$) intensities: ***P<0.001.


MDA-MB-231
1x Insulin
1xMethionine
1x Insulin
15xMethionine
RGB Merged, Contrast Adjusted
Lipid Synthesis and Flavins
Lipids
(2850cm-1, CH2 stretch)
de novo
Synthesized Lipids
(2135cm-1, CDL)
Unsaturated Lipids
(3010cm-1)
Heme Backbone cytC
(1580cm-1)
Lipid Arrangement
Indicator
(2850cm-1:2880cm-1)
Flavin Pools
(515nm)
NADH Pools
(460nm)
Region 1 (CH2 90th Percentile)
Region 2 (Flavins 98th Percentile)
Region 3 (CDL)
A
B
C
D
E
F
G
H
I
Normalized Intensity Profile
Normalized Intensity
Position Along Dotted Yellow Line
J
Unsaturated Lipid (3010cm-1) Frequency Distribution
MDA-MB-231 1x Insulin
1x Methionine
15x Methionine
823 ± 49
AUC [μm2] ± std dev
(Area of Cell)
1478 ± 72
142 ± 5
Mean ± S.E.M
(Single Cell)
134 ± 3
33
Std Dev
(Cell Population)
30
[39, 403]
Range
(Cell Population)
[19, 519]
Pixel Count/Cell
Intensity [A.U.]
K
Lipid Saturation (Each Cell)
MDA-MB-231 1x Insulin
Region 1
Region 2
Region 3
Lipid Arrangement Indicator Score
SE
Intensity 3010cm-1 [A.U.]
Intensity 2850cm-1:2880cm-1
1x Methionine
15x Methionine
L
Optical Redox Ratio
MDA-MB-231 1x Insulin
15x Methionine
1x Methionine
Intensity Flavins/(Flavins+NADH)

2.3 Discussion

For the first time, the unique lipid metabolism of triple negative breast cancer was studied under tandem excess methionine and insulin conditions, and revealed key insights that span the metabolic, spatial, and biochemical dimensions. Not only did this study confirm lipid droplets are reflective of cellular phenotypes and demonstrate their efficacy in classifying breast cancer subtypes, and even phenotypes, it improves morphological analysis using 3D imaging, as opposed to 2D, and efficiently displays relevant chemical disparities using the first demonstration of relative entropy for Raman data. Considering the critical impact lipid metabolism has on the progression of diseases such as cancer, the analyses on lipid saturation and peroxidation, optical redox status, and LD size and distribution solidify the effects of methionine and insulin, which may prove to be therapeutic targets for breast cancer in the future.

These experiments demonstrate the power of nearly label-free optical techniques to probe LD phenotypes for the study of TNBC's unique metabolism. Methionine dependence, also known as the Hoffman effect, has been explored in TNBC and other cancers, but fewer studies explored the effects of excess methionine, and fewer still, the tandem manipulation of methionine and insulin. Upon the addition of insulin in TNBC growth media, macromolecular changes appeared in the CH stretching region of excess methionine treated cells, as the CH_2:CH_3 ratio increased in TNBC, but decreased in MCF10A control cells. A potential pathway that involves both insulin and methionine in LD metabolism may be mediated by TNBC's elevated mTOR activity, and was explored through the chemometric, spatial, and molecular imaging dimensions with subcellular resolution. Currently the stoichiometric mass action of this pathway remains to be investigated in these breast cancer subtypes, but several studies have linked methionine, mTOR, and insulin signaling pathways [54–59], albeit transitively. Paramount in this investigation is the implication of these metabolites in the pursuit of TNBC diagnosis and treatment. Unmixing the interplay

between insulin and methionine may afford targeted therapies that address the rampant lipid metabolism that facilitates breast cancer progression.

LD chemical composition also demonstrated excellent classification ability, as lipid and protein Raman cross sections are not only larger, but also very diverse and highly implicated in metabolic cellular states. Classification of breast cancer subtypes, and even more so their phenotypic states, can be critical in improving patient outcomes due to the necessity of early diagnosis. MCF10A cells may exhibit differential protein metabolism by retaining scavenged methionine or not metabolizing proteins as much as TNBC, which is due, in part, to very different doubling times. Several other considerations including LD size may also contribute to these disparities, as larger LDs have a greater volume:surface area ratio, and thus a greater lipid:protein ratio since the apolipoproteins stud the phospholipid surface of the lipid core. LD fusion can affect this as well, since newly synthesized LDs may have a diluted CD signal if fused with older LDs. Further investigation is warranted to uncover the details of differential lipid metabolism in breast cancer subtypes using LDs, but this platform has set the stage for macroscopic observations using efficient optical techniques.

Both TNBC and MCF10A cells exhibited similar spatial information of LDs under these treatments as excess methionine conditions generally decreased the number of lipid droplets while increasing their volume in both cell types, while insulin generally increased both size and number of LDs. Insulin restriction appeared to increase LD number as well, and in TNBC, appeared to attenuate the effects of excess methionine on LD number. This interesting pattern not only suggests that TNBC has unique insulin-mediated lipid metabolism, but that insulin and methionine may have more complex concentration-dependent interactions in general as well. LD distribution also seemed to be more clustered in the excess methionine cases, and we intuit from the diminished spectral presence of cytC that these lipid droplets were less co-localized with mitochondria. Finally, the antioxidant properties of methionine expectedly diminished flavin autofluorescence and resulting lipid droplet spectra showed higher degrees of lipid unsaturation.

In **Fig. 2.9 K**, l the optical redox ratios and the lipid arrangement indicator ratios indicate that methionine plays a large role in lipid peroxidation and saturation. The degree of saturation of lipids is a critical consideration for cell membrane fluidity, especially in aggressive cancers that can alter their extra cellular matrix (ECM), or those that metastasize and migrate rapidly. For the first time, the dynamics of lipid saturation and peroxidation under nutritional control has been imaged with label-free subcellular resolution.

To broaden the scope of the investigation and capitalize on the rich chemical data of the Raman spectrum, relative entropy was used to rank the features that exhibited the greatest variance between different groups. As expected, there are several areas other than the CH stretching region that offer strong classification ability despite lower Raman intensities. This may be attributed to the higher intensity deviations at higher intensities typical of multiplicative scattering effects. Additionally, the effects of individual nutritional manipulations become clearer with all Raman peaks being visible simultaneously. From this insight, the interrogation of pathways with Raman spectra can be done more efficiently, as the relative entropy scores for each Raman shift can be seen at once, reducing the number of spectra and subplots that need to be displayed. With this demonstration of efficacy, more critical quantitative analyses, as well as algorithmic improvements will be conducted. For example, incorporation of directional shifts in intensities can be made visible on the relative entropy plot, as opposed to absolute distance metrics alone. This will not only identify discriminating variables but will also circumvent the need to manually determine significant ratios, ratio differences, and other trends as well. Further, this relative entropy plot may be useful in feature reduction, so that fewer hyperspectral images may be required for discriminating LD microenvironments and subpopulations. Other methods more directly identify the wavenumbers that contribute the most to a spectrum's classification, such as the hybrid variable combination population analysis (VCPA) and iterative retaining important variables (IRIV) approach [92]. However, due to the large number of variables, IRIV can be time and resource intensive.

The diverse pathophysiology of breast cancer may have important mechanisms involving methionine and insulin that can be studied with optical techniques such as spontaneous Raman spectroscopy and SRS/TPF microscopy. This study also emphasizes that LDs are organelles diverse in structure and function and can yield rich metabolic information when interrogated by Raman techniques. Future studies that involve automated high-throughput acquisitions of spectra and images at more finely tuned concentrations of insulin and methionine may increase the power of the results discussed here. Different distribution fits for the relative entropy algorithm, displaying directionality of peak intensity changes, as well as the multiplexing of dietary manipulations such as glucose, pyruvate, and glutamine may paint a clearer picture of the metabolic dynamics in breast cancer [93, 94]. This will also help make hyperspectral imaging more efficient in terms of disk space and clustering ability. Utilizing morphological characteristics and intensity changes to augment classification has not been performed in this study but will be a prudent next step in developing these optical techniques for classification purposes. Additionally, spatial distribution of LD by size and chemometric composition, as well as quantitative descriptions of LD distribution and co-localization would further enrich this investigation. This kind of quantitative hyperspectral image data will bolster the utility of LD analysis in the study of breast cancer, and ultimately improve not only our understanding of the complex disease, but patient outcomes in eventual translation as well.

2.4 Materials and Methods

2.4.1 Experimental Design

An experimental outline is shown below (**Fig. 2.10**). First, three cell subtypes were grown in media with either 1x methionine (0.03g/L) or 20x methionine (0.6g/L). Then the experiment was repeated with the addition of 3 insulin concentrations for each of the groups to investigate their relationship.

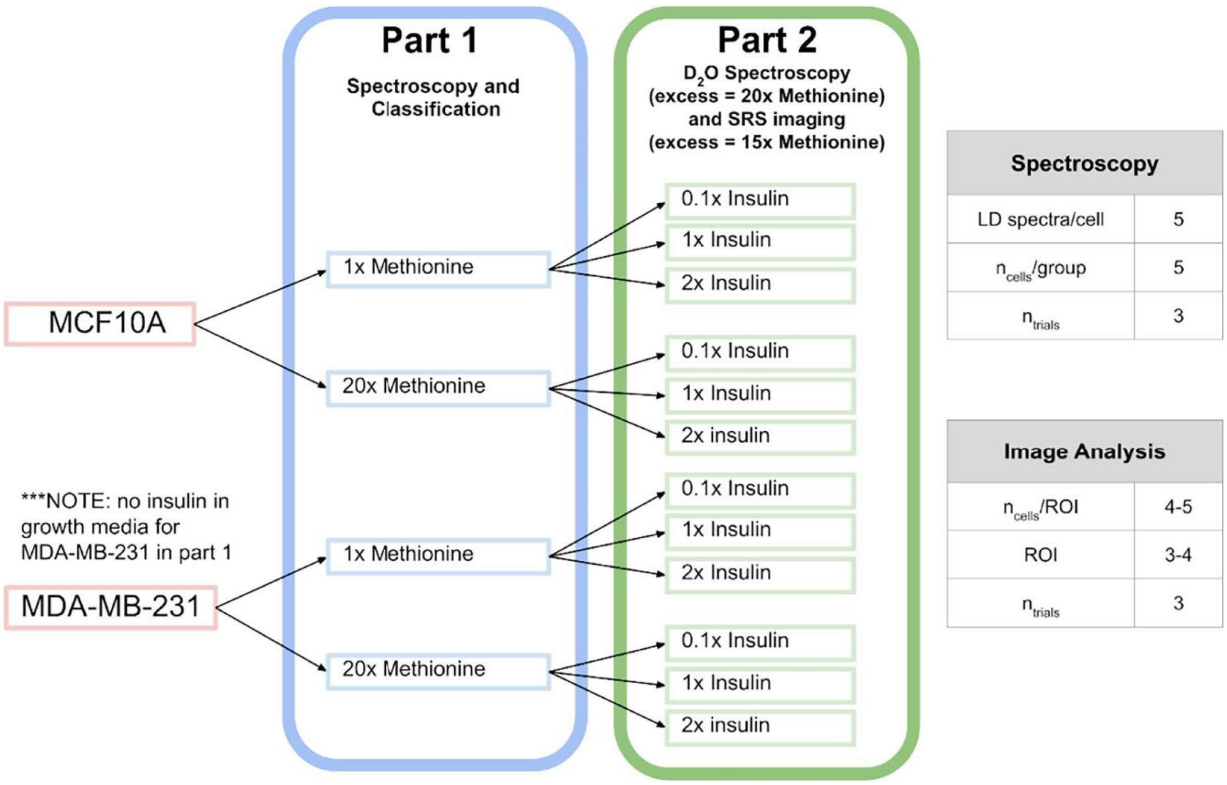


Figure 2.10 Experimental Design
Graphical illustration of the groups in this study. Only methionine concentration is modulated at first, and then insulin and methionine were modulated in tandem. DMEM used in this experiment already contains 0.03g/L of L-methionine, which corresponds to the 1x methionine group. NOTE: in part 2, the excess methionine concentration is 20x for the Raman spectroscopy, and 15x for SRS imaging.

2.4.2 Cell Culture

Human triple negative breast cancer cell line (MDA-MB-231) and normal-like breast epithelial cell line (MCF10A) were cultured in Dulbecco's modified Eagles' medium (DMEM), supplemented with 5% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Fisher Scientific, Waltham, MA), and incubated with 5% CO2 at 37°C. Cell cycles were synchronized using double thymidine block [29]. After passaging at 80% confluence, cells were seeded at a concentration of 2×10^5/mL atop 70% ethanol-soaked cover glass in 24-well plates and incubated for 8 hours. Then the growth media was changed to 50% heavy water (D_2O) and treatment media as follows.

For MDA-MB-231 and MCF10A cell culture media, 57 mg/L and 42 mg/L L-methionine (M8439, Sigma Aldrich) was added to DMEM for the excess methionine group for spontaneous Raman spectroscopy and SRS imaging (20x and 15x concentration), respectively. The DMEM

powder used in this study already contains 0.03g/L (1x concentration) L-methionine and corresponds to the physiological concentration group. The reason for lowering the excess methionine concentration for SRS imaging analysis is because the cell morphological changes were more varied and poorer with 20x methionine, making it more difficult to acquire quantitative metabolic activity from images on a per-cell basis. Insulin (Sigma Aldrich, St. Louis, MO) was added at 1µg/mL, 10 µg/mL, and 20 µg/mL for the 0.1x, 1x, and 2x insulin groups, respectively.

Cells were incubated for 48 hours, which corresponds to a deuterium-retarded cell cycle. Cyclin dependent kinase 1 (CDK1) inhibitor (RO-3066, Sigma) was added with 8 hours remaining to arrest growth before mitosis. Cells were gently rinsed with 1x PBS with Calcium and Magnesium ions at 4°C (Fisher Scientific, 14040216), and finally fixed in 4% methanol-free PFA solution (VWR, 15713-S) for 15 minutes. The cover glass was mounted on 1mm thick glass microscope slides with 120 µm spacers filled with 1x PBS without calcium and magnesium ions. These samples are stored at 4°C submerged in PBS when not in use.

2.4.3 Spontaneous Raman Spectroscopy

Spontaneous Raman scattering spectra were obtained by a confocal Raman microscope (XploRA PLUS, Horiba) equipped with a 532 nm diode laser source and 1800 lines/mm grating. The acquisition time is 30 s with an accumulation of 4. The excitation power is ~40 mW after passing through a 100x objective (MPLN100X, Olympus). The background spectra were taken for each LD at the same focus plane as the LD and were subtracted from each LD spectrum immediately. Spectra were preprocessed using vector normalization and simplex normalized. Peaks were normalized to the protein peak at 2935cm^{-1}. Previous studies suggest Raman microspectroscopy can quantify lipids non-invasively [95].

2.4.4 Stimulated Raman Scattering Imaging Microscopy

An upright laser-scanning microscope (DIY multiphoton, Olympus) with a 25x water objective (XLPLN, WMP2, 1.05 NA, Olympus) was applied for near-IR throughput. Synchronized

pulsed pump beam (tunable 720–990 nm wavelength, 5–6 ps pulse width, and 80 MHz repetition rate) and Stokes (wavelength at 1032nm, 6 ps pulse width, and 80MHz repetition rate) were supplied by a picoEmerald system (Applied Physics & Electronics) and coupled into the microscope. The pump and Stokes beams were collected in transmission by a high NA oil condenser (1.4 NA). A high O.D. shortpass filter (950nm, Thorlabs) was used that would completely block the Stokes beam and transmit the pump beam only onto a Si photodiode for detecting the stimulated Raman loss signal. The output current from the photodiode was terminated, filtered, and demodulated in X with a zero phaseshift by a lock-in amplifier (HF2LI, Zurich Instruments) at 20MHz. The demodulated signal was fed into the FV3000 software module FV-OSR (Olympus) to form the image during laser scanning. All 3D lipid droplet images were obtained with a pixel dwell time 40 µs with 3-frame averaging for a total imaging speed of ~10-15 min per image stack. Laser power incident on the sample is approximately 40mW.

2.4.5 Two Photon Fluorescence Microscopy

Autofluorescence of flavins was excited at 820 nm and autofluorescence of NADH was excited at 780nm using the same tunable picosecond laser described in chapter 1.3. Epi-detected emission of flavin autofluorescence was collected using a 460 nm filter cube (OCT-ET460/50M32, Olympus), and NADH was collected using a 515nm filter. These images were also 512x512 pixels and were acquired with a 12.5 µs/pixel dwell time using a 300mW power at the laser shutter. Autofluorescence images were background subtracted using a rolling ball algorithm with a radius of 50px, which is intended to approximate cell size in these images.

2.4.6 Data Analysis

Spectral Clustering

Previous studies have shown these breast cancer subtypes have unique Raman features [45, 96]. Machine learning was conducted to determine the extent to which these features can be used to segregate these subtypes and be augmented using the metabolic dimension of excess

methionine. Neural network classification was done using a simple multi-layer perceptron (MLP) model with 100 neurons in the hidden layer and a rectified linear unit (ReLU) activation function for each neuron. An L2 regularization term with hyperparameter α=0.0002 penalizes the model for incorrect classification during learning with cross-entropy loss minimization. The classification is stochastically optimized using an adaptive moment estimation algorithm called Adam. Advantages of this choice of activation function and solver in an MLP include invariance to rescaling gradients, the ability to learn non-linear models, and a natural simulated annealing to optimize the gradient [97]. MLP are, however, sensitive to parameter tuning, and all spectra were normalized to have the same range.

The input for the MLP model consists of a matrix of Raman spectra, and a vector of target classes. In this study, target classes are of categorical type and correspond to the cell subtype and metabolite concentration groups. Each spectrum Xi is represented as a vector containing m wavenumbers which are each input into a first layer of neurons. Each neuron in the hidden layer accepts the weighted linear combination of input features and applies the ReLU activation function, outputting the data to the output layer. Softmax is used to probabilistically determine the target class of the multiclass model. The model is trained via backpropagation to minimize cross-entropy loss with a maximum of 200 iterations in which subsequent weight vectors reflect a subtracted loss gradient according to equation 2.1 below.

$$\boldsymbol{W^{i+1} = W^{i} - \epsilon \nabla Loss_{W}^{i}} \qquad \textbf{2.1}$$

where, ε corresponds to the learning rate. A python implementation of model is readily available from scikit-learn v0.24.1 [98]. The width of the hidden layer, k=100, as well as its depth of 1 single layer, are tunable depending on the dataset. Larger datasets may require more neurons and deeper networks to perform better. The geometric mean of input variables and classes roughly totaled 100, and an underlying assumption of a simple binary effect of excess methionine and very distinct Raman spectra was comfortable with only a single layer. However, multiplexing of variables such as cell subtype, methionine concentration, and other manipulations may intuitively

justify additional hidden layers in future experiments. Classification in this study mainly attempts to highlight the higher dimensionality of methionine's non-linear effects on breast cancer subtypes, and discuss potential pathway interactions for further investigation. That is, if all breast cancer subtypes responded similarly and to a similar degree, more complex neural nets may not be necessary to achieve good performance.

Principal component analysis (PCA) is performed using Orange 3.26 on the pre-processed data. The first 10 PCs are used as the vectors for tSNE visualization without any exaggeration and a perplexity of 30.

2.4.7 Selected Raman Feature Analysis

Spectroscopic data is extracted using MATLAB and is plotted using MATLAB. To visualize the influence of all Raman peaks on classification simultaneously, the relative entropy method described in **section 1.2.1** is employed. To date, this method has not been previously reported on Raman spectroscopy applications. Second derivative barcode analysis may be coupled with this method, and improve it, as barcode analysis contains both amplitude and width information [99].

2.4.8 Image Analysis

Images were processed using MATLAB and ImageJ. 3D image stacks of lipid droplets underwent bandpass filtering to suppress horizontal noise artifacts from laser beam scanning, and smoothed. Lipid droplets received a sphericity score based on Euclidean distance from perfect spheres emanating from the center of mass of the lipid droplet to the surface of the lipid droplet. Those with low sphericity scores were discarded. Autofluorescence images underwent sliding paraboloid background subtraction before manual cell segmentation and measurement was conducted via ImageJ.

2.5 Supplementary Material

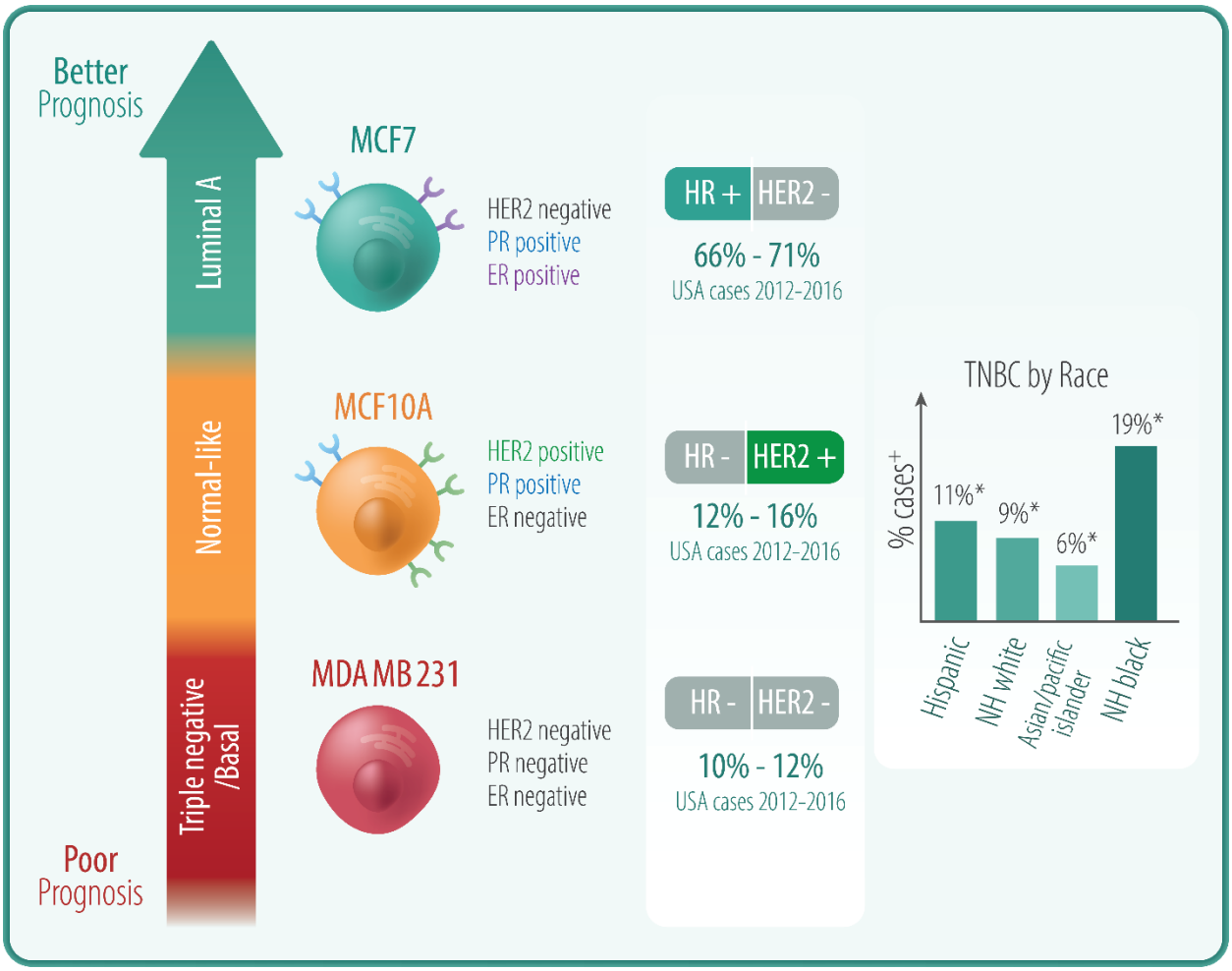


Figure 2.11 Epidemiology of breast cancer
Breast cancer prognosis may be heavily influenced by subtype. Subtypes may be classified by membrane proteins that afford hormone sensitivity but may be further delineated by other aspects of the proteome and metabolic phenotype. TNBC affects non-Hispanic black women more frequently than any other race, but the exact reasons for this discriminatory behavior are not currently known.

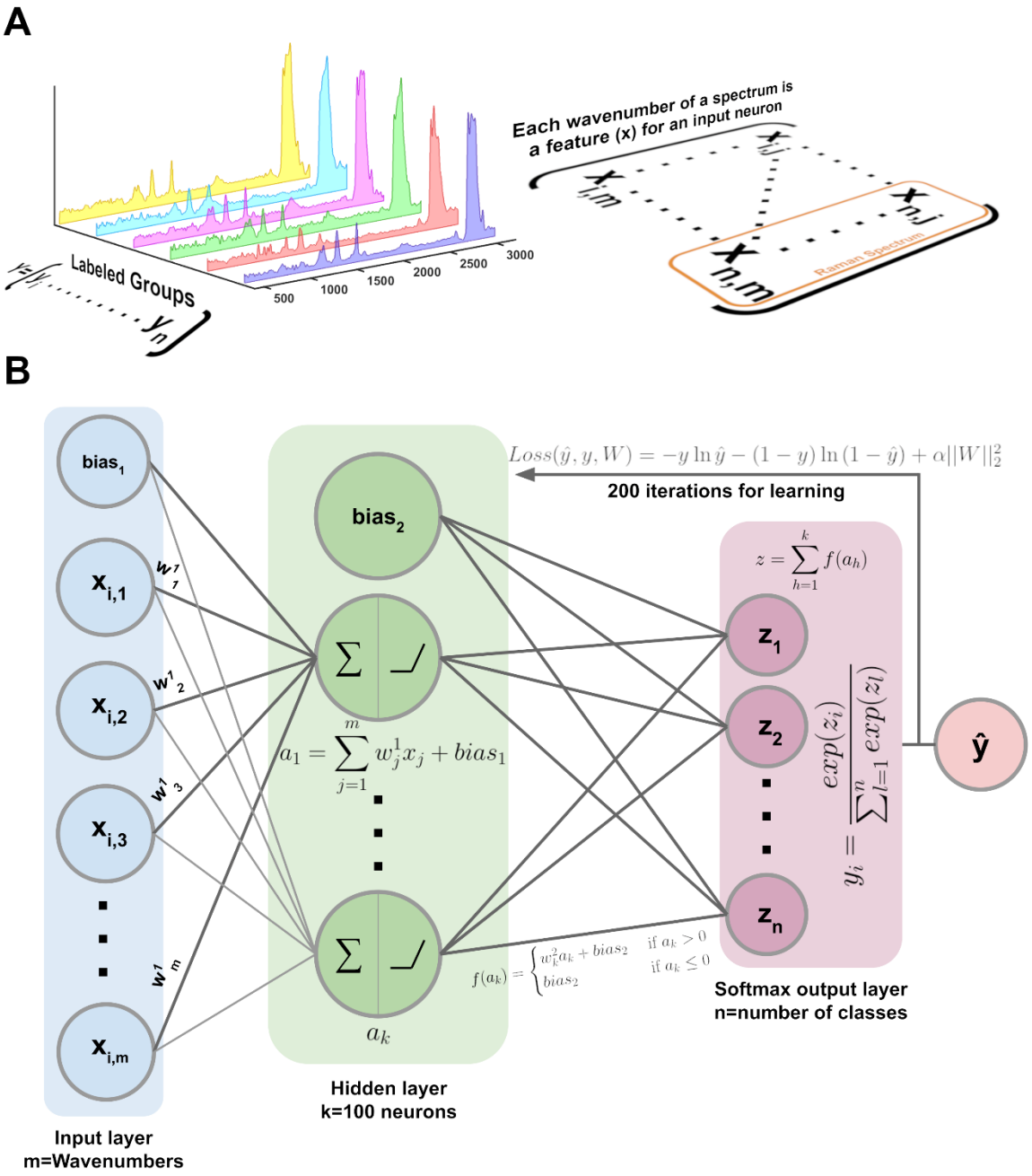


Figure 2.12 Preliminary classification architecture of breast cancer cells
(A) Input for the MLP neural network classification model. All n spectra are labeled as strings which are stored in target vector Y with n categorical variables. All spectra contain the same number of features, so that the final input matrix has size n **x** m corresponding to number of spectra, and wavenumber variables, respectively. **(B)** MLP model for multiclass prediction of breast cancer subtypes and metabolic phenotypes. Raman intensities are sent to the input layer, multiplied by a weight vector of the same size as wavenumbers, and summed in each hidden layer neuron. A ReLU activation function determines what value is sent to the output layer, where the hidden layer neuron outputs are summed for each class. Softmax calculates the predicted target class ŷ based on the which z corresponds to the highest probability. The model learns by minimizing cross-entropy loss in which the gradient $W^{i+1} = W^i - \epsilon \nabla Loss_W^i$.

Confusion Matrix

Confusion matrix for Neural Network (showing number of instances)

		Predicted						
		MBMDA231 D2O	MBMDA231 D2O high Met	MCF10A D2O	MCF10A D2O high Met	MCF7 D2O	MCF7 D2O high Met	Σ
Actual	MBMDA231 D2O	10	1	0	0	0	0	11
	MBMDA231 D2O high Met	4	4	0	0	0	0	8
	MCF10A D2O	0	0	5	1	0	1	7
	MCF10A D2O high Met	0	0	1	5	0	1	7
	MCF7 D2O	0	0	0	0	9	0	9
	MCF7 D2O high Met	0	0	0	0	0	8	8
	Σ	14	5	6	6	9	10	50

Figure 2.13 Simple breast cancer classifier results
Confusion matrix illustrating the 10-fold cross validation accuracy of the ReLU neural network. 100 neurons in the hidden layer and 200 iterations with the Adam solver achieved AUC: 0.9811, CA: 0.82, F1: 0.812, Precision: 0.823, Recall: 0.82. Classification between dietary methionine concentrations in TNBC cell line MDA-MB-231 was the poorest.

Table 2.1 2-way ANOVA of lipid:protein ratio affected by insulin and methionine in TNBC
2-way ANOVA of Fig. 4B shows significant interaction term Insulin*Methionine in TNBC.

ANOVA MDA-MB-231 CH_2:CH_3

	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Insulin	0.023695	2	0.011848	2.841	0.0741
Methionine	0.084655	1	0.084655	20.2999	**0.0001**
Insulin*Methionine	0.034647	2	0.017324	4.1541	**0.0256**
Error	0.12511	30	0.0041702		
Total	0.2681	35			

Table 2.2 2-way ANOVA of de novo lipogenesis affected by insulin and methionine
From Fig. 2.5 A-B, Relative lipid metabolism characterized by the ratio of de novo synthesized and total lipids was not significantly sensitive to methionine concentration in normal-like breast cells, while there is a significant interaction term Insulin*Methionine in TNBC. In both cell types, methionine concentration was the significant independent variable in terms of relative lipid and protein synthesis.

A

	MCF-10A CD_L:CD_P					MCF-10A CD_L:CH_2				
	Sum Sq.	d.f.	Mean Sq.	F	Prob>F	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Insulin	0.008009	2	0.0040045	0.36443	0.69523	0.0010429	2	0.0005215	15.0812	**1.13E-06**
Methionine	0.061128	1	0.061128	5.563	**0.019688**	3.46E-06	1	3.46E-06	0.10001	0.75227
Insulin*Methionine	0.0031567	2	0.0015784	0.14364	0.86633	3.34E-05	2	1.67E-05	0.48336	0.61771
Error	1.5823	144	0.010988			0.004979	144	3.46E-05		
Total	1.6546	149				0.0060588	149			

B

	MDA-MB-231 CD_L:CD_P					MDA-MB-231 CD_L:CH_2				
	Sum Sq.	d.f.	Mean Sq.	F	Prob>F	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Insulin	0.046629	2	0.023315	1.9588	0.14477	0.04666	2	0.0065004	16.2293	**4.40E-07**
Methionine	0.25799	1	0.25799	21.675	**7.27E-06**	0.0031618	1	0.0090534	22.6031	**4.78E-06**
Insulin*Methionine	0.038589	2	0.019295	1.621	0.20128	0.0008114	2	0.0013273	3.3137	**0.039174**
Error	1.714	144	0.011903			0.059362	144	0.0004005		
Total	2.0572	149				0.10999	149			

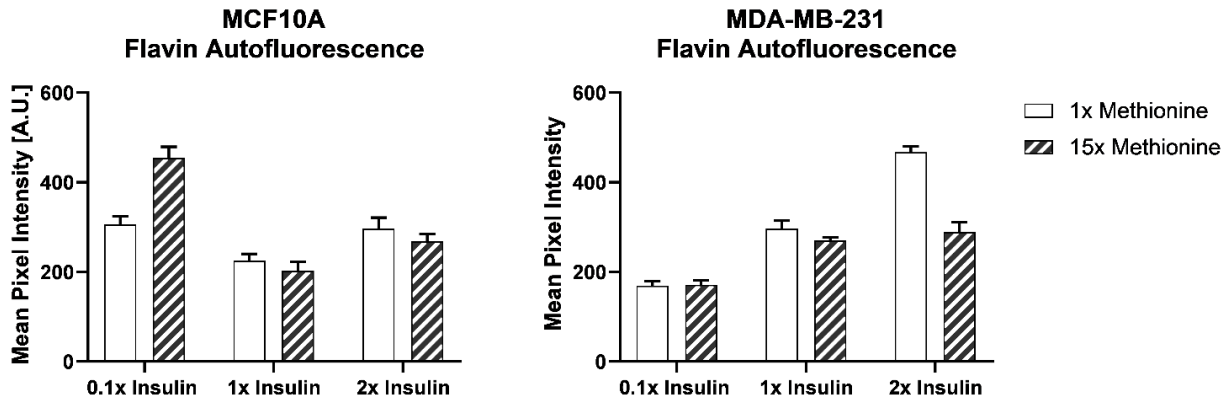


Figure 2.14 Flavin autofluorescence affected by insulin and methionine
Quantitative summary of flavin autofluorescence from TPF image analysis. Excess methionine generally decreased flavin autofluorescence, whereas higher insulin concentrations generally increased flavin autofluorescence in both MCF10A (left) and MDA-MB-231 (right).

Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears in Frontiers in Oncology Volume 12, 2022. Fung, Anthony A.; Hoang, Khang; Zha, Honghao; Chen, Derek; Zhang, Wenxu; Shi, Lingyan, Frontiers, 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 3: Multimolecular Lipid Subtype Detection Using PRM-SRS in Cells and Tissues

3.1 Introduction

Current lipidomic technologies, such as shotgun lipidomics, can quickly identify hundreds of lipids from small samples. Albeit highly sensitive, such methods rely on mass spectrometry (MS), nuclear magnetic resonance (NMR), or other techniques that are destructive to cells and tissues [100–103]. Conventional matrix-assisted laser desorption/ionization (MALDI)-MS imaging enables label-free lipid imaging but it has a lateral resolution on the order of cell diameters (~10 μm) and destroys the sample during the imaging process. In addition, 3D MALDI imaging relies on serial sections of the sample, and the lipid species that are resolvable are limited to those with the highest ion yields. Other optical techniques have been developed to non-destructively visualize spatial distributions of lipid pools [104] as well as metabolic flux [105] at the subcellular resolution, but they rely on markers such as fluorescently labeled antibodies and transfected biosensors, which may alter the native distribution of lipids in cells or tissues. It is difficult to use labeled optical imaging to differentiate diverse molecular species simultaneously, since the diversity of lipid species far exceeds the specificity and availability of optical tags and dyes. Therefore, label-free optical imaging is instrumental. Stimulated Raman scattering (SRS) microscopy has demonstrated advantages of non-destructive 3D imaging with subcellular resolution in a label-free manner [106, 107]. Recent work has even demonstrated quantitative mass concentration measurements of lipids, proteins, and water [108]. For label-free SRS imaging microscopy, multiple subcellular organelles can be the chemical specificity is achieved through hyperspectral imaging (HSI) or training of a deep learning model [109]. Lock-in free multiplex SRS imaging can rapidly extract hundreds of morphological or metabolic features in situ to understand lipid metabolism in cancer cells [110]. Despite these advancements, there has been

no report on distinguishing multiple lipid subtypes in cells and tissue samples by using nondestructive label-free optical imaging methods.

In addition to imaging technologies, post-processing methods/algorithms also contribute to producing high resolution and high-quality images. Recent work on Raman HSI analysis using multivariate curve resolution alternating least squares (MCR-ALS) algorithm has demonstrated effective unmixing of chemical species without disturbing the native distribution of biomolecules [22]. However, a higher spectral resolution may entail prohibitively long imaging time. In addition, unmixing lipid species using unsupervised methods can be computationally expensive and lack the ability to directly identify a chemical species without manual association *posteriori*. For example, the MCR-ALS approach converts a complex spectrum to a linear combination of component spectra, but it can take 30 minutes to process a 512 x 512-pixel hyperspectral image and presupposes the number of chemical species in a sample. The result displays a pixel's identity by its relative proportional composition of reference species. However, this is not always feasible in a complex biological sample. Singular Value Decomposition (SVD) can estimate the number of components, but analytical results may be sensitive to slight deviations from the exact number of components. Clustering and segmentation of image pixels may be informed by MCR-ALS, however, the precise molecular identities of the highlighted pixels may still be unknown, as there is no guarantee that the unmixed components correspond to a specific molecular type.

Spectral reference matching approaches, also known as spectral angle mapping, have been widely applied to Raman spectral analyses by quantifying the spectral similarity between an image pixel spectrum and a known reference spectrum [111]. **Fig. 3.1** shows the general process of reference matching approach applied to hyperspectral imaging. First, spectra of the target analytes (the reference standards of interest) are acquired using spontaneous Raman spectroscopy (**Fig. 3.1 A**) and preprocessed for background removal and normalization (**Fig. 3.1 B,** see Methods for details). Hyperspectral Raman microscopy imaging is next performed on the

sample of interest, in which each pixel contains specific spectral information (**Fig. 3.1 C-D**). Then each pixel's spectrum is preprocessed in the same way as the reference spectrum (**Fig. 3.1 E**) and is analyzed with respect to the reference spectrum by calculating the cosine similarity score (**Fig. 3.1 F**).

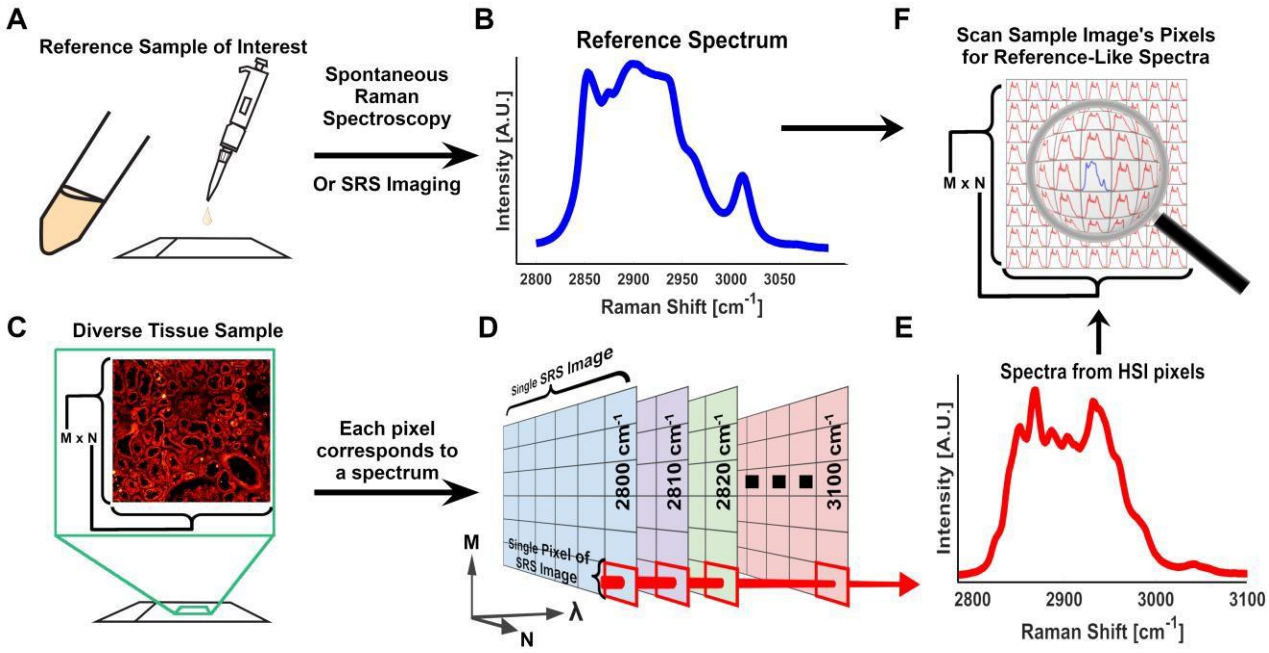


Figure 3.1 General Reference Matching Method
(A) A reference spectrum of a lipid subtype standard is acquired by spontaneous Raman spectroscopy and **(B)** preprocessed to remove background and normalize. **(C-D)** A sample is imaged using SRS to generate a HSI. **(E)** Each pixel of the HSI is a vector of intensity values that represent the Raman spectrum at that pixel. **(F)** Similarity scores are calculated between each pixel and the reference spectra

However, the general spectral reference matching approach has low specificity and the high incidence of false positives makes it difficult to implement in vibrational spectroscopy. This is because the peak position and intensity differences of spectra generated by various equipment can produce uncertainty that overshadows the subtle differences between lipid subtypes. To enhance the specificity for accurately distinguishing lipid species, we developed a Penalized Reference Matching (PRM) algorithm and applied it to SRS (PRM-SRS) microscopy, and accumulated a diverse library of more than 38 biomolecules for potential detection. This method is efficient and can process a 512 pixels x 512 pixels x 76 hyperspectral image stack within one

minute. In this study, we demonstrate broad applications of PRM-SRS on differentiating lipid subtypes and mapping their spatial distributions in cells and tissues. This new method will provide quantitative and qualitative insights into different roles of lipid species in multiple biological processes and can augment other unmixing techniques as well.

3.2 Results

Most images collected in this study were taken from the Raman CH stretching region (2700 cm^{-1} to 3150 cm^{-1}) with 75 total Raman shifts (a spectral distance of 6 cm^{-1} between images). The position deviation Δx was the shift of peaks in the spectrum. We assessed several values for the penalty coefficient and chose $\alpha = 1 \times 10^{-4}$. At a higher value ($4 \times 10^{-4}$) of α, the image contrast was too high to show the full-range signals whereas a lower α (0.25×10^{-4}) caused over-saturation in images (**Fig. 3.2 A-C).** This is because if the penalty is too low, the pixel and reference spectra are free to shift themselves relative to each other until the highest similarity score is returned, no matter how far that shift may be from the original position.

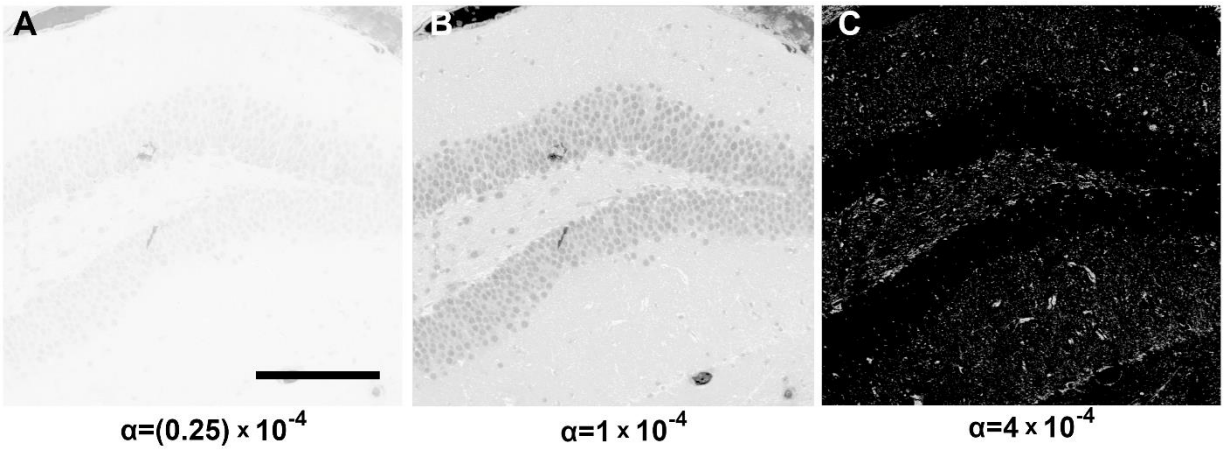


Figure 3.2 Empirical Determination of the Penalty Coefficient
(A) Images of the PRM-SRS result of cardiolipin standard detected in a mouse dentate gyrus SRS HSI. A penalty term that is too low causes over saturated images. **(B)** The empirically chosen penalty coefficient provides a satisfactory contrast. **(C)** A penalty term that is too high prevents the adjustment of spectral position, causing pixels that would otherwise have correspondent spectra to yield an artificially low similarity score. Scale bar 200 microns.

3.2.1 Mapping cholesterol levels in Drosophila fat body using reference spectra

As a proof of concept, we applied the PRM algorithm to detecting and comparing cholesterol levels in fat body tissues of young and old Drosophila. Analogous to mammalian liver and adipose tissue, Drosophila fat body has been used extensively to study lipid metabolism. We collected fat body tissue spectra from young and old flies using spontaneous Raman spectroscopy and compared them to reference spectra of cholesterol at the fingerprint (750 cm^{-1} to 1650 cm^{-1}) and CH-stretching (2700 cm^{-1} to 3150 cm^{-1}) regions (**Fig. 3.3 A-D**). Although the fat body is known to be enriched in triacylglycerides (TAGs), PRM enabled us to extract cholesterol-matched signals in a TAG-rich environment using the cholesterol reference standard. Compared to samples from young flies, fat body samples from old flies showed significantly higher similarity scores to the cholesterol reference spectra in both regions (**Fig. 3.3 E-F**), indicating elevated cholesterol content in old flies. This result is consistent with the published data [112]. This analysis demonstrates our PRM algorithm as an effective method for rapid in situ lipid mapping in tissues.

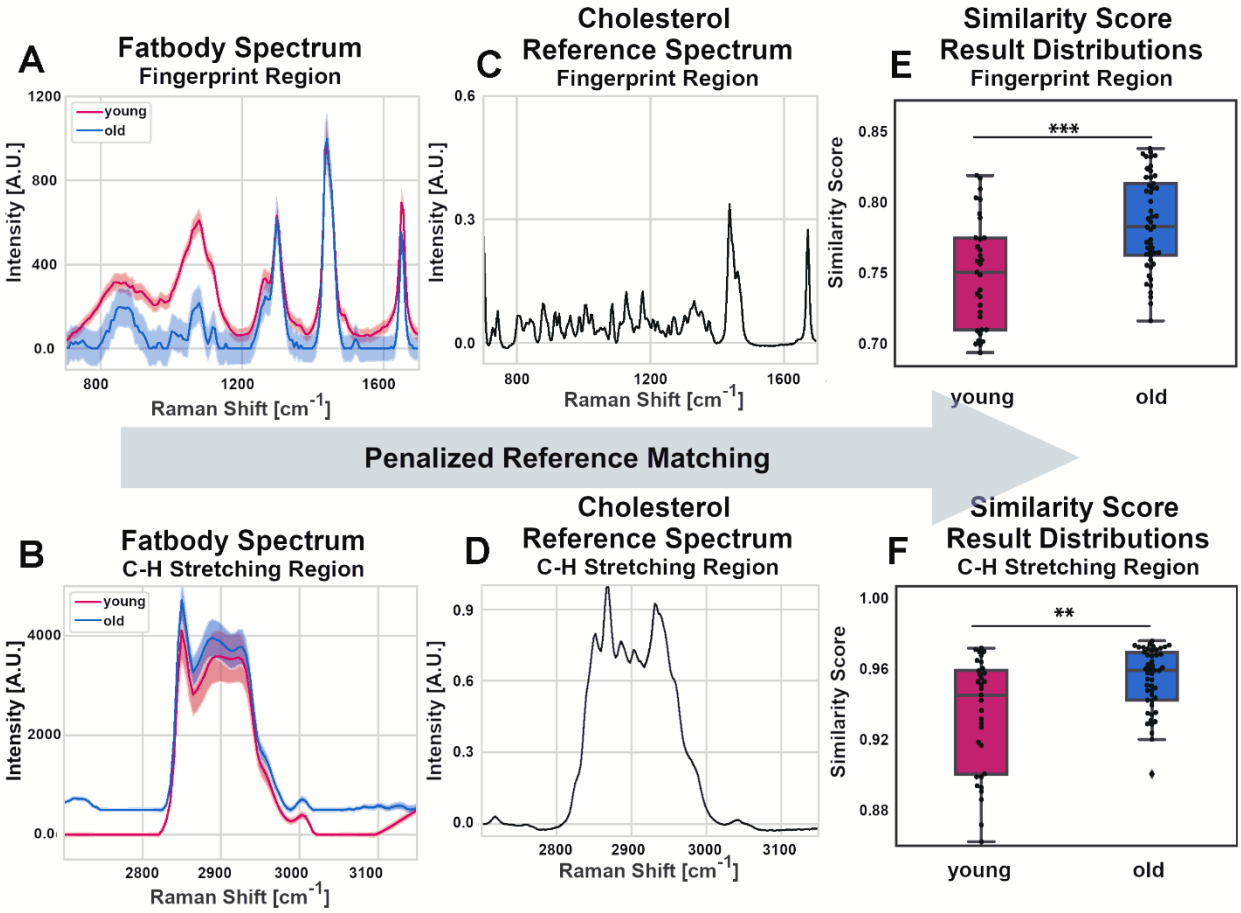


Figure 3.3 Spectral PRM cross correlation in the fingerprint and CH stretching regions
(A, B) Raman signals from fat body tissues of young and old Drosophila in the fingerprint and CH regions. (**C, D)** Raman signals of the cholesterol reference standard in fingerprint and CH stretching regions, respectively. (**E)** Fingerprint region similarity scores of young and old Drosophila fat body samples to the cholesterol reference standard. $p = 4.85 \times 10^{-4}$ by Wilcoxon rank sum test. (**F)** CH-stretch similarity scores of young and old Drosophila fat body samples to the cholesterol reference standard. $p = 0.0037$ by Wilcoxon rank sum test. **, $p<0.01$; ***, $p<0.001$.

Depending on the biological questions to address and Raman scattering equipment available, either the CH-stretching or fingerprint region in a Raman spectrum may be the focus of a study. Both regions can be used to analyze changes in biomolecule distribution, pathological structures (such as amyloid plaques), and other morphological characteristics [2, 113–115]. Although both spectral regions yielded similar results, the fingerprint region generated results with a lower rejection level of $p<0.001$ (**Fig. 3.3**). This is likely because the fingerprint region contained more definitive features, and the CH-stretching region possessed low intensity shoulders below 2800 cm^{-1} and above 3000 cm^{-1}, which may lead to a higher similarity score between samples

since both spectral data sets matched in those regions where the intensity was zero. Importantly, this demonstrates that similarity scores generated from spectra data by PRM-SRS can be used to estimate the levels of biomolecules in the samples.

3.2.2 Using PRM-SRS to detect cardiolipin changes in cells

After validating the efficacy and robustness of PRM on spontaneous Raman spectral analyses, we next extended the algorithm to the analysis of Stimulated Raman scattering (SRS) images. To evaluate the spatial accuracy and quantitative approximation of PRM-SRS, we first benchmarked it against fluorescence microscopy images. Using PRM-SRS imaging, we examined cardiolipin (CL), an essential phospholipid in the inner mitochondrial membrane, in cultured HEK293 cells. CL is synthesized in the inner mitochondrial membrane in consecutive reactions catalyzed by enzymes including phosphatidylglycerophosphate synthase 1 (PGS1), phosphatidylglycerophosphate phosphatase (PTPMT1) and cardiolipin synthase (CLS1) [116, 117]. PGS1 is essential for CL synthesis, and expression of an enzyme-deficient mutant PGS1 leads to a reduction of PGP (Phosphatidylglycerophosphate) and CL in CHO cells [118]. We generated stable HEK293 cell lines with downregulated PGS1 (shPGS1). PGS1 down-regulation was confirmed by immunofluorescence analysis using a PGS1-specific antibody (**Supplementary Fig. 3.9**). Following staining using nonyl acridine orange (NAO), a fluorescent dye with high affinity for CL [119], cells were analyzed using both two-photon fluorescence (TPF) microscopy and PRM-SRS. To demonstrate the specificity of SRS signals for CL, we compared control HEK293 cells with shPGS1 cells. PRM-SRS analysis of the hyperspectral images was consistent with TPF images in both control and shPGS1 cells (**Fig. 3.4 A-B**). Quantitative analyses of both PRM-SRS images and fluorescence images showed significant decreases of CL signals in shPGS1 cells compared with control cells (**Fig. 3.4 C-D**). Importantly, the similarity score image of the reference-matched CL is distinct from any single Raman shift images in the

CH symmetric stretching regions (**Supplementary Fig. 3.9 C**). These results demonstrate the ability of PRM-SRS to quantitatively detect CL changes in cells, and its potential for visualizing protein/lipid metabolic dynamics at the subcellular scale.

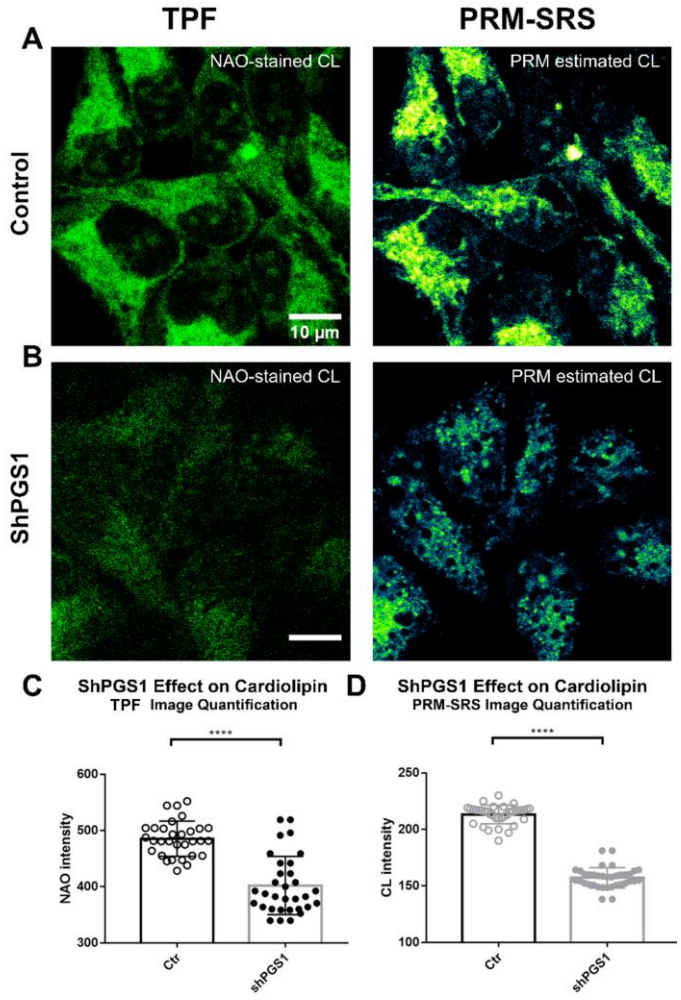


Figure 3.4 PRM-SRS and fluorescence staining show similar results
(A, B) Comparison of PRM-SRS and fluorescence staining in control cells expressing shCtr (Ctr; **A**) and PGS1 knockdown (shPGS1; **B**) HEK293 cells. Panels on the left, two photon fluorescence microscopy (TPF) images following nonyl acridine orange (NAO)-labeling of CL. Panels on the right, label-free SRS hyperspectral images of CL at the CH-stretching region. Quantitative analyses of NAO staining signal intensity (**C**) and PRM-SRS imaging signal intensity (similarity score in eight-bit integer image) (**D)** of CL in control and shPGS1 cells. Significantly decreased signals in shPGS1 cells were detected by both TPF and PRM-SRS microscopy. Data are presented as mean ± SEM and analyzed by One-way ANOVA with Bonferroni post hoc test; ****p<0.0001. Scale bar, 10 μm.

3.2.3 PRM-SRS Visualizes Clinically Relevant Lipid Subtypes in Human Kidney

We then applied PRM-SRS to characterizing lipid subtypes in human kidney tissues, a structurally and functionally highly complex tissue composed of more than 50 cell types [120]. Cholesterol, ceramides (Cer), and triacylglycerides are among the most abundant lipid species in the kidney. Dyslipidemia is frequently observed in nephrotic syndrome (NS) and various types of chronic kidney disease (CKD) [121]. The glomerulus, the filtration unit of the nephron, is a network of capillaries that sequesters lipid species as an initial step of filtration and is decorated with lipid droplets. Wrapping around the capillary of the glomerular tuft are podocytes, making up the epithelial lining of the Bowman's capsule. We used healthy human kidney tissue sections as a control to showcase the application of our PRM-SRS in imaging different lipid subtypes in structurally complex tissue samples.

SRS imaging detected the overall distribution of lipids in the morphologically distinct structures in these kidney tissues, such as glomeruli, tubules, and blood vessels (**Fig. 3.5 A**). Using PRM-SRS, we estimated relative concentrations of lipids in different structures, such as lipid droplets in podocytes, and eosinophilic bodies near tubules (**Figs. 3.5 A-B).** PRM-SRS imaging revealed distributions of distinct lipid subtypes in the glomerulus and surrounding structures in situ, including TAG, cholesterol, cholesterol ester, and C12 ceramide, with the 90^{th} percentile similarity scores to the corresponding pure lipid reference spectra (**Fig. 3.5 C**).

Dyslipidemia is manifested as elevated levels of serum TAGs, cholesterol, and very low to intermediate density lipoproteins. Common initial abnormalities include decreased production and activity of lecithin-cholesterol acyltransferase which decreases high-density lipoprotein (HDL) levels and maturation of HDL cholesterol [122]. The regulation of HDL cholesterol is tightly controlled by several organs, but generally entails the esterification of cholesterol into cholesterol esters, which move towards the center of HDL particles, along with neutral TAGs. This maintains a favorable cholesterol gradient as these HDL particles become enriched by sequestering cholesterols and fatty acids from other lipoproteins. Although mature lipoproteins are too large to pass the glomerular filtration barrier, lipids and lipid-bound proteins from lipoproteins may affect

overall renal lipid metabolism [123]. Our ratiometric imaging revealed that there is a greater amount of non-esterified cholesterol in the lipid particles than neighboring structures. These pools of cholesterol may represent those yet to be enriched or ectopic deposits (**supplementary Fig. 3.10**). Ceramides are also abundant in the kidney and play a crucial role in regulating cellular processes and binding cholesterol and other lipoproteins [124]. Ceramides, e.g., C12 ceramide, show high similarity with pixel spectra in lipid droplets and lipoprotein particles (**Fig. 3.5 C**). In nephropathies, ectopic lipid deposits in the glomerular mesangium and proximal tubules are typically concurrent with low HDL levels [122]. Other characteristics of glomerular nephropathies are depicted in **Fig. 3.5 D**. The ability of PRM-SRS to track the lipidomic profile in tissues collected from patients at various stages of diseases will generate critical data for changes in these macromolecules over time, and with associated biological variables. Such studies will provide insights into assessing severity, progression or prognosis of various lipid metabolic diseases.

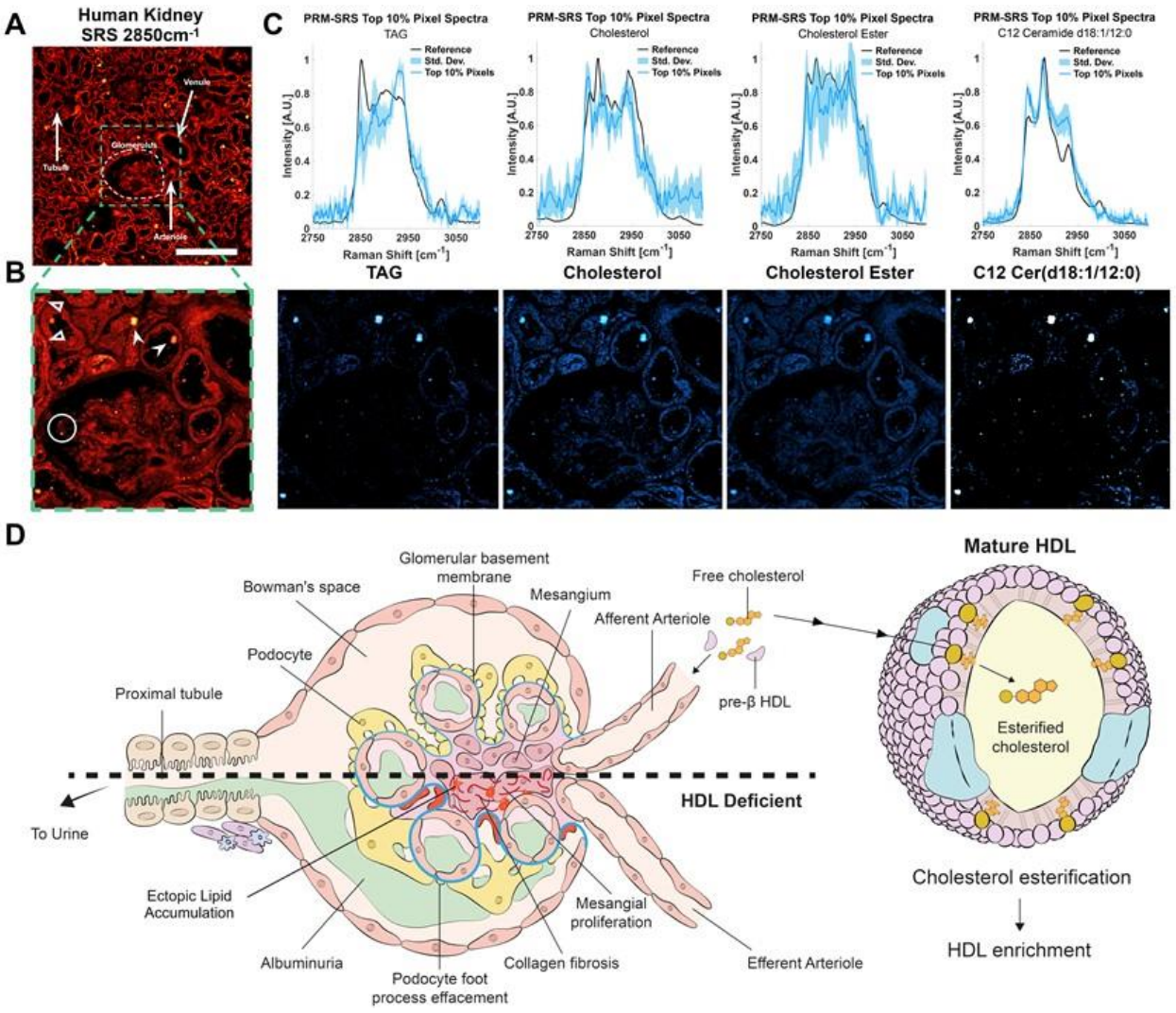


Figure 3.5 PRM-SRS Detection of Lipids in Kidney Biopsy
(A-B) SRS image of a human kidney tissue section at 2850 cm^{-1}. Panel B shows the enlarged image of the boxed area in panel A. Hollow arrowheads, intracellular lipid droplets in tubules. Solid arrowheads, eosinophilic bodies. Circles, lipid droplets sequestered by podocytes in the glomerulus. Scale bar, 200μm. **(C)** PRM-SRS spectra (top panels) and images (bottom panels) of different lipid subtypes of interest show the distribution of the similarity scores, each with the same contrast levels. Spectra of the top 10% of similarity score pixels overlaid on the reference spectrum for each lipid subtype show consistent matches. Resulting similarity score images were background subtracted to improve the contrast. **(D)** Schematic diagram of glomerular pathologies associated with dyslipidemia in kidney diseases. (Adapted with permission from [125]). Scale bar, 200 μm.

3.2.4 Mapping lipid subtype distributions in Drosophila fat body

Using maximum intensity projection (MIP) of the PRM-SRS hyperspectral image of total lipids, we visualized lipid droplets in Drosophila fat body cells (**Fig. 3.6 A**). We also detected lipid subtypes using different lipid reference standards, including TAG and phosphatidylethanolamine

(PE). In addition to detecting lipid subtypes, PRM-SRS can also provide information on subcellular distribution, including co-localization, of different lipid subtypes. Comparison of MIP of the PRM-SRS hyperspectral image of total lipids (**Fig. 3.6 A**) with mono-unsaturated triacylglycerol (TAG 18:1) reference-matched image (**Fig. 3.6 B**) revealed abundant TAG in lipid droplets (**Fig. 3.6 A, B**). A critical tenet of unmixing techniques such as PRM is that spectral shape, not intensity, is what drives the similarity score of normalized spectra. The MIP in **Fig 3.6 A** shows several lipid droplets with non-uniform maximum intensities, yet the TAG reference matched image shows a more uniform result. This was because despite intensity differences that may have arisen from the sample focus plane, the spectral shapes were still consistent.

Drosophila fatbody cells contain lysosome-like structures that regulate their lipid anabolism and were detected using reference spectra of lysosome-like structures of fatbody (**Fig. 3.6 C**). These spectra, unlike the lipid subtypes, have a more dominant CH_3 stretching peak at 2935 cm-^{1}, and a more pronounced olefinic peak at 3065 cm^{-1}. PE is one of the most prominent lipid subtypes in cell/organelle membranes and can be visualized by taking PRM-SRS images using their corresponding reference standard. **Fig. 3.6 D** shows the spatial distribution of the ratio of the PE and TAG similarity scored images. Upon closer inspection, **Fig. 3.6 F** shows the spectra of the apparent pixels are similar to TAG, and therefore appear darker in those regions in both the TAG and PE similarity scored images. However, the lipid cores have a greater disparity in these similarity scores, with an even greater similarity to TAG and lesser similarity to PE. Therefore, the apparent pixels are visible because of relative concentrations, as discussed in supplementary**.** The intensity profiles (**Fig. 3.6 G**) support the notion that signal intensity of images based on similarity scores varies by spectral shape, whereas signal intensity in SRS images varies by chemical bond concentrations. Thus, a lipid droplet core may appear more uniform in a TAG reference-matched PRM-SRS image than in an original SRS image of the CH_2 stretching region. Together, these data demonstrate that PRM-SRS is useful in detecting different lipid subtypes and their distribution at the subcellular scale.

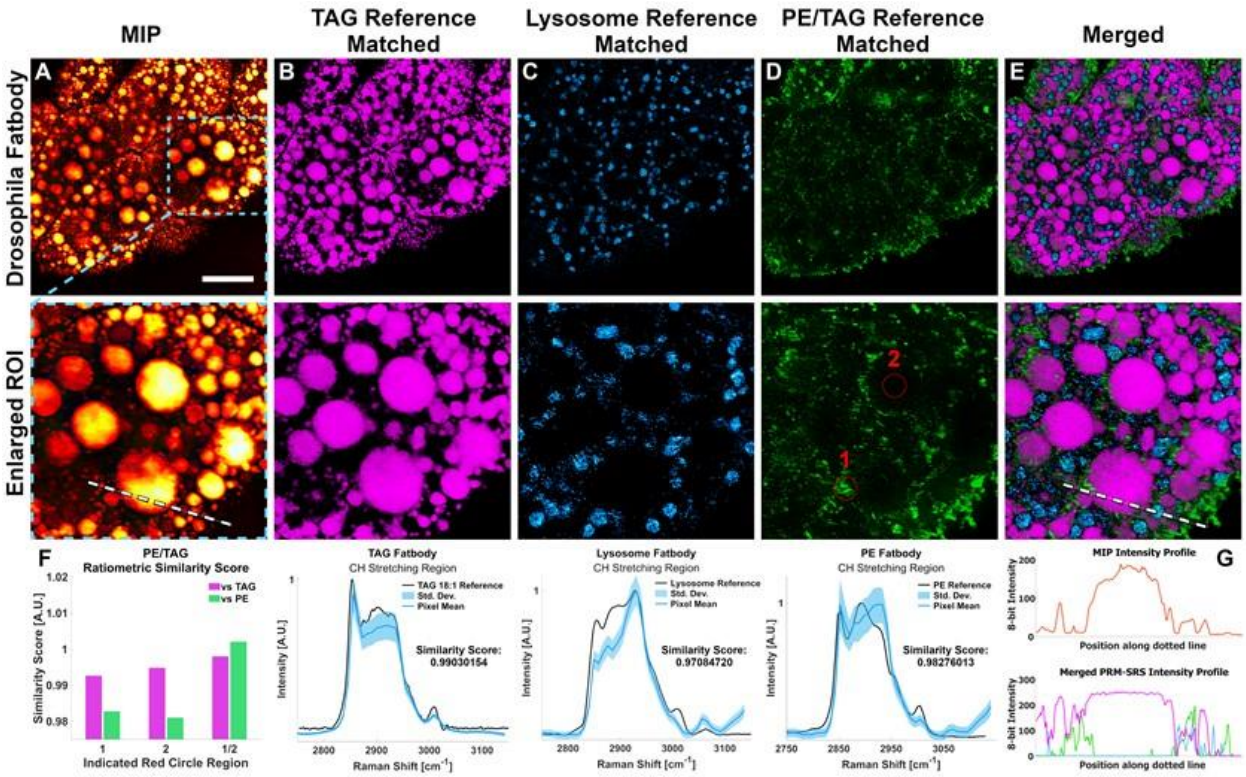


Figure 3.6 PRM-SRS Detection of Lipid Subtypes in Drosophila Fatbody
(**A**) Maximum intensity projection (MIP) of the PRM-SRS hyperspectral image of total lipids reveals lipid droplets. (**B**) PRM-SRS detected TAG in lipid droplet cores. **(C)** Lysosome-like structures detected by PRM-SRS using reference spectra measured from lysosome-like structures in drosophila fatbody. **(D)** PE:TAG ratiometric images show that the interstitium between lipid droplet cores and lysosome-like structures has higher relative levels of PE. **(E)** PRM-SRS subtype images are merged to detect co-localization of different lipid subtypes. **(F)** Similarity scores from areas marked by red circles 1 and 2 in the lower part of panel D highlight the necessity of evaluating relative concentrations as opposed to absolute concentrations. **(G)** Intensity profiles along the dotted white lines in (**A**) and (**E**), upper and lower panels respectively, show how signal intensity varies with spectral shape, rather than concentration in a PRM-SRS image. Scale bar is 20 μm.

3.2.5 Analyzing lipid subtypes in mouse brain samples

We next applied PRM-SRS to analyzing lipid metabolism in the context of the aging using mouse hippocampal samples. We visualized and compared cholesterol, PC, and PE levels in hippocampal samples from young (3 months) and old (18 months) mice (**Figs. 3.7 A-C; F-G**). We also generated ratiometric images for quantitative analysis, since the signal intensity has a linear relationship with the concentration of chemical bonds of the molecules detected. Ratiometric imaging analyses showed increased Cholesterol/PE ratio in subregions of granule cell nuclei (**Fig. 3.7 D, I**; red circles). This increase in the Cholesterol/PE ratio was more prominent and detected

in more granule cells in the old brain samples as compared with the young ones (compare **Fig. 3.7 D** with **7I**), indicative of altered cholesterol and/or PE metabolism in the old brains. These results show that ratiometric PRM-SRS imaging can detect changes in differential spatial distribution of various lipid subtypes even when such changes are not obvious in images of individual lipid subtypes.

Ratiometric images of PC/PE showed higher levels of PC relative to PE in the granule cell nuclei of the dentate gyrus in both young and old mice, but lower levels outside the nuclear regions (**Fig. 3.7 E, J**). Compared to the young brain sample, the old brain sample showed no significant changes in the average PC or PE levels in the granule cells in both individual imaging channels (**Fig. 3.7 B, C, G, H**) and ratiometric images (**Fig. 3.7 E, J**). This is consistent with the results from Gas Chromatography Mass Spectrometry (GC-MS) (see **Fig. 3.7 K, L**). However, we noticed spatial distribution differences in the PC to PE ratio between young and old samples. The ratiometric images reveal that more granule cell nuclei had uniformly higher PC/PE ratio in the old brain sample, whereas the nuclei in the young sample showed less even distribution of the PC/PE ratio (red areas; see those nuclei marked by purple arrows) (**Fig. 3.7 E, J**). These data suggest altered synthesis, accumulation or clearance of PC and/or PE in the granule cells in the old brains, consistent with a previous report [126]. Since PE is a precursor of PC, higher PC to PE ratios inside the older hippocampal granule cells suggest that aging brains may have altered CTP:phosphocholinecytidylyl transferase (CCT) activity – a rate limiting PC synthesis enzyme with a predominantly nuclear localization [127]. This finding is significant because both PRM-SRS imaging and GC-MS analysis show that the younger brain samples contain less cholesterol relative to PE than old ones. However, only through ratiometric analysis were we able to detect differential subcellular distribution of lipids, including cholesterol/PE and PC/PE ratio in the nuclei (**Fig. 3.7 D, I**, **E, J**).

For comparison, we analyzed the same samples using GC-MS to quantify cholesterol/PE and PC/PE ratios (**Fig. 3.7 K**), which demonstrated an increased cholesterol/PE ratio and no

changes in the PC/PE ratio in in old brain samples compared with young ones. The PRM-SRS images of nuclei in the tissues were manually segmented using ImageJ for quantification of cholesterol/PE and PC/PE ratios (**Fig. 3.7 L**). These data suggest that PRM-SRS may be used for quantitative lipidomic imaging analyses in tissue samples in the future.

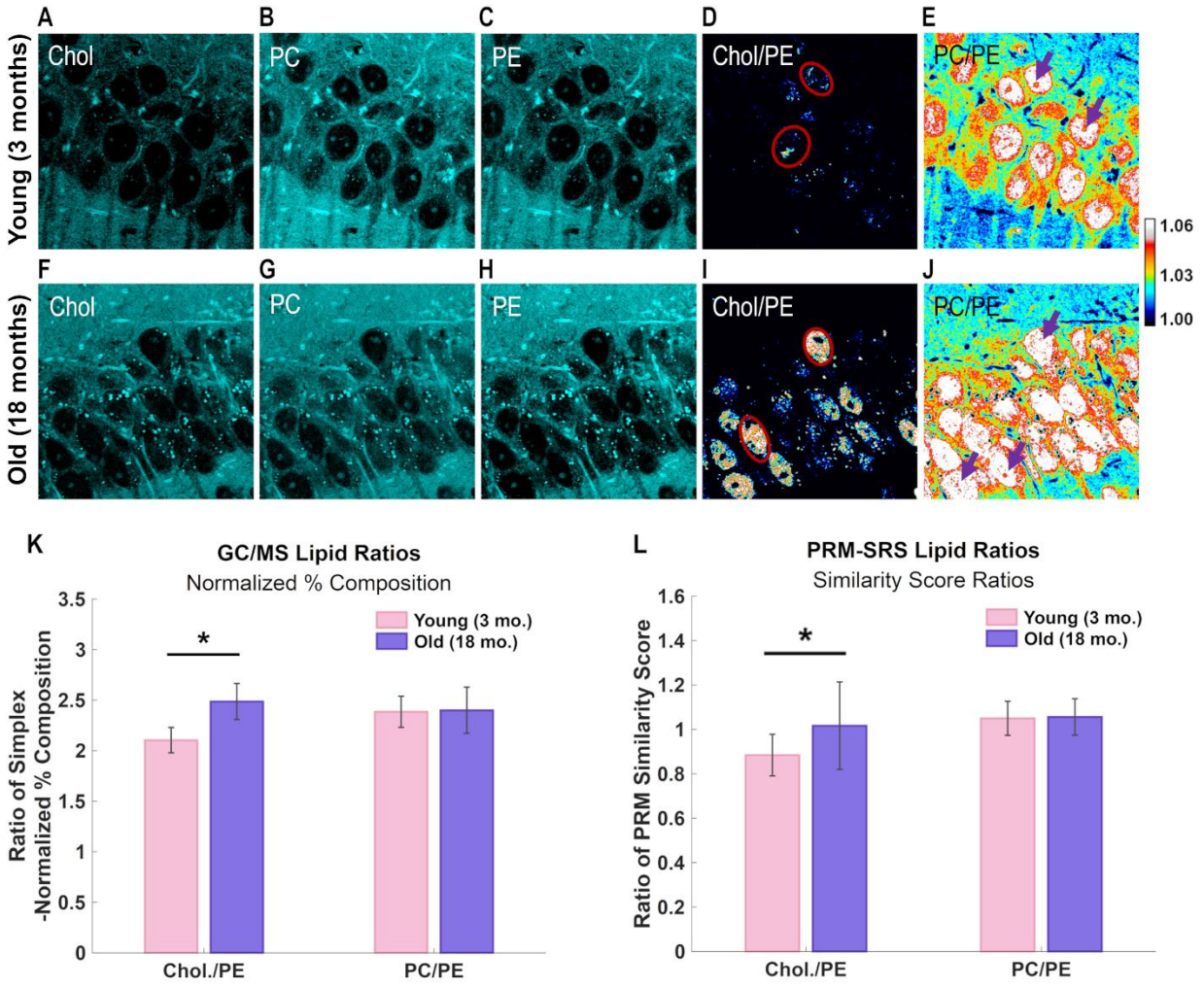


Figure 3.7 PRM-SRS imaging of mouse hippocampus
(A-J) PRM-SRS hyperspectral detection of cholesterol, PC, and PE in hippocampus samples from young and old mice. Overall intensity of detected lipid subtypes shows distinct patterns, with old brains showing higher cholesterol to PE ratio, but relatively consistent levels of PE and PC. **(D, I)** Ratiometric images of cholesterol to PE shows more nuclei with higher cholesterol/PE ratio in the old brains. Selected nuclei are marked by red circles **(E, J)** Ratiometric images of PC relative to PE show higher PC/PE ratio in granule cell nuclei of both young and old brains, but the spatial distribution of the ratio is more heterogeneous in young samples (see nuclei marked by purple arrows). Scale bar, 20μm. **(K, L)** Mass spectrometry (K) shows results consistent with that obtained by ratiometric PRM-SRS imaging (**L**). Summed concentrations of lipids were simplex normalized and displayed in the ratio form. **(L)** Ratiometric image intensities, corresponding to the ratio of PRM similarity scores of lipid subtypes, are plotted. Error bars represent standard deviation.

3.2.6 Detecting lipid subtype distributions in human brain tissues

Sphingosine is another crucial lipid subtype whose metabolic alteration has been suggested as a biomarker for neurodegenerative diseases, such as Alzheimer's, Parkinson's, and Huntington's diseases [103, 128]. To visualize individual cells, we used label-free optical SRS histology (SRH) imaging of human brain sample (**supplementary Fig. 3.11**) to create virtual histology images similar to that of hematoxylin-and-eosin (H&E) staining as previously reported [129]. Using PRM-SRS, we visualized sphingosine and CL simultaneously in the human brain tissue sections (**Fig. 3.8 A-B**). Superimposition of sphingosine and CL images illustrates their relative distribution in brain cells (**Fig. 3.8 C-D**). Ratiometric imaging (**Fig. 3.8 E**) and quantitative analyses (**Fig. 3.8 F**) demonstrated a clear reduction in the CL to sphingosine ratio inside the nucleus, consistent with the fact that CL is mainly localized at the inner mitochondrial membrane but not in the nucleus. These results show that PRM-SRS can be used to visualize the subcellular distribution of different lipid subtypes.

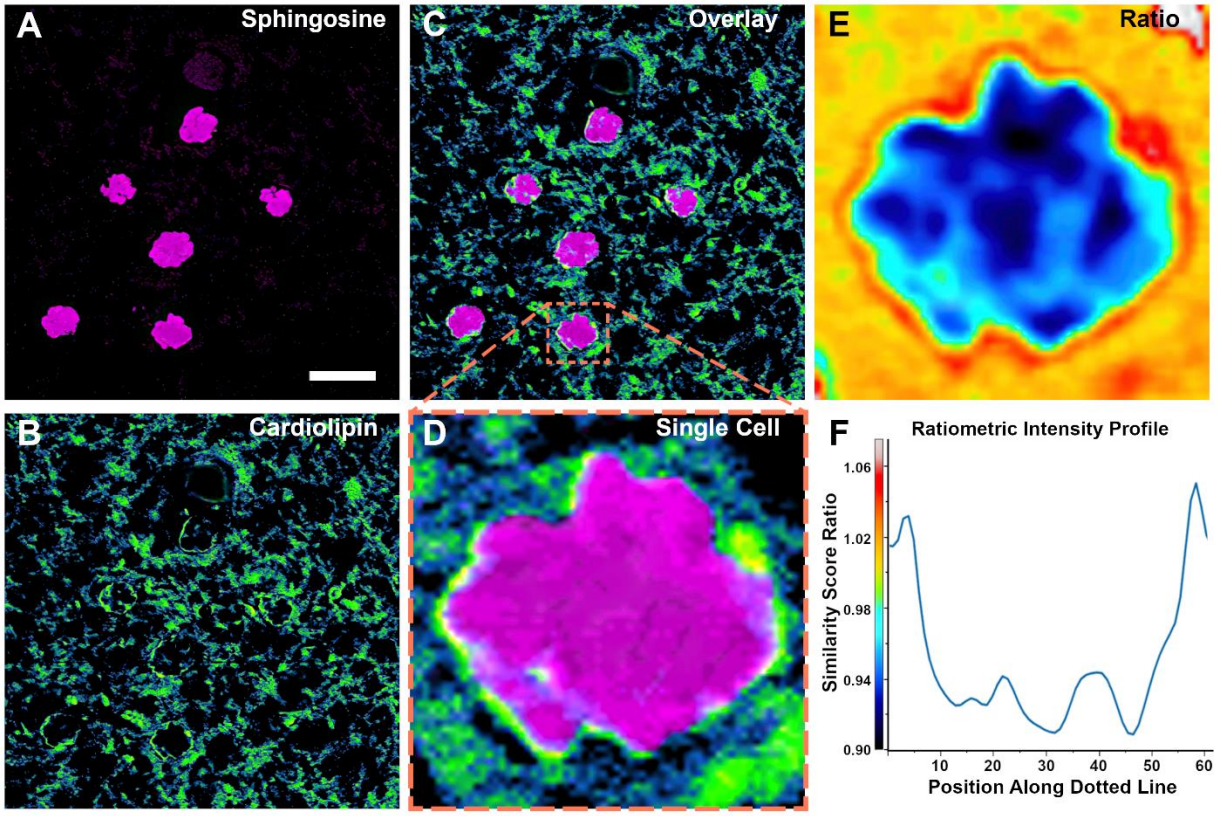


Figure 3.8 PRM-SRS Detection of Cardiolipin and Sphingosine in Human Brain Biopsy
(A) Sphingosine signals in the brain tissue; **(B)** CL in the same region of interest; **(C)** merged image of CL and Sphingosine; **(D)** Zoomed-in image of a single brain cell with CL and Sphingosine signals. **(E)** Ratiometric image of CL to Sphingosine signals. **(F)** Intensity profile of **(E)** along the indicated white dashed line. Scale bar, 10 μm

3.3 Discussion

In this study, we developed a PRM algorithm that can efficiently unmix and distinguish a variety of lipid subtypes from single SRS HSI stacks. Compared with fluorescence imaging, our PRM-SRS platform shows advantages of multiplexed lipid subtype visualization from single label-free HSI sets. This also represents a significant expansion in applications compared with traditional SRS imaging, which often relies on detecting total lipids in the CH-stretching mode at 2850 cm^{-1}.

With an improved contrast, PRM-SRS imaging enables us to identify different lipid subtypes. The spectra of lipid subtype-standards collected from spontaneous Raman scattering

microscopy can be utilized in analyzing HSI data collected from SRS microscopy. Our PRM-SRS can generate both co-localization and ratiometric data of individual lipid subtypes simultaneously by mapping their spatial distributions and quantifying their relative concentrations. In this study, we established a Raman spectra library with 38 lipid subtype standards and demonstrated the simultaneous detection of a few selected lipid subtypes by PRM-SRS in cells and tissues. Analyses of human kidney tissue samples indicate that PRM-SRS can be used to identify different lipid subtypes associated with renal diseases, suggesting potential application of PRM-SRS in diagnosis and prognosis of these diseases, including those associated with dyslipidemia. Such label-free methods may be instrumental in early detection of kidney diseases by detecting and measuring relative levels of different lipid biomarkers without the need to stain biopsied samples or perform destructive imaging, especially on limited clinical samples. Analyses of Drosophila fat body samples show that PRM-SRS can be used effectively in mapping spatial distributions of lipid subtypes at the subcellular scale. These results highlight the ability of PRM-SRS to selectively visualize multiple lipid subtypes in a single image with ease and freedom akin to individual-subtype labeled imaging without the need to actually label them. Analyses of mouse and human brain tissues demonstrate the importance of measuring relative lipid concentrations through ratiometric imaging, which reveals regionally different concentrations of lipid subtypes that may not be readily apparent in single-channel images. Although lipid subtypes are not measured in absolute concentrations, their relative levels are consistent with results from other modalities such as mass spectrometry (MS).

The brain is a lipid-rich organ. Lipid subtypes such as cholesterol and sphingolipids are important components of the brain. Alteration in lipid subcellular distribution and metabolism impact on brain cell function have been associated with neurological diseases. Our analyses of mouse and human brain tissues illustrate the capability of PRM-SRS in quantitatively mapping and analyzing distribution of different lipid subtypes within single cells. These analyses confirm the cross-applicability of the fingerprint and CH-stretching spectral regions for quantitative

analyses. Further, our PRM-SRS imaging shows that sphingosine, a catabolic product of sphingomyelin, has a predominantly nuclear localization. Nuclear sphingomyelinase and sphingosine kinases regulate the release of ceramides and sphingosine, as well as the conversion of sphingosine to sphingosine-1-phosphate. These processes regulate cell proliferation and cell death [130]. Sphingosine kinases may shift from a cytosolic to a nuclear localization in brain samples from Alzheimer's disease patients [131]. Development of new technologies in imaging distinct lipid subtypes and their metabolism will enhance our ability to investigate molecular mechanisms underlying different brain disorders.

PRM-SRS has sufficient capability to provide quantitative information on lipid subtype distribution. Theoretically, the similarity scores and concentrations of dominant molecules have a linear relationship in a certain dynamic concentration range. However, spectral shifts caused by multiple interfering factors (such as changes chemical environment or instrumentation) may distort the relationship between the similarity scores and concentrations. Since the spectral shifts can be caused by multiple factors, it is difficult to define an exact function of a spectral shift. Nevertheless, such spectral shifts can be corrected by using the penalized regression analysis. In brief, the penalty term helps us calculate the similarity score to more precisely describe the linear relationship with concentration. Depending on the equipment used and the samples of interest, careful tuning of the penalty term in the PRM algorithm is necessary. At present, the PRM-SRS platform should be used in well controlled experiments to limit external chemometric dimensions. In this way, spectral signals are more likely from molecular subtypes in the samples, rather than from noises. Since Raman peak intensities are multiplexed in the sense that a specific peak shape may be influenced by multiple molecules, it is critical that molecular makeup is as consistent as possible when using PRM-SRS to estimate relative concentrations of different molecular subtypes. In the current PRM platform, users should determine the optimal penalty coefficient experimentally to avoid artificial increases or decreases of similarity scores. An extremely low penalty coefficient would allow the comparison between reference spectra and pixel

spectra to occur at any offset, which may inflate the overall similarity score. For example, ceramide has a notably high peak at 2880 cm^{-1}, while pixel spectra typically have the most prominent peak at the 2935 cm^{-1} area. Allowing spectral offsets without penalty may result in a high similarity score because the ceramide reference could yield a high similarity score with the pixels with a sharp 2935 cm^{-1} peak. On the other hand, an extremely high penalty coefficient would be akin to not allowing spectral offsets during comparison at all, which would be similar to traditional reference matching. This is disadvantageous because spectra acquired with different equipment may not be exactly calibrated on the same x axis, which could artificially decrease the similarity scores. While these cases do not occur often, the greatest similarity score does occur at very small offsets. Fig. 3.2 shows PRM-SRS images with different penalty coefficients. Although the PRM-SRS pipeline can be further enhanced by including more comprehensive reference standards and further increasing analysis speed, this platform is robust for analyzing different lipid subtypes. Reference matching could also be a useful tool to detect the presence of representative mixtures of compounds, not simply individual molecules.

The main advantages of PRM-SRS include multiplexed molecular subtype visualization, positive values and fast speed of similarity score calculation. Similarity scores are always positive values since Raman intensities cannot be negative. On the other hand, pseudo-inverse matrix (PINV) coefficients can be negative, which will make it difficult to normalize the output. Our similarity score calculation is faster than other methods, such as the pseudo-inverse matrix (**Fig. 1.9**). PRM and PINV show similar results in correlation with relative concentrations. However, PINV-based calculation time increases exponentially with the number of spectra in the original matrix and image size. When analyzing 1024x1024 hyperspectral images using PINV, there are millions of spectra in a single experiment. On the other hand, similarity score calculation using PRM-SRS is based on the inner-product, which is easily vectorized and split in a parallel pool. The calculation time difference between PRM and PINV is shown in Figure 1.9. Considering the number of spectra in a single HSI stack, the calculation speed is an important factor as a practical

analysis tool for image analysis. Although the PRM-SRS result cannot provide absolute concentrations as precisely calibrated linear unmixing methods, this approach clearly shows advantages over other methods.

PRM-SRS can also be used together with other chemometric methods, such as GC-MS for cross validation, as the incidence of false positive may still be high. Finally, detection of the vast variety of lipid subtypes may require further improvement in unmixing methods and spectral resolution, as the lipid subtype reference library is expanded, and more lipid subtypes are further evaluated. With some adjustments, such as using different reference libraries, this PRM-SRS platform can be extended to analyzing other molecules, including proteins, nucleic acids, and even clinically relevant molecular complexes (such as protein aggregates or oligomers).Using heavy water (D_2O) probed SRS (DO-SRS), metabolic imaging can also distinguish de novo newly synthesized biomolecules, including lipids, proteins, and DNA [17, 129, 132], from old existing biomolecules in cells and tissues at the subcellular resolution. This ever-expanding library of molecular subtype references may warrant broader spectral regions, including the fingerprint, CH-stretching, and O-H stretching regions to increase the chemometric dimensions. Integration of statistical denoising and regression methods will help increase the power of molecular subtype matching. Application of higher order signal manipulations such as digital derivatives and wavelet analyses, will enhance the ability to extract the most prominent as well as subtle but important features.

In summary, this study presents a new hyperspectral imaging platform - PRM-SRS - that allows for direct-identification of multiple molecular species in situ with subcellular resolution and high chemical specificity by leveraging the cross-applicability of spectral reference libraries and HSI methods. This PRM-based method can be applied to various microscopy setups, such as SRS, FTIR, and spontaneous Raman scattering spectroscopy. Compared with existing HSI methods, PRM-SRS shows a much-enhanced speed and efficiency. With appropriate reference spectra established, PRM-SRS can be used to detect a wide range of different biomolecules. This

new platform can also be applied to studying metabolism of diverse types of biomolecules in cell and tissue samples, which will be highly useful for investigating metabolic changes under different pathophysiological conditions. With its easy implementation, PRM-SRS can be combined with high-throughput methods, such as microfluidic/nanofluidic devices and single-cell apparatuses, or with large-area HSI mapping methods. The application of deep learning algorithms, such as DeepChem, may further improve the imaging speed in femtosecond SRS imaging [109]. PRM-SRS could benefit from additional instrumentation improvements, such as distortion-free polygon scanning and spectral focusing, as well as from machine learning to enhance the SNR [133]. Finally, PRM could easily augment other unmixing methods, including MCR-ALS by providing fast initial component spectra. Thus, PRM-SRS has great potentials in multiplex cell and tissue imaging with a broad spectrum of applications.

3.4 Materials and Methods

3.4.1 HEK293 Cell Cultures

The parental HEK293 cell line was obtained from the American Type Culture Collection (ATCC). Cells were cultured in DMEM supplemented with 5% fetal bovine serum (FBS), 1% penicillin/streptomycin (Fisher Scientific, Waltham, MA).

The control shRNA construct was as previously described [134]. A shPGS1 construct was designed PGS1 by expressing shRNA against PGS1 (target sequence: 5'-TCGGGTTCCATCCGTTTAAAT-3') in the plasmid vector Tet-pLKO-puro (Vector Builder Inc.) to specifically down-regulate expression of PGS1. The control and shPGS1 constructs were transfected into HEK293 cells using lipofectamine (Invitrogen). Following transfection, cells were selected using puromycin (1ug/ml) and stably expressing cell clones were obtained. Control and shPGS1 cells were passaged at 80% confluence and plated on #1 thickness laminin coated coverglasses (GG12-laminin, VWR). After allowing cells to adhere to the coverglasses for 2 hours,

cells were fixed using 4% v/v PFA for 15 min and stained with 100 nM NAO in the dark for 30 min. Cells were SRS imaged transmissively through coverglasses.

Immunofluorescence staining was performed following using a polyclonal rabbit anti-PGS1(Sigma-Aldrich, Cat# AV48896) and secondary antibody conjugated with Alexa-488 (Abcam, Cat# ab150081).

3.4.2 Human Kidney Tissue Preparation

De-identified human kidney tissue sections (30μm) were prepared from 4% v/v PFA-fixed frozen biopsy samples using a Compresstome (VF-210-0Z, Precisionary). The kidney cortex was isolated for imaging. Samples were imaged between 1mm thick glass slide and #1 thickness coverglass, submerged in 1x PBS.

3.4.3 Human Brain Tissue Preparation

De-identified post-mortem autopsy human brain sections (6 μm) were prepared from formalin-fixed and paraffin-embedded cortex tissue of control subject without detectable neuropathology as previously published [135, 136]. The sections were deparaffinized following a published protocol [111]. Subsequent SRS imaging was conducted with the tissue sections sandwiched in PBS between 1mm thick glass slides and #1 thickness cover glass.

3.4.4 Mouse Brain Tissue Preparation

Young (3 months) and aged (18 months) mice were euthanized with 5% isoflurane, and then perfused with 4% paraformaldehyde. The brains were harvested and fixed in 4% paraformaldehyde at 4°C for overnight. The fixed brains were washed with PBS and cut into 120-μm thickness slices with Vibratomes (Precisionary). The brain slices were placed in the center of a spacer and sandwiched between glass slides and coverslip for hyperspectral SRS imaging.

3.4.5 Drosophila Fat body Samples

Wild type (w^{1118} stock #5905) were originally obtained from the Bloomington Stock Center and have been maintained in the lab for several generations. Fat bodies were dissected from day 7 adult flies and fixed in 4% PFA (in 1xPBS) for 15 min. Samples were imaged immediately using SRS microscopy for hyperspectral imaging.

3.4.6 Spontaneous Raman Spectroscopy

Spontaneous Raman scattering spectra were obtained by a confocal Raman microscope (XploRA PLUS, Horiba) equipped with a 532 nm diode laser source and 1800 lines/mm grating. The acquisition time is 30 s with an accumulation of 4. The excitation power was ~40 mW after passing through a 100x objective (MPLN100X, Olympus). Output spectra were background subtracted and vector and simplex normalized. The pure lipid reference standards were placed on glass slides for spontaneous Raman spectra measurement. All lipid subtype reference spectra were acquired in the same manner.

3.4.7 Stimulated Raman Scattering Microscopy

An upright laser-scanning microscope (DIY multiphoton, Olympus) with a 25x water objective (XLPLN, WMP2, 1.05 NA, Olympus) was applied for near-IR throughput. Synchronized pulsed pump beam (tunable 720–990 nm wavelength, 5–6 ps pulse width, and 80 MHz repetition rate) and Stokes (wavelength at 1032 nm, 6 ps pulse width, and 80 MHz repetition rate) were supplied by a picoEmerald system (Applied Physics & Electronics) and coupled into the microscope. The pump and Stokes beams were collected in transmission by a high NA oil condenser (1.4 NA). A high O.D. shortpass filter (950 nm, Thorlabs) was used that would completely block the Stokes beam and transmit the pump beam only onto a Si photodiode for detecting the stimulated Raman loss signal. The output current from the photodiode was terminated, filtered, and demodulated in X with a zero phase shift by a lock-in amplifier (HF2LI, Zurich Instruments) at 20 MHz. The demodulated signal was fed into the FV3000 software module FV-OSR (Olympus) to form the

image during laser scanning. All SRS images were obtained with a pixel dwell time 40 µs and a time constant of 30 µs. A stack of 512 pixel x512 pixel x76 images in the C-H stretching region took approximately 15 minutes to acquire. The PRM analysis of this image-stack took less than 1 min. Laser power incident on the sample was approximately 40mW. Stimulated Raman histology was performed following published protocol [137].

3.4.8 Gas Chromatography Mass Spectrometry (GC-MS)

Hippocampal slices (n=4 per group) from 3 month old and 18 month old mice were homogenized in ethanol/water1:1 (v/v) and the homogenate were sent to Lipotype GmbH (Dresden, Germany) for mass spectrometry-based lipid analysis [138]. Lipids were extracted using a two-step chloroform/methanol procedure [139]. Samples were spiked with internal lipid standard mixture containing: cardiolipin 14:0/14:0/14:0/14:0 (CL), ceramide 18:1;2/17:0 (Cer), diacylglycerol 17:0/17:0 (DAG), hexosylceramide18:1;2/12:0 (HexCer), lyso-phosphatidate 17:0 (LPA), lyso-phosphatidylcholine 12:0 (LPC), lyso-phosphatidylethanolamine 17:1 (LPE), lyso-phosphatidylglycerol 17:1(LPG), lyso-phosphatidylinositol 17:1 (LPI), lyso-phosphatidylserine 17:1 (LPS), phosphatidate 17:0/17:0 (PA), phosphatidylcholine 17:0/17:0 (PC),phosphatidylethanolamine 17:0/17:0 (PE), phosphatidylglycerol 17:0/17:0 (PG),phosphatidylinositol 16:0/16:0 (PI), phosphatidylserine 17:0/17:0 (PS), cholesterolester 20:0 (CE), sphingomyelin 18:1;2/12:0;0 (SM), triacylglycerol 17:0/17:0/17:0 (TAG) and cholesterol D6 (Chol). After extraction, the organic phase was transferred to an infusion plate and dried in a speed vacuum concentrator. First step dry extract was re-suspended in 7.5 mM ammonium acetate in chloroform/methanol/propanol (1:2:4, V:V:V) and 2nd step dry extract in 33% ethanol solution of methylamine in chloroform/methanol (0.003:5:1; V:V:V). All liquid handling steps were performed using Hamilton Robotics STARlet robotic platform with the Anti Droplet Control feature for organic solvents pipetting.

Samples were analyzed by direct infusion on a QExactive mass spectrometer (ThermoScientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences). Samples were analyzed in both positive and negative ion modes with a resolution of Rm/z=200=280000 for MS and Rm/z=200=17500 for MSMS experiments, in a single acquisition. MSMS was triggered by an inclusion list encompassing corresponding MS mass ranges scanned in 1 Da increments [140]. Both MS and MSMS data were combined to monitor CE, DAG and TAG ions as ammonium adducts; PC, PC O-, as acetate adducts; and CL, PA, PE, PE O-, PG, PI and PS as deprotonated anions. MS only was used to monitor LPA, LPE, LPE O-, LPI and LPS as deprotonated anions; Cer, HexCer, SM, LPC and LPC O- as acetate adducts and cholesterol as ammonium adduct of an acetylated derivative [141]. Data were analyzed with in-house developed lipid identification software based onLipidXplorer [142]. Data post-processing and normalization were performed using an in-house developed data management system. Only lipid identifications with a signal-to-noise ratio >5, and a signal intensity 5-fold higher than in corresponding blank samples were considered for further data analysis.

3.4.9 Image Processing

SRS images were converted to unsigned 16 bit images via MATLAB, and were filtered using a morphological top-hat algorithm with 8 structuring elements, where appropriate. Unless used in ratiometric calculations, images for display were background subtracted using a sliding paraboloid with a radius of one tenth the image length. Intensity profiles and color maps were generated from ImageJ. All images within a figure have the same contrast unless specified otherwise. Ratiometric and overlaid images were created using the Image Calculator function and Overlay function, respectively, in ImageJ. Statistical analyses were performed using SPSS.

3.4.10 Penalized Reference Matching Algorithm Computation

Computation was conducted as described in the main text using MATLAB R2021b using an 8 Core Intel i9-9880H CPU, NVIDIA Quadro RTX 4000, and 64GB of RAM. Spectra were

intensity normalized from 0 to 1 following baseline correction using arPLS (if background spectra were not available for subtraction). All spectra were interpolated at every integer wavenumber using the interp1 function to avoid dimension mismatch errors during inner product calculations. Spectra were also Euclidean normalized using the standard vecnorm or norm functions in MATLAB. If a spectral shift of the reference spectrum exceeded the range of the original analyte spectrum, it was padded with zeros on the leading side, and trimmed on the lagging side. For timing and efficiency calculations, no parallel workers were used to split the spectral dataset for processing, but more workers are possible in MATLAB if supported by the hardware if the dataset is exceptionally large.

3.5 Supplementary Material

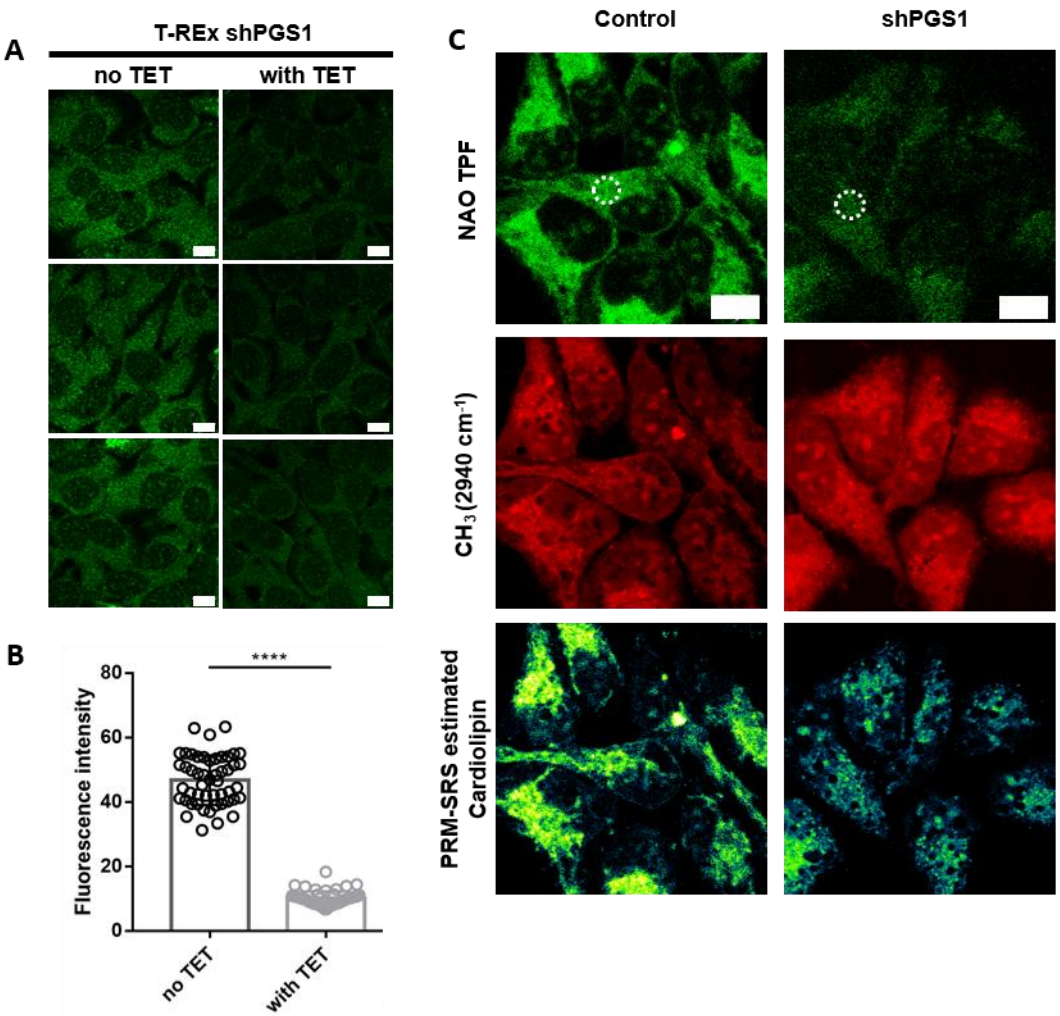


Figure 3.9 IF staining of PGS1 expression in shPGS1 HEK cells
(A) Immunofluorescence staining of stable shPGS1 cells using the specific PGS1 antibody. **(B)** Quantification of immunofluorescence signal intensity in images shown in panel A. Compared to the control group, the PGS1 protein level was significantly reduced by shPGS1 following induction with tetracycline (Tet). **(C)** TPF (following NAO staining), SRS, and PRM-SRS images from Fig. 4 demonstrate that the similarity score-based PRM-SRS image for CL distribution is similar to that shown by NAO staining (TPF) but distinct from single Raman shift SRS images. Scale bar: 10 µm

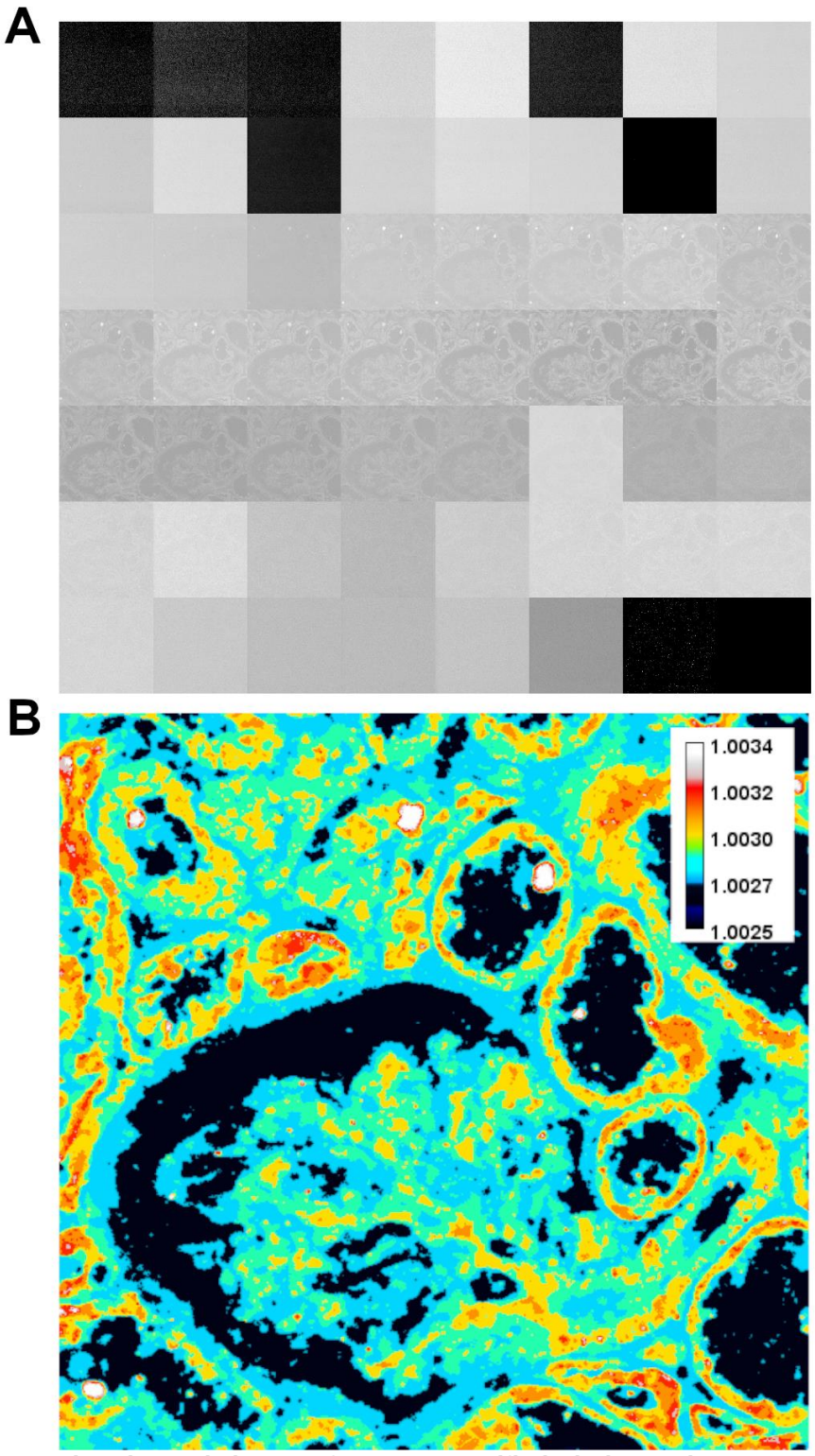


Figure 3.10 HSI and Ratiometric image of Free to Esterified Cholesterol PRM-SRS images
(A) Hyperspectral image frames of the human kidney tissue section from 2750cm^{-1} to 3050cm^{-1}. **(B) Cholesterol:cholesterol ester ratiometric image.** Ratio values are based on the similarity scores from the PRM-SRS images. Higher similarity scores for free cholesterol occur intracellularly in tubules and glomerular epithelial cells that line the arteriole and capillaries of the mesangium, as well as the larger deposits indicated by solid arrows in Figure 5B.

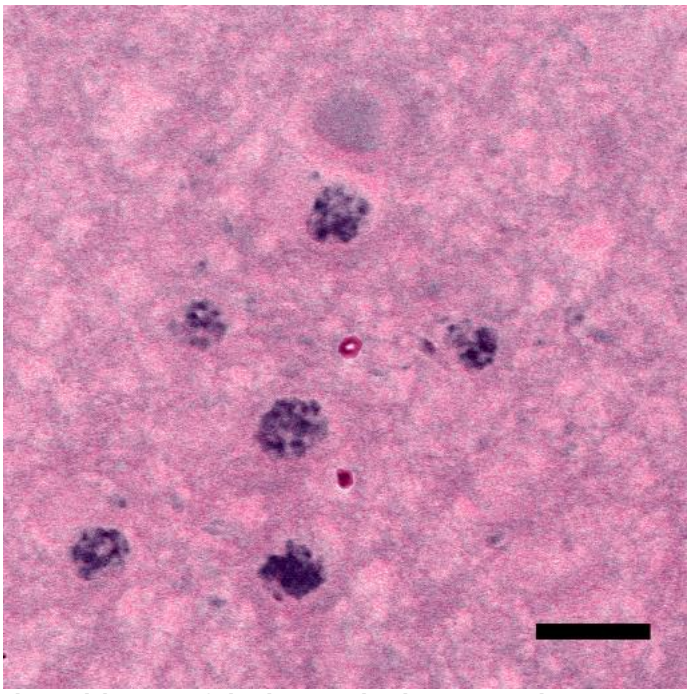


Figure 3.11 SRH (virtual H&E) image of the human brain temporal cortex sample
Scale bar: 10 µm.

Acknowledgements

Chapter 3, in full, is a reprint of the material as it appears in BiorXiv 2022. Anthony A. Fung[1+], Wenxu Zhang[1+], Yajuan Li[1+], Zhi Li[1#], Hongje Jang[1#], Honghao Zha[1], Xiaoping Chen[2], Fangyuan Gao[3], Jane Y. Wu[2], Huaxin Sheng[4], Junjie Yao[5], Dorota Skowronska-Krawczyk[3], Sanjay Jain[6] & Lingyan Shi[1*]. The dissertation author was the primary investigator and author of this paper. Additionally, we thank Drs. K. Zhang, C. Metallo, F. Liu, and G. Schmid-Schoenbein for helpful discussions. We are grateful for the support of the Washington University Kidney Translational Research Center (KTRC) for kidney samples and the HuBMAP grant U54HL145608. We thank Dr. E. Bigio and Dr. M-M. Mesulam from Mesulam Center for Cognitive Neurology and Alzheimer's Disease (MCCNAD) for providing the de-identified autopsy brain samples; and MCCNAD is supported by NIH P30 AG013854.

Chapter 4: Tissue-Level Analysis of Diabetic Kidney Disease

4.1 Introduction

Metabolic disorders such as diabetes can affect kidney redox balance, lipid regulation, and collagen homeostasis. With over 40% of diabetic patients developing diabetic kidney disease (DKD), the leading cause of chronic kidney disease (CKD) [143], diabetes is one of the most significant risk factors in kidney-related mortality. Independent reports of both glomerular [144] and tubulointerstitial [143, 145] changes in DKD include hyalinosis of capillaries and arterioles, collagen accumulation, glomerular basement membrane (GBM) thickening, glomerular mesangium volume fraction (Vv(Mes/Glom)) increases, and ectopic lipid accumulation. In practice, however, there are two major limitations to the conventional staining methods. First, multiplexing capabilities are limited by the chroma of the stain or absorption/emission wavelengths of the fluorescent probe. Several washes in alcohols and various acids or alcohols [146, 147] may be required to re-color the sample with a different stain, altering the native lipid distribution [148], and affect future fluorescent staining. To circumvent this, adjacent sections are used,but may not necessarily exhibit the same morphologies and require larger biopsies from the patient. Second, the power of morphological analysis is greater with 3D imaging. Both limitations may require excessive tissue availability to increase the power of the analyses, imposing additional risk to the patient. To address these issues, a label-free imaging platform using a single microscopy setup (**Fig. 4.1A**) was developed to capture spatial biomolecular aspects of DKD in 2D and 3D (**Fig. 4.1B**). With this platform, multiple views of the same sample are achieved (**Fig. 4.1C**), and can be extended to 3D samples without the need to co-register and stitch serial sections. Using this platform, morphological and biochemical differences in DKD tissues were identified, such as collagen anisotropy, collagen fiber thickness, mesangial proliferation, lipid subtype accumulation, and diminished optical redox ratio and lipid saturation scores.

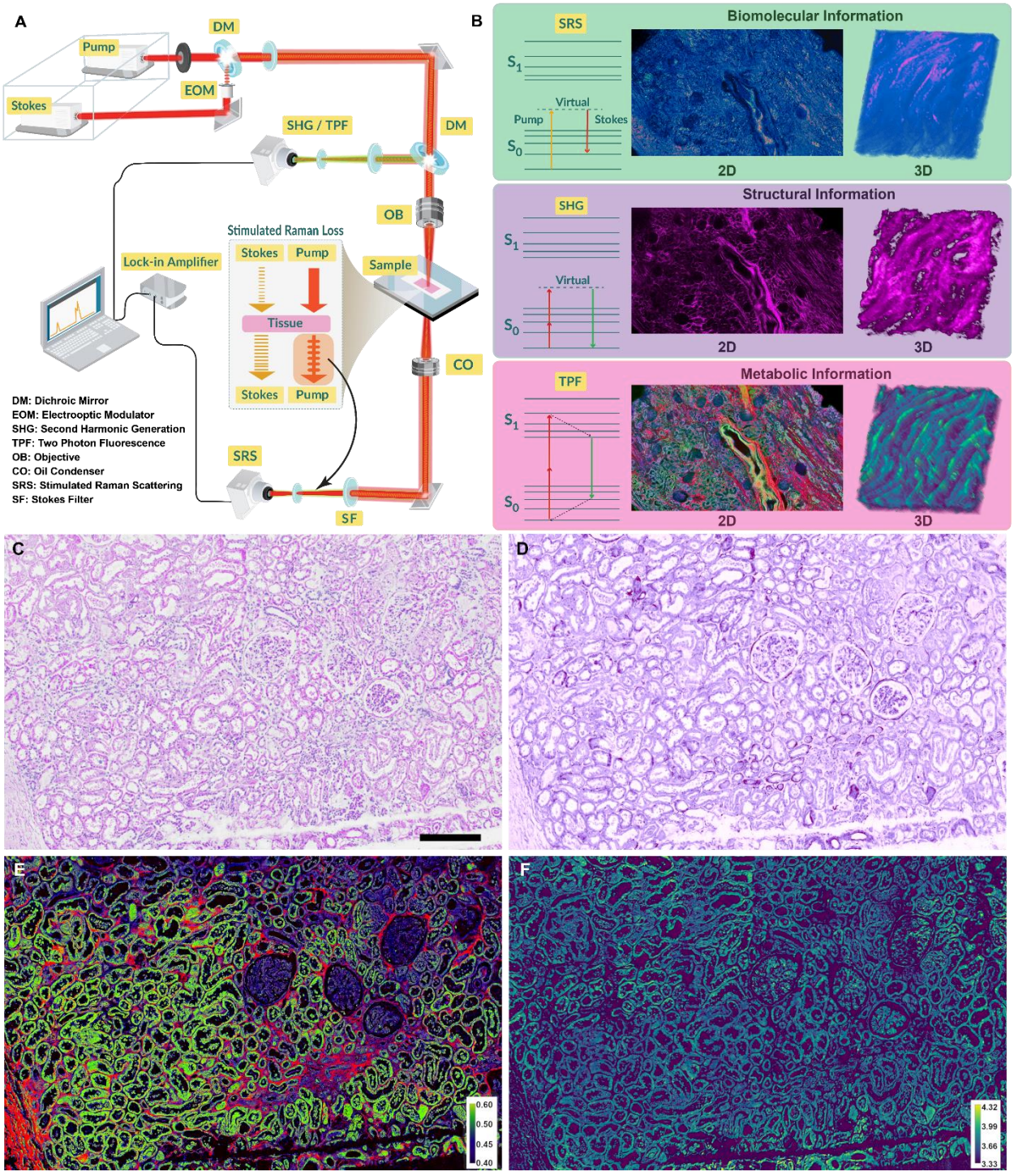


Figure 4.1 Label-free multimodal imaging platform analyzes human kidney biopsy
(A) a single microscope and laser system with back-scattered fluorescence and SHG photodiode, and transmissive SRS photodiode. **(B)** Example outputs from the platform including 2D and 3D SRS, SHG, and TPF. **(C)** SRH of pseudo-colored H&E stain. **(D)** SRH of pseudo-colored PAS stain. **(E)** Overlaid Collagen type 1-3 SHG and ORR. **(F)** Lipid saturation SRS ratio. Scale bar 200 microns.

Morphological differences in DKD include inflamed Vv(Mes/Glom) - the ratio of the capillary mesangium to the entirety of the glomerulus, including the Bowman's space – and thicker collagen fibers. Studies show that CKD tissues exhibit a larger Vv(Mes/Glom), as the mesangial cell matrix expands [149–151]. This phenomenon has also been linked to high density cholesterol dyslipidemia [152–154]. However, this volume ratio is often measured from cross sections of tissue using pixel density, point sample intercept (PSI), or lineal analysis [155–158]. In a biopsy, however, glomerular units are often found in different focal planes, which may complicate diagnoses by reducing the statistical power of the analysis. Studies show that with sufficient sampling, variations within a patient sample become negligible [159], but presently there is no study comparing various 2D glomerular volume fraction estimation methods with 3D imaging results [160]. Briefly, the PSI method, also referred to as the Cavalieri method, reports a smaller volume than most other estimation methods, including 3D methods based on serial section reconstruction [160, 161]. This method relies on random sampling, encompassing both medial glomerular slices with a high mesangial volume fraction, and polar slices with a low mesangial volume fraction. Interestingly, mesangial expansion is relatively slow to develop and is typically more considered for advanced CKD, rather than DKD, although it has been frequently reported [144, 162]. Could this be due to the large variance in Vv(Mes/Glom) measurements? More recently, it was found that 2-plane sections, 20 microns apart, are sufficient to estimate mean glomerular volume [163]. In the context of mesangial matrix volume fraction, however, there is an intuitive chance that the physical slicing could distort these structures or fail to capture representative sections of glomeruli, thus increasing the reliance on more random sampling to account for this variance. This is less problematic for larger cross sections, but there is also the added step of ensuring the slices are co-localized during analysis. In figure 1 below, a stimulated Raman scattering (SRS) microscopy image of the CH_3 stretching region ($2935cm^{-1}$) shows a human kidney biopsy with glomeruli in various focal planes (**Fig. 4.2A**). The glomerulus in the middle represents an ideal case, with the maximum cross-sectional area and arteriole present. In

the upper left and right sides of Fig. 4.2A, polar slices of glomeruli can be seen, which have starkly contrasted mesangium-to-glomerulus area ratios. Serial slices on the order of 10 microns can cause significant differences in this area fraction as shown in **Fig. 4.2B**. Therefore, a focus on a more inclusive field of view is critical to fully utilize precious biopsy samples.

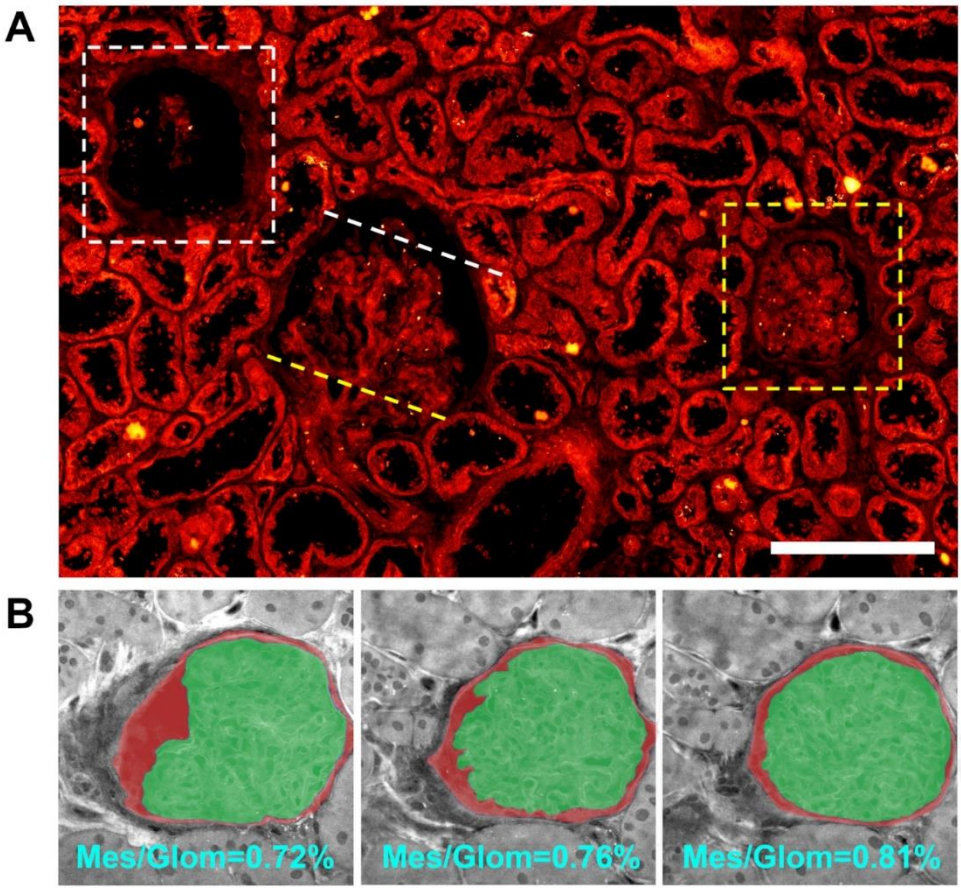


Figure 4.2 2D human kidney biopsy highlights variance of mesangial fraction
(A) glomeruli in various focus planes. White dashed box shows a glomerulus that was perhaps sliced distal to the arteriole, indicated by the white dashed line on the glomeruli sliced near the center. Yellow dashed box shows glomerulus that was perhaps sliced proximally to the arterioles, indicated by the yellow dashed line. **(B)** focus planes of a single glomerulus only 15microns apart from each other show significant differences in the area fraction of the mesangium (highlighted in green) in the glomerulus (highlighted in red). Scale bar 200um.

Biomolecular differences include lipid dysregulation, proteinuria, or altered NAD[P]H and flavin levels. These features are not entirely disjoint from the morphological differences. Fibrogenesis is heavily influenced by endothelial cell function, particularly in peritubular

microvessels [164]. In mice, the mitochondrial deacetylase, Sirtuin 3 (SIRT3), is not only a tight regulator of glucose and lipid metabolism, but a key metabolic programmer of renal fibrosis as well [165]. Kidney tissue has a relatively low level of fatty acid synthase so much of the lipid metabolism is undertaken by the incorporation of exogenous lipids and the mitochondrial-mediated breakdown of fatty acids. Dyslipidemia is frequently observed in nephrotic syndromes and kidney diseases and often presents as elevated levels of LDL cholesterol and decreased levels of HDL cholesterol in sera [166]. Oxidative stress has been associated with both fibrosis and lipid peroxidation. Previous studies have used the normalized optical redox ratio (ORR) as an indicator of oxidative stress, and can be measured using the label-free autofluorescence of NAD[P]H (390nm/460nm) and flavins (430nm/525nm) [167].

4.2 Results

4.2.1 Morphological Features of DKD

One of the most widely used morphological analysis methods involves histopathological staining and inspection with inverted light microscopy. To test the stainless staining capabilities of SRH, two Raman shifts were used, CH_2 stretching (2850 cm^{-1}) and CH_3 stretching (2935 cm^{-1}). These vibrational modes are broadly correlated with lipid and protein biomolecules, respectively. Custom lookup tables (LUTs) were applied to these channels and blended in MATLAB to reveal subcellular morphologies such as the nucleoli within the epithelial cells of tubules in a similar fashion to H&E histology. (**Fig. 4.3A-B**). This method was extended to 3D volumetric scanning of glomeruli (**Fig. 4.3C-D**) to analyze the structure of these functional tissue units. While some studies have delved into the capillary network of glomeruli using scanning electron microscopy (SEM) [168], most reports examine these structures in 2D using light microscopy (LM) and PSI analysis, which has been shown to underestimate mesangial volume by as much as 30 percent, relative to 3D reconstruction methods [161]. Still, these 3D methods require separate imaging of serial slices of tissue. Since SRS and two-photon fluorescence

microscopies are highly localized multiphoton processes, a 200 micron thick sample can acquire glomerular images at many focus planes, like the one shown in **Fig. 4.3E** by simply adjusting the height of the objective lens. **Figure 4.3F** shows how nuclei of glomeruli and tubules can be segmented in 3D images so they can be counted. This would help establish the correlation between cell proliferation in the mesangium and disease states such as DKD, but further segmentation of cell types within the glomerulus may be required.

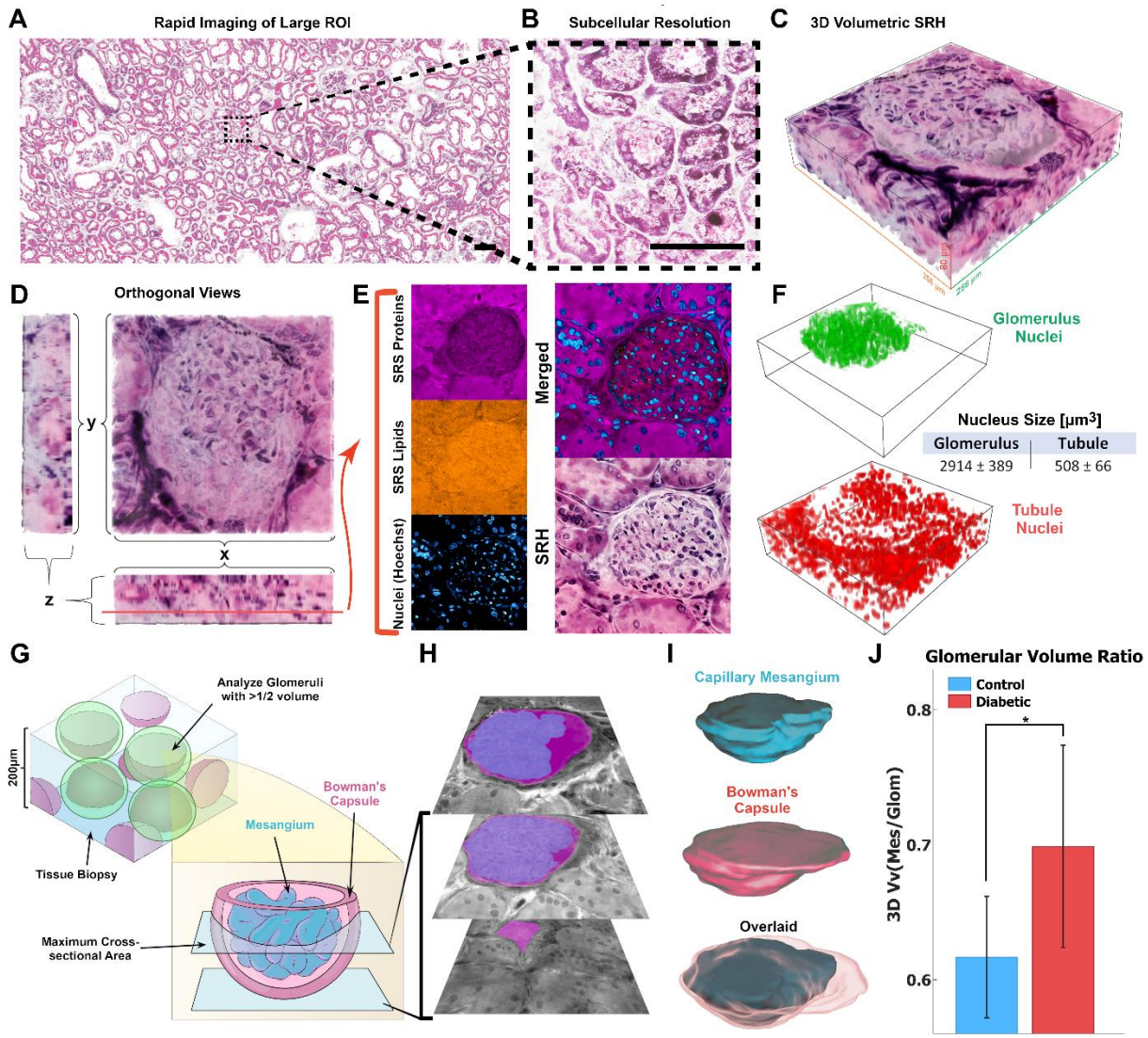


Figure 4.3 SRH and 3D analysis of glomeruli
(A) large human kidney ROI and small subcellular resolution **(B)** in which nucleoli are visible. **(C)** Confocal scanning in the Z-direction can produce 3D pseudo-colored SRH images. **(D)** Orthogonal projections of (C). **(E)** SRS and TPF images of lipids, protein, and nuclei juxtaposed to the same image slice of the 3D stack shows that multiple visualization schema are possible. **(F)** Glomerular volumes can be segmented to inform the number of nuclei within the glomerulus and in the surrounding tubule cortex. Scale bar 200um. **(G)** a schematic depicting the process of obtaining confocal image planes of glomeruli with at least a hemisphere within the imaging volume. Image planes between the maximum cross-sectional area and the vanishing point are retained. **(H)** Example of a few image planes of a glomerulus with the mesangium (cyan) and total glomerulus (Bowman's capsule) (magenta) overlaid on the CH_3 stretching images. **(I)** The image planes are used to reconstruct volumes in ImageJ and MeshMixer. **(J)** Summary of glomerular volume ratio [Vv(Mes/Glom)] between control samples and DKD samples shows that mean DKD glomeruli have significantly higher mesangial volume fractions compared to control specimens (n=4 for each group). Error bars represent one standard deviation of the mean. **P < 0.01.

Glomeruli are relatively spherical tissue structures with a diameter of roughly 200 microns [169].Taking a 200 micron thick 3D image of a cortex sample could yield whole glomeruli, but

more likely will encompass several hemispheres, or at least as many medial cross sections as a 2D laminar slice with all glomeruli at exactly the same focus plane. In this thick sample, glomeruli were segmented from the image stack between the slice with the largest glomerular cross section, and the slice where the glomerulus is no longer visible (**Fig. 3G-H**). These slices were interpolated to a volume using ImageJ and MeshMixer (**Fig. 3H**). Only glomeruli with at least one hemisphere in the kidney biopsy volume, or with a distance of at least 100 microns between the largest and smallest cross sectional area, were retained. Results confirmed that DKD specimen glomeruli have a significantly greater mesangial volume fraction than control specimens (**Fig. 3I-J**). However, these volume fractions were greater than previously reported results that used fine and coarse point counting such as the Cavalieri method [157, 158]. The same samples were also analyzed using a single plane, mimicking traditional methods, and each glomerulus' maximal cross section within the volumes. Table 1 below summarizes the comparison between these three methods.

Table 4.1 Comparison of various Vv(Mes/Glom) measurement methods.
2D single plane measurements were calculated using the ratio of pixels in the mesangium to the entire glomerulus in a single plane view, similar to traditional methods. 2D max individual measurements were taken from a 3D image of each glomerulus, selecting for the plane with the largest cross-sectional area, and using this to calculate the ratio of pixels. 3D hemisphere measurements were taken from 3D images of glomeruli with either a defined polar and local maximum cross-section, or those with at least 100 microns contained within the 3D image volume.

	2D Single Plane		2D Max Individual		3D Hemispheres	
	Control	DKD	Control	DKD	Control	DKD
Mean Vv(Mes/Glom)	0.5482	0.5833	0.6344	0.7344	0.6167	0.6987
Standard Deviation	0.2138	0.2005	0.1783	0.0762	0.0450	0.0756
Standard Error	0.1073	0.0102	0.0844	0.0117	0.0016	0.0523

Interestingly, the DKD samples exhibited a smaller variance between glomeruli in the 2D methods. This observation was in tandem with a lower incidence of polar cross sections showing a very small Vv(Mes/Glom). The lowest variance in measurements came from the 3D hemispherical method, which discarded several glomeruli based on the criteria previously described. Although volume measurements are more sensitive to small changes in radii than area

measurements due to their cubed dimension, if only hemispheres are collected and analyzed, the variance within the population is mitigated, while the variance across populations, such as control and DKD, is widened.

To rapidly identify glomeruli for analysis, unsupervised clustering may also be applied to hyperspectral SRS images in the CH stretching region, using SHG of collagen as a mask. The unique tubulointerstitial network of type I-III collagen in kidneys provides an excellent natural segmentation mask with minimal manual annotation that can be used to watershed these structures and intersect them with the SRS image to isolate the tissue pixels from the background pixels (**Fig. 4.4A-C**). This is because tissue samples may have varying proportions of background area in the image sets that may artifactually influence the chemometric data of those pixels. Each colored area in Fig. 4.4C is a datum that can be classified using k-means initialized with 3 clusters. These clusters were segmented and overlaid, along with the collagen SHG, in **Fig. 4.4D**. The spectral profiles of the unprocessed image stack and their corresponding second derivative barcodes (**Fig. 4.4E-F**) show the CH-stretching region of the pixels corresponding to each of the segmented images in Fig. 4D. This reiterates how little sample preparation and signal processing is required for such a hyperspectral image. A large DKD kidney sample that was processed in this manner (**Fig. 4.4G**) illustrates some broad morphological differences such as a denser tubulointerstital collagen network (green), tubules (red) with less space in their centers, and glomeruli (cyan) with a seemingly higher mesangial area fraction. While the spectral resolution is not sensitive enough to separate tubules, arterioles, collecting ducts, and other structures perfectly, unsupervised clustering of SRS HSI shows competitive utility compared with neural network models used for the same purpose with histological stains, such as Indica Labs software.

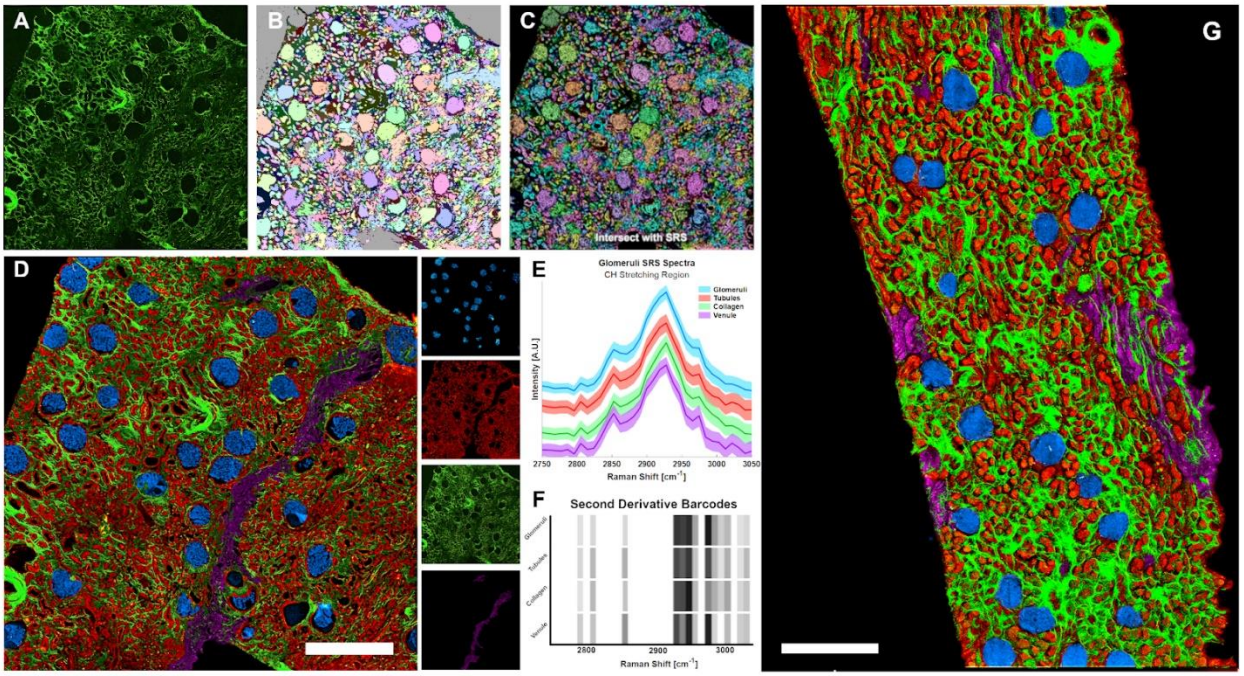
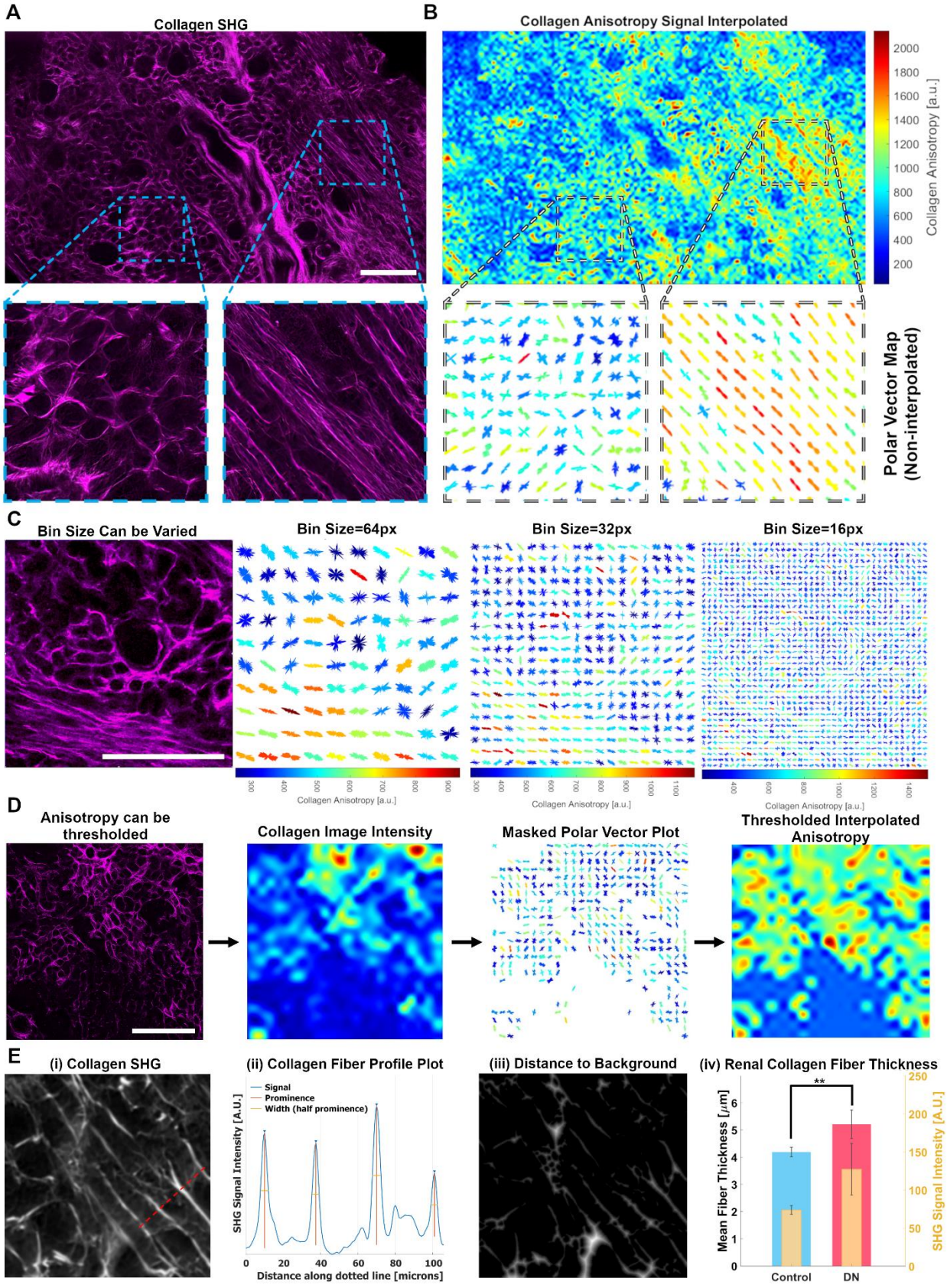


Figure 4.4 SRS hyperspectral imaging clusters glomeruli
(A) collagen SHG, used as a segmentation mask, to separate glomeruli/tubules with a watershed segmentation shown in **(B)**. **(C)** The resulting mask in (B) is intersected with the maximum intensity projection of the SRS € to remove background pixels and is then loaded into MATLAB. **(D)** Overlaid and separated pseudo-colored morphologies base on K-means clustering initialized with 4 clusters, with the pixels covered by collagen SHG (A) included as wel€**(E)** A plot of the mean and standard deviation of the pixel spectra from each labeled cluster. **(F)** A second derivative barcode improves the visualization of the wavenumbers of intere€in (E). **(G)** Diabetic Kidney sample analyzed with the same method reveals a much denser collagen intertubular network and inflamed tubule epithelial cells. Scale Bar 600um

Renal fibrosis is another morphological hallmark of renal failure and kidney diseases and has been studied using fluorescent probes for both collagen and collagen binding protein [170]. However, spatial analysis of tissue is often described by area covered by collagen fluorescence and manually-measured fiber thicknesses using correlative SHG and TEM [11]. Collagen fibrosis has other dimensions such as orientation and relative density. Anisotropy analysis of SHG images has been used to quantify collagen fiber orientation and recruitment to assess alterations in ocular microstructure [171],. Types 1-3 collagen fiber angles can be quantifiably visualized using polar vectors and label-free SHG imaging using the same stokes beam (1031nm) used in aforementioned SRS imaging. From this analysis alone, distinct regions of highly oriented fibers and tubule interstitium are highlighted (**Fig. 4.5A-B**). These regions, separated by a large vessel,

correspond to the cortex and medulla of the kidney, from left to right. For tissues with complex collagen structures like the kidney, binning the image into smaller areas can capture more structural information (**Fig. 4.5C**). Furthermore, regions without fibrillar collagen can be masked out using thresholded intensity images or anisotropy itself to avoid including background pixels in the Fourier analysis (**Fig. 4.5D**). Whole-slide collagen SHG images were transformed into distance maps with pixel intensity corresponding to distance to background. Even at this high level, DN samples are nearly 25 percent thicker than the control samples, with higher SHG intensity as well. By applying anisotropy analysis to fiber thickness measurements of a large tissue sample dataset, it is feasible to address the complexity of the collagen network by acquiring region-specific fiber thickness measurements. This may distinguish between fibrosis near glomeruli, in the tubulointerstitial matrix, or of medullary rays, each of which have distinct local anisotropy. Future studies will likely marry anisotropy analysis and machine learning algorithms to segment structural aberrations and aid in fibrosis diagnoses.

Figure 4.5 Collagen SHG anisotropy and diameter analysis
(A) SHG image of collagen in a 50 micron-thick human kidney sample using a 1031nm stokes laser. **(B)** Fourier-based analysis of collagen fiber orientation represented as polar vector plots (plot separation: 80um x 80um). Plot shapes reveal preferential directions of fibers, whereas colors reflect anisotropy scaling (warmer colours = higher anisotropy) **(C)** Bin sizes can be varied to capture polar vector representations of different portions of tissue. **(D)** Anisotropy can be thresholded using either the collagen signal intensity (depicted) or polar vector length to eliminate unwanted signal. **(E)** Collagen fiber thickness can be measured either manually via intensity profile (i-ii), or the image as a whole by selecting the local maxima of a distance map (iii). DN samples had a mean collagen fiber thickness 24.4% greater than the control samples (iv). Scale bar 500um.


A
Collagen SHG
B
Collagen Anisotropy Signal Interpolated
Collagen Anisotropy [a.u.]
Polar Vector Map
(Non-interpolated)
C
Bin Size Can be Varied
Bin Size=64px
Bin Size=32px
Bin Size=16px
Collagen Anisotropy [a.u.]
D
Anisotropy can be thresholded
Collagen Image Intensity
Masked Polar Vector Plot
Thresholded Interpolated Anisotropy
E
(i) Collagen SHG
(ii) Collagen Fiber Profile Plot
Signal
Prominence
Width (half prominence)
SHG Signal Intensity [A.U.]
Distance along dotted line [microns]
(iii) Distance to Background
(iv) Renal Collagen Fiber Thickness
Mean Fiber Thickness [μm]
SHG Signal Intensity [A.U.]
**
Control
DN

4.2.2 Biomolecular Features of DKD

As the link between structural and metabolic alterations in DKD becomes clearer, the next goal is to determine whether label-free imaging is capable of visualizing altered lipidomic profiles in the kidney tissues that expressed such drastic structural differences.

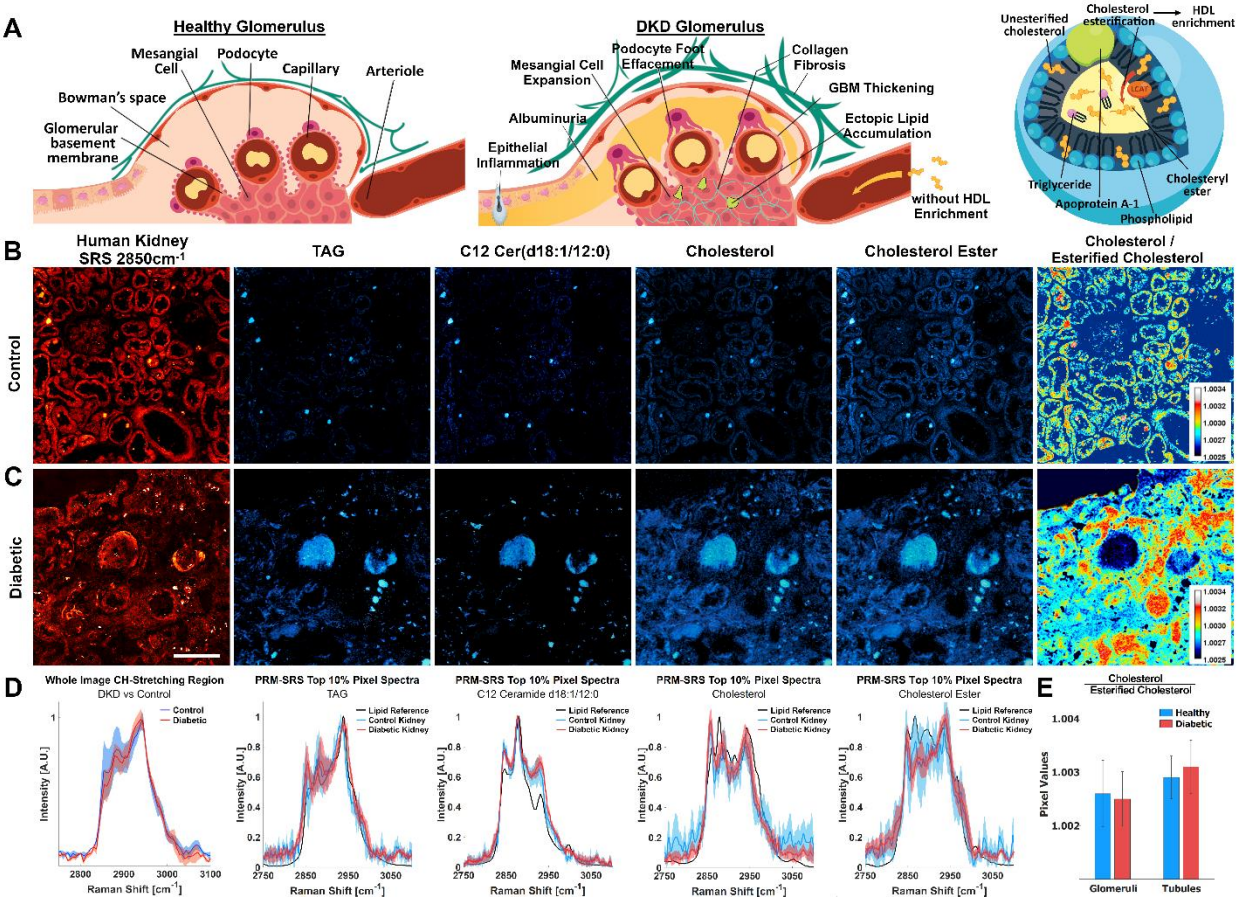


Figure 4.6 Label-free Lipid subtype visualization in situ using PRM-SRS
(A) A depiction illustrates common DKD nephropathies in relation to a hypothesized link with high free cholesterol and low serum HDL caused by lack of HDL enrichment. **(B-C)** SRS images of human kidney at the CH_2 stretching peak (Left) and PRM-SRS images of reference-matched lipid subtypes of interest showing the spatial distribution of similarity scores at the same contrast level. **(D)** HSI profile (Left) shows the Raman Spectra for all pixels in the HSI, with a red dashed line to indicate the 2850cm^{-1} image above. Spectra of the top 10% of similarity score pixels overlaid on the reference spectrum for each lipid subtype show confer PRM-SRS images. **(E)** Ratiometric image of cholesterol and esterified cholesterol similarity scores show highest levels in eosinophilic bodies and within tubule endothelial cells, and not glomerular capillaries and podocytes. DKD tissues show higher relative free cholesterol in the tubulointerstital region, and lower relative free cholesterol in the glomeruli. This may be due to hyaline or albuminuric material in the glomerulus of DKD, leading to poorer reference matching of lipid subtypes in that region. Scale bar, 200 µm.

The differences in lipid subtype levels in DKD may also affect the lipid metabolism within the tissue. In addition to multi-modal imaging capabilities (**Fig. 4.7A**), metabolic indicators become clearer when certain images are analyzed with respect to each other in the form of ratios. Two metabolic indicators that have been associated with lipid peroxidation are the normalized optical redox ratio and the lipid saturation score indicator [167]. Both indicators are lower in DKD tissues (**Fig. 4.7 B-E**). A lower normalized ORR may indicate less oxidation of fatty acids. A lower lipid saturation score may indicate higher levels of saturated fats relative to unsaturated fats, which can be typical of poorer diet and impaired HDL maturation.

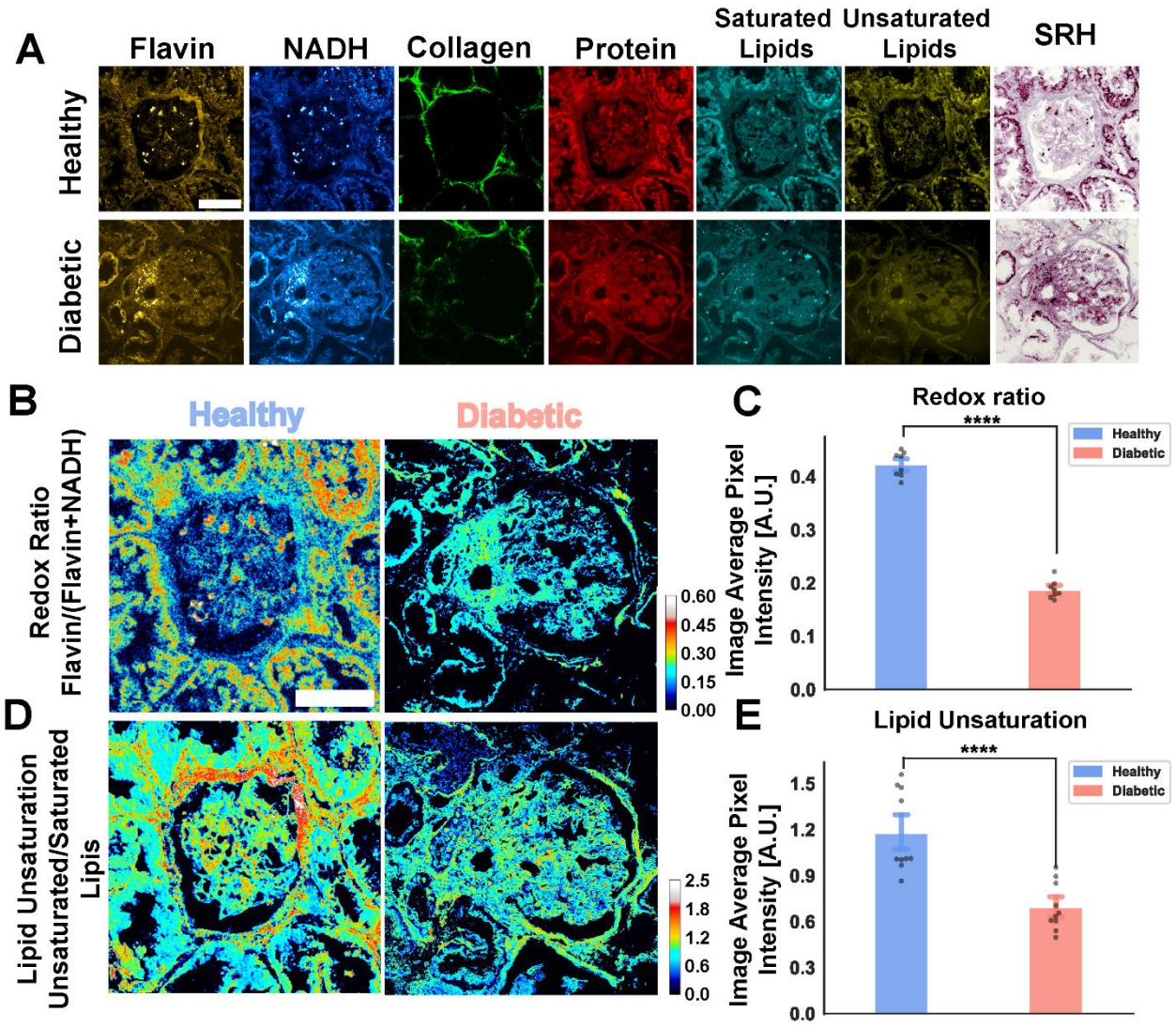


Figure 4.7 Oxidative stress and lipid saturation
(A) Glomerular subregion images using SRS and TPF imaging **(B)** Normalized Optical Redox Ratio as a measure of oxidative stress and metabolism in healthy and diabetic kidneys. **(C)** Quantitative image analysis of multiple ROI's. (D) Lipid saturation ratiometric images of healthy and diabetic kidneys. **(E)** Quantitative image analysis of multiple ROI's. Scale bar, 150 μm.

In certain tissues such as the heart, liver, and kidneys, a lower redox ratio has been found to correlate with disease states such as diabetes. This could be counter-intuitive because a greater pool of NAD[P]H reductive potential may warrant less oxidative stress, yet the opposite is observed. In diabetes, the reduced forms of pyridine nucleotides such as NADH and NADPH were elevated [172]. This could be due to a shift toward anaerobic pathways and away from oxidative phosphorylation. Furthermore, in studies that examined oxidative stress and renal

fibrosis, it was found that NOX activity was elevated in fibrotic cases [173]. Together, these studies suggest that greater levels of NAD[P]H contribute to oxidative stress and renal fibrosis.

Likewise, a decrease in flavins such as riboflavins, flavin mononucleotide (FMN), flavin monooxygenase (FMO), and flavin adenine dinucleotide (FAD), could also account for a lower ORR. In diabetic retinopathy, for example, a flavin deficiency hampers the production of glutathione, a primary antioxidant defense [174]. . In DN, which is known to undergo oxidative stress, it is therefore unsurprising that the tissue exhibits a lower optical redox ratio.

4.3 Discussion

To our knowledge, at the time of writing this study is the first label-free multi-modal imaging analysis of structural and molecular biomarkers in diabetic kidney disease, using a single microscopy platform. It highlights the potential to uncover a multiplexed data to improve patient outcomes by conserving tissue samples, and to inspire new research into the connection between dyslipidemia and kidney disease.

Structurally, we demonstrated 3D digital histology and segmentation of nuclei in distinct structures, as well as verified mesangial matrix expansion in DKD glomerulopathy using 3D confocal SRS imaging. This platform better preserves the native orientation of the tissues by avoiding excessive physical slicing and allows for more representative volumetric data of the mesangium volume fraction. Furthermore, hyalinosis and inflammation of the tubulointerstitial matrix and collagen fibrosis were confirmed without any labels or dyes. Anisotropy bin sizing and thresholding may also be a useful tool to dissect region-specific collagen metrics to examine fibrosis in DN in greater detail.

Biochemically, DKD tissues appear to have different relative lipid subtype levels. DKD samples have more intense PRM-SRS images for some of the most relevant lipid subtypes, indicating that their pixels have a more “lipid-like” hyperspectral profile. Furthermore, DKD

samples appear to have a slightly higher degree of cholesterol esterification in the tubulointerstitial matrix than in the glomeruli, though this result should be cross-validated using other methods. The glomerulus also had a more intense CH_3-symmetric stretching signal, which may indicate albuminuria. Finally, the normalized optical redox ratio and lipid saturation score suggest DKD kidneys have dysregulated lipid metabolism, with less oxidation of lipids and a higher degree of lipid saturation and oxidative stress in the DKD samples.

Future studies may iterate on these first demonstrations. For example, 3D digital histology may benefit from further segmentation of cell types, not just nuclei in certain regions. The isolation of glomeruli and their mesangial matrices could also be automated using AI, such as the HALO Platform (Indica Labs) [175], to further improve the efficiency of these measurements. Additional stains can also improve diagnostic power and fully leverage the label-free hyperspectral imaging platform. These stains can be further improved using cycle generative adversarial networks (cycle GANs) to make these SRH images appear more standardized so that a database can be established. It should be noted that SHG of collagen only visualizes types 1-3 due to their non-centrosymmetric structure, and thickness measurements and SHG signal intensity depend on fiber orientation and position relative to the focal volume. Since collagen fibrosis is tightly regulated by lipid metabolism and oxidative stress, future studies are necessary to determine the extent to which the SHG intensity and collagen anisotropy may predict other nephropathies such as lipid subtype differences, normalized optical redox ratio, and mesangial volume fraction.

Finally, since the biomarkers examined in this study don't require staining, the big question is whether the tissue needs to be excised at all. The samples examined above were of the outer kidney cortex. It is exciting to think about future endoscopic probes that may be able to observe these biomarkers using a epi-detection back-scattering configuration, thus circumventing the need for tissue excision, transit to pathology departments, sample preparation, and analyses.

4.4 Materials and Methods

4.4.1 HEK293 Cell Cultures

Human kidney samples were obtained from The Kidney Translational Research Center (KTRC) at the Washington University School of Medicine in St. Louis under an approved IRB protocol. Samples were fixed using 4% PFA () and stored in 1x PBS at -20deg C when not in use. Samples were sliced using a sliding microtome (HM 450, Epredia) at 200 µm and cleared using 8M urea for 48 hours at room temperature. Samples were submerged in 1ug/mL Hoechst 33342 (Thermofisher) for 15 minutes to stain nuclei for co-localization verification. Samples were imaged between 1mm thick glass slides and number 1.5 thickness cover glass (Erie Scientific).

4.4.2 3D Microscopy

An upright laser-scanning microscope (DIY multiphoton, Olympus) with a 25x water objective (XLPLN, WMP2, 1.05 NA, Olympus) was applied for near-IR throughput. Synchronized pulsed pump beam (tunable 720–990 nm wavelength, 5–6 ps pulse width, and 80 MHz repetition rate) and Stokes (wavelength at 1032nm, 6 ps pulse width, and 80MHz repetition rate) were supplied by a picoEmerald system (Applied Physics & Electronics) and coupled into the microscope. The pump and Stokes beams were collected in transmission by a high NA oil condenser (1.4 NA). A high O.D. shortpass filter (950nm, Thorlabs) was used that would completely block the Stokes beam and transmit the pump beam only onto a Si photodiode for detecting the stimulated Raman loss signal. The output current from the photodiode was terminated, filtered, and demodulated in X with a zero phase shift by a lock-in amplifier (HF2LI, Zurich Instruments) at 20MHz. The demodulated signal was fed into the FV3000 software module FV-OSR (Olympus) to form the image during laser scanning. All SRS images were obtained with

a pixel dwell time 40 µs and a time constant of 30 µs. Laser power incident on the sample is approximately 40mW.

Second Harmonic Generation (SHG) was used to capture type 1-3 collagen images. The 1031nm stokes laser described above, with 300mW and a dwell time of 10us per pixel, was used with 5-frame averaging. Backscattered SHG signals were filtered using 465nm filter.

NADH and Flavin autofluorescence images were captured using the 800nm pump beam with 350mW and a pixel dwell time of 10us/px with 5-frame averaging. Backscattered signals were filtered using a dual filter cube of 465nm and 515nm. All wide-view tile-stitching and confocal z-scanning was controlled by the Fluoview software (Olympus).

4.4.3 Stimulated Raman Histology

Stimulated Raman histology was performed following section 1.3.3. For 3D representations, the opacity of white interstitium was achieved using a sigmoidal transfer function in the MATLAB VolumeViewer alphamap field.

4.4.4 Collagen Anisotropy and Diameter Measurements

Collagen fiber direction and anisotropy, along with total collagen fiber signal, was extracted into 80um x 80um regional bins using a bespoke MATLAB code. Details of the methods are published elsewhere [10]. In brief, the method applies a 2D discrete Fourier transform (DFT)–based algorithm to the SHG images and features a novel integrated periodicity plus smooth image decomposition to correct DFT edge discontinuity artefacts, minimizing the loss of peripheral image information that limits more commonly used DFT methods. Collagen fiber thickness measurements were achieved by transforming whole-slide SHG images into distance maps in which pixel intensity values correspond to its distance to the nearest background pixel using the bwdist function in MATLAB. Local Maxima was recorded and multiplied by 2, since the pixel

distance is half the thickness. Background subtraction was done using a flatfield correction and top-hat filter in MATLAB with a disk morphological structuring element.

4.4.5 3D Vv(Mes/Glom)

The CH_3 Raman shift ($2935cm^{-1}$) was used for the 3D image analysis of glomeruli because of its high intensity and optimal contrast. Glomeruli were manually identified and the planes between their maximal cross-sectional area and vanishing point were analyzed. The thickness of the glomerulus is the distance between those focus planes. The radius of the glomerulus is the average distance from the centroid of the glomerulus to the edge of the Bowman's capsule at the plane with the maximum cross-sectional area. If no maximum cross-sectional area could be defined (i.e. the plane with the largest area occurs at the first or last plane of the 3D image) then the radius is assumed to be 100 microns. Those glomeruli that showed thicknesses of less than 90 percent of their respective radii were discarded. These may have been due to the region being a polar slice (less than a hemisphere), or the glomerulus not fully contained within the imaging volume, respectively.

Auxiliary analysis using the maximum cross-sectional area plane for the volume fraction estimation was also performed. For these measurements, only glomeruli clipped at the edges of the image were discarded, just as with the traditional 2D method.3.4.6 Spontaneous Raman Spectroscopy

4.4.6 Glomerular Segmentation

Single plane hyperspectral SRS images of human kidney were acquired along with the collagen SHG images. The SRS hyperspectral image was projected to 2D using maximum intensity projection (MIP) for ease of viewing. The collagen SHG image was transformed into a mask by maximizing the contrast and closing gaps manually using the pen tool in ImageJ. The muscle MorphoLibJ plugin was used to generate a segmentation of the collagen mask, which was then intersected with the MIP image and thresholded to remove the empty spaces such as those

between tubules. Pixels within the segmented regions correspond to a datum, which were then k-means clustered initialized with 3 clusters. Segments were pseudo-colored by cluster identity using lookup tables for visualization.

4.4.7 Spontaneous Raman Spectroscopy

Spontaneous Raman scattering spectra were obtained by a confocal Raman microscope (XploRA PLUS, Horiba) equipped with a 532 nm diode laser source and 1800 lines/mm grating. The acquisition time is 20 s with an accumulation of 4. The excitation power was ~40 mW after passing through a 100x objective (MPLN100X, Olympus). Output spectra were background subtracted and vector and simplex normalized. The pure lipid reference standards were placed on glass slides for spontaneous Raman spectra measurement. All lipid subtype reference spectra were acquired in the same manner.

4.4.8 PRM-SRS Lipid Subtyping

PRM-SRS was conducted according to previously published methods [176]. Briefly, pure lipid subtypes such as those shown in figure 1.13 and 1.14 were spectroscopically analyzed using spontaneous Raman scattering. These spectra were used as reference spectra. SRS HSI had each pixel scored using cosine similarity in the CH stretching region at multiple offsets. Similarity scores were had a corresponding penalty that varied according to the offset to account for non-linear, lensing, or other non-standardized equipment effects. Resulting images contain pixels with simplex-normalized similarity scores, whose intensity corresponds to the relative similarity of that pixel's spectrum to the reference spectrum. These relative intensities are interpreted as relative concentrations.

Computation was conducted as described in the main text using MATLAB R2021b using an 8 Core Intel i9-9880H CPU, NVIDIA Quadro RTX 4000, and 64GB of RAM. Spectra were intensity normalized from 0 to 1 following baseline correction using arPLS (if background spectra were not available for subtraction). All spectra were interpolated at every integer

wavenumber using the interp1 function to avoid dimension mismatch errors during inner product calculations. Spectra were also Euclidean normalized using the standard vecnorm or norm functions in MATLAB. If a spectral shift of the reference spectrum exceeded the range of the original analyte spectrum, it was padded with zeros on the leading side, and trimmed on the lagging side. For timing and efficiency calculations, no parallel workers were used to split the spectral dataset for processing, but more workers are possible in MATLAB if supported by the hardware if the dataset is exceptionally large.

4.4.9 Statistical Analysis

A total of 2 DKD samples and 3 control samples were used in this study. For lipid subtype, collagen anisotropy, and optical redox ratio calculations, 5 regions of interest were obtained from each sample. For glomerular volume fraction calculations, 5 representative hemispherical glomeruli were used in each group. A simple student's t-test was used for all measures of significance.

Acknowledgements

Chapter 4, in full is currently being prepared for submission for publication of the material. Anthony A. Fung, Zhi Li, Craig Boote, Petar Markov, Sanjay Jain, Lingyan Shi. The dissertation author was the primary researcher and author of this material.

Chapter 5: Concluding Remarks and Outlooks

Raman scattering was introduced to the world less than a century ago, and to the biomedical space in the past few decades. In this fast-growing field it is important to consider the translation of these powerful technologies and concepts at every opportunity. The integration of multiple modalities, including Raman scattering, culminates in a unique platform capable of visualizing and analyzing subcellular organelles such as LDs, and entire FTUs, at both a molecular and structural level. Through this platform, breast cancer lipid metabolism was characterized under tandem nutrient modulation *in vitro*, non-destructive lipid subtyping was demonstrated, and diabetic kidney disease biomarkers were confirmed *ex vivo*.

While the development of tools such as relative entropy and PRM-SRS can expedite the workflow for finding potential molecules of interest and estimate lipidomic changes, there is always room for improvement. Due to the small Raman scattering cross section, imaging times are still considered slow, especially when a high image quality is demanded. New technologies are continuously being developed to improve the Raman signal quality and imaging speed, while new algorithms are developed in tandem to improve analytical power and processing time. Recently, a super-resolution deconvolution algorithm that can be performed on a single image frame was developed, which can achieve high image qualities without excessive imaging times [177]. Furthermore, 3D volumetric imaging SRS tomography (SRST) methods have been developed to perform label-free deep tissue imaging without the need for a z-scanning stage, making 3D imaging as fast as 2D imaging [178]. Future improvements to this label-free optical imaging platform will surely incorporate novel algorithms and technological advancements such as these.

References

1. Lee KS, Palatinszky M, Pereira FC, Nguyen J, Fernandez VI, Mueller AJ, Menolascina F, Daims H, Berry D, Wagner M, Stocker R (2019) An automated Raman-based platform for the sorting of live cells by functional properties. Nat Microbiol 4:1035–1048. https://doi.org/10.1038/s41564-019-0394-9

2. Suzuki Y, Kobayashi K, Wakisaka Y, Deng D, Tanaka S, Huang C-J, Lei C, Sun C-W, Liu H, Fujiwaki Y, Lee S, Isozaki A, Kasai Y, Hayakawa T, Sakuma S, Arai F, Koizumi K, Tezuka H, Inaba M, Hiraki K, Ito T, Hase M, Matsusaka S, Shiba K, Suga K, Nishikawa M, Jona M, Yatomi Y, Yalikun Y, Tanaka Y, Sugimura T, Nitta N, Goda K, Ozeki Y (2019) Label-free chemical imaging flow cytometry by high-speed multicolor stimulated Raman scattering. Proceedings National Academy of Science 116:15842–15848. https://doi.org/10.1073/pnas.1902322116

3. Wang J, Lin K, Zheng W, Yu Ho K, Teh M, Guan Yeoh K, Huang Z (2015) Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy improves in vivo diagnosis of esophageal squamous cell carcinoma at endoscopy. Sci Rep 5:12957. https://doi.org/10.1038/srep12957

4. Goulart ACC, Silveira L, Carvalho HC, Dorta CB, Pacheco MTT, Zângaro RA (2022) Diagnosing COVID-19 in human serum using Raman spectroscopy. Lasers Med Sci 37:2217–2226. https://doi.org/10.1007/s10103-021-03488-7

5. Ma X, Sun X, Wang H, Wang Y, Chen D, Li Q (2018) Raman Spectroscopy for Pharmaceutical Quantitative Analysis by Low-Rank Estimation. Front Chem 6:

6. Paudel A, Raijada D, Rantanen J (2015) Raman spectroscopy in pharmaceutical product design. Adv Drug Deliv Rev 89:3–20. https://doi.org/10.1016/j.addr.2015.04.003

7. Lin L, Ye J (2021) Spontaneous Raman and Surface-Enhanced Raman Scattering Bioimaging. Adv Exp Med Biol 3233:177–195. https://doi.org/10.1007/978-981-15-7627-0_9

8. Shen Y, Xu F, Wei L, Hu F, Min W (2014) Live-cell quantitative imaging of proteome degradation by stimulated Raman scattering. Angew Chem Int Ed Engl 53:5596–5599. https://doi.org/10.1002/anie.201310725

9. Miller DR, Jarrett JW, Hassan AM, Dunn AK (2017) Deep Tissue Imaging with Multiphoton Fluorescence Microscopy. Curr Opin Biomed Eng 4:32–39. https://doi.org/10.1016/j.cobme.2017.09.004

10. Pijanka JK, Markov PP, Midgett D, Paterson NG, White N, Blain EJ, Nguyen TD, Quigley HA, Boote C (2019) Quantification of collagen fiber structure using second harmonic generation imaging and two-dimensional discrete Fourier transform analysis: Application to the human optic nerve head. J Biophotonics 12:e201800376. https://doi.org/10.1002/jbio.201800376

11. Bancelin S, Aimé C, Gusachenko I, Kowalczuk L, Latour G, Coradin T, Schanne-Klein M-C (2014) Determination of collagen fibril size via absolute measurements of second-harmonic generation signals. Nat Commun 5:4920. https://doi.org/10.1038/ncomms5920

12. Sugita S, Suzumura T, Nakamura A, Tsukiji S, Ujihara Y, Nakamura M (2021) Second harmonic generation light quantifies the ratio of type III to total (I + III) collagen in a bundle of collagen fiber. Sci Rep 11:11874. https://doi.org/10.1038/s41598-021-91302-3

13. Strupler M, Pena A-M, Hernest M, Tharaux P-L, Martin J-L, Beaurepaire E, Schanne-Klein M-C (2007) Second harmonic imaging and scoring of collagen in fibrotic tissues. Opt Express 15:4054–4065. https://doi.org/10.1364/OE.15.004054

14. Kullback S, Leibler RA (1951) On Information and Sufficiency. Ann Math Stat 22:79–86. https://doi.org/10.1214/aoms/1177729694

15. Sgarro A (1981) Informational divergence and the dissimilarity of probability distributions. Calcolo 18:293–302

16. Garcıa-Garcıa D, Williamson RC Divergences and Risks for Multiclass Experiments. 20

17. Shi L, Zheng C, Shen Y, Chen Z, Silveira ES, Zhang L, Wei M, Liu C, de Sena-Tomas C, Targoff K, Min W (2018) Optical imaging of metabolic dynamics in animals. Nat Commun 9:2995. https://doi.org/10.1038/s41467-018-05401-3

18. Orringer DA, Pandian B, Niknafs YS, Hollon TC, Boyle J, Lewis S, Garrard M, Hervey-Jumper SL, Garton HJL, Maher CO, Heth JA, Sagher O, Wilkinson DA, Snuderl M, Venneti S, Ramkissoon SH, McFadden KA, Fisher-Hubbard A, Lieberman AP, Johnson TD, Xie XS, Trautman JK, Freudiger CW, Camelo-Piragua S (2017) Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. Nat Biomed Eng 1:1–13. https://doi.org/10.1038/s41551-016-0027

19. Freudiger CW, Pfannl R, Orringer DA, Saar BG, Ji M, Zeng Q, Ottoboni L, Wei Y, Ying W, Waeber C, Sims JR, De Jager PL, Sagher O, Philbert MA, Xu X, Kesari S, Xie XS, Young GS (2012) Multicolored stain-free histopathology with coherent Raman imaging. Lab Investig J Tech Methods Pathol 92:1492–1502. https://doi.org/10.1038/labinvest.2012.109

20. Khan AM, Rajpoot N, Treanor D, Magee D (2014) A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans Biomed Eng 61:1729–1738. https://doi.org/10.1109/TBME.2014.2303294

21. Fu D, Xie XS (2014) Reliable Cell Segmentation Based on Spectral Phasor Analysis of Hyperspectral Stimulated Raman Scattering Imaging Data. Anal Chem 86:4115–4119. https://doi.org/10.1021/ac500014b

22. Felten J, Hall H, Jaumot J, Tauler R, de Juan A, Gorzsás A (2015) Vibrational spectroscopic image analysis of biological material using multivariate curve resolution-alternating least squares (MCR-ALS). Nat Protoc 10:217–240. https://doi.org/10.1038/nprot.2015.008

23. Schaefer PM, Kalinina S, Rueck A, von Arnim CAF, von Einem B (2019) NADH Autofluorescence—A Marker on its Way to Boost Bioenergetic Research. Cytometry A 95:34–46. https://doi.org/10.1002/cyto.a.23597

24. Blacker TS, Mann ZF, Gale JE, Ziegler M, Bain AJ, Szabadkai G, Duchen MR (2014) Separating NADH and NADPH fluorescence in live cells and tissues using FLIM. Nat Commun 5:3936. https://doi.org/10.1038/ncomms4936

25. Zhang Y, Guldner IH, Nichols EL, Benirschke D, Smith CJ, Zhang S, Howard SS (2021) Instant FLIM enables 4D in vivo lifetime imaging of intact and injured zebrafish and mouse brains. Optica 8:885–897. https://doi.org/10.1364/OPTICA.426870

26. Boote C, Palko JR, Sorensen T, Mohammadvali A, Elsheikh A, Komáromy AM, Pan X, Liu J (2016) Changes in posterior scleral collagen microstructure in canine eyes with an ADAMTS10 mutation. Mol Vis 22:503–517

27. Lutz V, Sattler M, Gallinat S, Wenck H, Poertner R, Fischer F (2012) Characterization of fibrillar collagen types using multi-dimensional multiphoton laser scanning microscopy. Int J Cosmet Sci 34:209–215. https://doi.org/10.1111/j.1468-2494.2012.00705.x

28. Camp Jr. CH, Lee YJ, Cicerone MT (2016) Quantitative, comparable coherent anti-Stokes Raman scattering (CARS) spectroscopy: correcting errors in phase retrieval. J Raman Spectrosc 47:408–415. https://doi.org/10.1002/jrs.4824

29. Chen G, Deng X (2018) Cell Synchronization by Double Thymidine Block. Bio-Protoc 8:e2994. https://doi.org/10.21769/BioProtoc.2994

30. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Sauer AG, Jemal A, Siegel RL (2019) Breast cancer statistics, 2019. CA Cancer J Clin 69:438–451. https://doi.org/10.3322/caac.21583

31. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, Shi B (2015) Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res 5:2929–2943

32. Olzmann JA, Carvalho P (2019) Dynamics and functions of lipid droplets. Nat Rev Mol Cell Biol 20:137–155. https://doi.org/10.1038/s41580-018-0085-z

33. Cruz ALS, Barreto E de A, Fazolini NPB, Viola JPB, Bozza PT (2020) Lipid droplets: platforms with multiple functions in cancer hallmarks. Cell Death Dis 11:1–16. https://doi.org/10.1038/s41419-020-2297-3

34. Daniëls VW, Smans K, Royaux I, Chypre M, Swinnen JV, Zaidi N (2014) Cancer cells differentially activate and thrive on de novo lipid synthesis pathways in a low-lipid environment. PloS One 9:e106913. https://doi.org/10.1371/journal.pone.0106913

35. Franco D, Trusso S, Fazio E, Allegra A, Musolino C, Speciale A, Cimino F, Saija A, Neri F, Nicolò MS, Guglielmino SPP (2017) Raman spectroscopy differentiates between sensitive and resistant multiple myeloma cell lines. Spectrochim Acta A Mol Biomol Spectrosc 187:15–22. https://doi.org/10.1016/j.saa.2017.06.020

36. Butler LM, Mah CY, Machiels J, Vincent AD, Irani S, Mutuku S, Spotbeen X, Bagadi M, Waltregny D, Moldovan M, Dehairs J, Vanderhoydonc F, Bloch K, Das R, Stahl J, Kench J, Gevaert T, Derua R, Waelkens E, Nassar ZD, Selth LA, Trim PJ, Snel MF, Lynn DJ, Tilley WD, Horvath LG, Centenera MM, Swinnen JV (2020) Lipidomic profiling of clinical prostate cancer reveals targetable alterations in membrane lipid composition. bioRxiv 2020.10.27.356634. https://doi.org/10.1101/2020.10.27.356634

37. Dawaliby R, Trubbia C, Delporte C, Noyon C, Ruysschaert J-M, Van Antwerpen P, Govaerts C (2016) Phosphatidylethanolamine Is a Key Regulator of Membrane Fluidity in Eukaryotic Cells. J Biol Chem 291:3658–3667. https://doi.org/10.1074/jbc.M115.706523

38. Bompard J, Rosso A, Brizuela L, Mebarek S, Blum LJ, Trunfio-Sfarghiu A-M, Lollo G, Granjon T, Girard-Egrot A, Maniti O (2020) Membrane Fluidity as a New Means to Selectively Target Cancer Cells with Fusogenic Lipid Carriers. Langmuir 36:5134–5144. https://doi.org/10.1021/acs.langmuir.0c00262

39. Rysman E, Brusselmans K, Scheys K, Timmermans L, Derua R, Munck S, Van Veldhoven PP, Waltregny D, Daniëls VW, Machiels J, Vanderhoydonc F, Smans K, Waelkens E, Verhoeven G, Swinnen JV (2010) De novo Lipogenesis Protects Cancer Cells from Free Radicals and Chemotherapeutics by Promoting Membrane Lipid Saturation. Cancer Res 70:8117–8126

40. Schug ZT, Peck B, Jones DT, Zhang Q, Grosskurth S, Alam IS, Goodwin LM, Smethurst E, Mason S, Blyth K, McGarry L, James D, Shanks E, Kalna G, Saunders RE, Jiang M, Howell M, Lassailly F, Thin MZ, Spencer-Dene B, Stamp G, van den Broek NJF, Mackay G, Bulusu V, Kamphorst JJ, Tardito S, Strachan D, Harris AL, Aboagye EO, Critchlow SE, Wakelam MJO, Schulze A, Gottlieb E (2015) Acetyl-CoA Synthetase 2 Promotes Acetate Utilization and Maintains Cancer Cell Growth under Metabolic Stress. Cancer Cell 27:57–71. https://doi.org/10.1016/j.ccell.2014.12.002

41. Lisec J, Jaeger C, Zaidi N (2018) Cancer cell lipid class homeostasis is altered under nutrient-deprivation but stable under hypoxia. bioRxiv 382457. https://doi.org/10.1101/382457

42. Jarc E, Petan T (2019) Lipid Droplets and the Management of Cellular Stress. Yale J Biol Med 92:435–452

43. Wolins NE, Quaynor BK, Skinner JR, Schoenfish MJ, Tzekov A, Bickel PE (2005) S3-12, Adipophilin, and TIP47 package lipid in adipocytes. J Biol Chem 280:19146–19155. https://doi.org/10.1074/jbc.M500978200

44. Schott MB, Weller SG, Schulze RJ, Krueger EW, Drizyte-Miller K, Casey CA, McNiven MA (2019) Lipid droplet size directs lipolysis and lipophagy catabolism in hepatocytes. J Cell Biol 218:3320–3335. https://doi.org/10.1083/jcb.201803153

45. Abramczyk H, Surmacki J, Kopeć M, Olejnik AK, Lubecka-Pietruszewska K, Fabianowska-Majewska K (2015) The role of lipid droplets and adipocytes in cancer. Raman imaging of cell cultures: MCF10A, MCF7, and MDA-MB-231 compared to adipocytes in cancerous human breast tissue. Analyst 140:2224–2235. https://doi.org/10.1039/C4AN01875C

46. Li X, Li Y, Jiang M, Wu W, He S, Chen C, Qin Z, Tang BZ, Mak HY, Qu JY (2019) Quantitative Imaging of Lipid Synthesis and Lipolysis Dynamics in *Caenorhabditis elegans* by Stimulated Raman Scattering Microscopy. Anal Chem 91:2279–2287. https://doi.org/10.1021/acs.analchem.8b04875

47. Paar M, Jüngst C, Steiner NA, Magnes C, Sinner F, Kolb D, Lass A, Zimmermann R, Zumbusch A, Kohlwein SD, Wolinski H (2012) Remodeling of Lipid Droplets during

Lipolysis and Growth in Adipocytes. J Biol Chem 287:11164–11173. https://doi.org/10.1074/jbc.M111.316794

48. Sun X, Wang M, Wang M, Yu X, Guo J, Sun T, Li X, Yao L, Dong H, Xu Y (2020) Metabolic Reprogramming in Triple-Negative Breast Cancer. Front Oncol 10:. https://doi.org/10.3389/fonc.2020.00428

49. Petővári G, Dankó T, Tőkés A-M, Vetlényi E, Krencz I, Raffay R, Hajdu M, Sztankovics D, Németh K, Vellai-Takács K, Jeney A, Kulka J, Sebestyén A (2020) In Situ Metabolic Characterisation of Breast Cancer and Its Potential Impact on Therapy. Cancers 12:2492. https://doi.org/10.3390/cancers12092492

50. Lampa M, Arlt H, He T, Ospina B, Reeves J, Zhang B, Murtie J, Deng G, Barberis C, Hoffmann D, Cheng H, Pollard J, Winter C, Richon V, Garcia-Escheverria C, Adrian F, Wiederschain D, Srinivasan L (2017) Glutaminase is essential for the growth of triple-negative breast cancer cells with a deregulated glutamine metabolism pathway and its suppression synergizes with mTOR inhibition. PLOS ONE 12:e0185092. https://doi.org/10.1371/journal.pone.0185092

51. Jung SM, Hung C-M, Hildebrand SR, Sanchez-Gurmaches J, Martinez-Pastor B, Gengatharan JM, Wallace M, Mukhopadhyay D, Martinez Calejman C, Luciano AK, Hsiao W-Y, Tang Y, Li H, Daniels DL, Mostoslavsky R, Metallo CM, Guertin DA (2019) Non-canonical mTORC2 Signaling Regulates Brown Adipocyte Lipid Catabolism through SIRT6-FoxO1. Mol Cell 75:807-822.e8. https://doi.org/10.1016/j.molcel.2019.07.023

52. Yee LD, Mortimer JE, Natarajan R, Dietze EC, Seewaldt VL (2020) Metabolic Health, Insulin, and Breast Cancer: Why Oncologists Should Care About Insulin. Front Endocrinol 11:58. https://doi.org/10.3389/fendo.2020.00058

53. Sanderson SM, Gao X, Dai Z, Locasale JW (2019) Methionine metabolism in health and cancer: a nexus of diet and precision medicine. Nat Rev Cancer 19:625–637. https://doi.org/10.1038/s41568-019-0187-8

54. Cai H, Dong L, Liu F (2016) Recent Advances in Adipose mTOR Signaling and Function: Therapeutic Prospects. Trends Pharmacol Sci 37:303–317. https://doi.org/10.1016/j.tips.2015.11.011

55. Yoon M-S (2017) The Role of Mammalian Target of Rapamycin (mTOR) in Insulin Signaling. Nutrients 9:. https://doi.org/10.3390/nu9111176

56. Hay N (2011) Interplay between FOXO, TOR, and Akt. Biochim Biophys Acta BBA - Mol Cell Res 1813:1965–1970. https://doi.org/10.1016/j.bbamcr.2011.03.013

57. Shi X, Wang J, Lei Y, Cong C, Tan D, Zhou X (2019) Research progress on the PI3K/AKT signaling pathway in gynecological cancer. Mol Med Rep 19:4529–4535. https://doi.org/10.3892/mmr.2019.10121

58. Kitada M, Xu J, Ogura Y, Monno I, Koya D (2020) Mechanism of Activation of Mechanistic Target of Rapamycin Complex 1 by Methionine. Front Cell Dev Biol 8:. https://doi.org/10.3389/fcell.2020.00715

59. Zhou Y, Zhou Z, Peng J, Loor JJ (2018) Methionine and valine activate the mammalian target of rapamycin complex 1 pathway through heterodimeric amino acid taste receptor (TAS1R1/TAS1R3) and intracellular Ca2+ in bovine mammary epithelial cells. J Dairy Sci 101:11354–11363. https://doi.org/10.3168/jds.2018-14461

60. Costantino A, Milazzo G, Giorgino F, Russo P, Goldfine ID, Vigneri R, Belfiore A (1993) Insulin-resistant MDA-MB231 human breast cancer cells contain a tyrosine kinase inhibiting activity. Mol Endocrinol Baltim Md 7:1667–1676. https://doi.org/10.1210/mend.7.12.8145772

61. Gupta C, Tikoo K (2013) High glucose and insulin differentially modulates proliferation in MCF-7 and MDA-MB-231 cells. J Mol Endocrinol 51:119–129. https://doi.org/10.1530/JME-13-0062

62. Wanders D, Hobson K, Ji X (2020) Methionine Restriction and Cancer Biology. Nutrients 12:. https://doi.org/10.3390/nu12030684

63. Borrego SL, Fahrmann J, Datta R, Stringari C, Grapov D, Zeller M, Chen Y, Wang P, Baldi P, Gratton E, Fiehn O, Kaiser P (2016) Metabolic changes associated with methionine stress sensitivity in MDA-MB-468 breast cancer cells. Cancer Metab 4:9. https://doi.org/10.1186/s40170-016-0148-6

64. Jeon H, Kim JH, Lee E, Jang YJ, Son JE, Kwon JY, Lim T-G, Kim S, Park JHY, Kim J-E, Lee KW (2016) Methionine deprivation suppresses triple-negative breast cancer metastasis in vitro and in vivo. Oncotarget 7:67223–67234. https://doi.org/10.18632/oncotarget.11615

65. Morén B, Fryklund C, Stenkula K (2020) Surface-associated lipid droplets: an intermediate site for lipid transport in human adipocytes? Adipocyte 9:636–648. https://doi.org/10.1080/21623945.2020.1838684

66. Covington JD, Johannsen DL, Coen PM, Burk DH, Obanda DN, Ebenezer PJ, Tam CS, Goodpaster BH, Ravussin E, Bajpeyi S (2017) Intramyocellular Lipid Droplet Size Rather Than Total Lipid Content is Related to Insulin Sensitivity After 8 Weeks of Overfeeding. Obes Silver Spring Md 25:2079–2087. https://doi.org/10.1002/oby.21980

67. DeBose-Boyd RA, Ye J (2018) SREBPs in Lipid Metabolism, Insulin Signaling, and Beyond. Trends Biochem Sci 43:358–368. https://doi.org/10.1016/j.tibs.2018.01.005

68. Borrego SL, Fahrmann J, Hou J, Lin D-W, Tromberg BJ, Fiehn O, Kaiser P (2021) Lipid remodeling in response to methionine stress in MDA-MBA-468 triple-negative breast cancer cells. J Lipid Res 62:100056. https://doi.org/10.1016/j.jlr.2021.100056

69. Murata Y, Watanabe T, Sato M, Momose Y, Nakahara T, Oka S, Iwahashi H (2003) Dimethyl Sulfoxide Exposure Facilitates Phospholipid Biosynthesis and Cellular Membrane Proliferation in Yeast Cells. J Biol Chem 278:33185–33193. https://doi.org/10.1074/jbc.M300450200

70. Zou K, Rouskin S, Dervishi K, McCormick MA, Sasikumar A, Deng C, Chen Z, Kaeberlein M, Brem RB, Polymenis M, Kennedy BK, Weissman JS, Zheng J, Ouyang Q, Li H (2020) Life span extension by glucose restriction is abrogated by methionine supplementation:

Cross-talk between glucose and methionine and implication of methionine as a key regulator of life span. Sci Adv 6:eaba1306. https://doi.org/10.1126/sciadv.aba1306

71. Weber FL, Veach GL, Friedman DW (1981) Effects of insulin and glucagon on the uptake of amino acids from arterial blood by canine ileum. Dig Dis Sci 26:113–118. https://doi.org/10.1007/BF01312226

72. Hou J, Williams J, Botvinick EL, Potma EO, Tromberg BJ (2018) Visualization of Breast Cancer Metabolism Using Multimodal Nonlinear Optical Microscopy of Cellular Lipids and Redox State. Cancer Res 78:2503–2512. https://doi.org/10.1158/0008-5472.CAN-17-2618

73. Souba WW (1993) Glutamine and cancer. Ann Surg 218:715–728

74. Wise DR, Thompson CB (2010) Glutamine Addiction: A New Therapeutic Target in Cancer. Trends Biochem Sci 35:427–433. https://doi.org/10.1016/j.tibs.2010.05.003

75. Charidemou E, Ashmore T, Li X, McNally BD, West JA, Liggi S, Harvey M, Orford E, Griffin JL (2019) A randomized 3-way crossover study indicates that high-protein feeding induces de novo lipogenesis in healthy humans. JCI Insight 4:e124819. https://doi.org/10.1172/jci.insight.124819

76. Muthusamy T, Cordes T, Handzlik MK, You L, Lim EW, Gengatharan J, Pinto AFM, Badur MG, Kolar MJ, Wallace M, Saghatelian A, Metallo CM (2020) Serine restriction alters sphingolipid diversity to constrain tumour growth. Nature 586:790–795. https://doi.org/10.1038/s41586-020-2609-x

77. MDA-MB-231 (ATCC® HTB-26™). https://www.atcc.org/products/all/htb-26.aspx#culturemethod

78. Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M (2015) Raman spectroscopy of lipids: a review. J Raman Spectrosc 46:4–20. https://doi.org/10.1002/jrs.4607

79. Jamieson LE, Li A, Faulds K, Graham D Ratiometric analysis using Raman spectroscopy as a powerful predictor of structural properties of fatty acids. R Soc Open Sci 5:181483. https://doi.org/10.1098/rsos.181483

80. Deevska GM, Nikolova-Karakashian MN (2017) The expanding role of sphingolipids in lipid droplet biogenesis. Biochim Biophys Acta BBA - Mol Cell Biol Lipids 1862:1155–1165. https://doi.org/10.1016/j.bbalip.2017.07.008

81. Li Y, Zhang W, Fung AA, Shi L (2022) DO-SRS imaging of diet regulated metabolic activities in Drosophila during aging processes. Aging Cell n/a:e13586. https://doi.org/10.1111/acel.13586

82. (2018) The assembly of lipid droplets and their roles in challenged cells. EMBO J 37:e98947. https://doi.org/10.15252/embj.201898947

83. Benador IY, Veliova M, Liesa M, Shirihai OS (2019) Mitochondria Bound to Lipid Droplets: Where mitochondrial dynamics regulate lipid storage and utilization. Cell Metab 29:827–835. https://doi.org/10.1016/j.cmet.2019.02.011

84. Cui L, Liu P (2020) Two Types of Contact Between Lipid Droplets and Mitochondria. Front Cell Dev Biol 8:1589. https://doi.org/10.3389/fcell.2020.618322

85. Kitt JP, Bryce DA, Minteer SD, Harris JM (2017) Raman Spectroscopy Reveals Selective Interactions of Cytochrome c with Cardiolipin That Correlate with Membrane Permeability. J Am Chem Soc 139:3851–3860. https://doi.org/10.1021/jacs.7b00238

86. Sato ET, Martinho H (2018) First-principles calculations of Raman vibrational modes in the fingerprint region for connective tissue. Biomed Opt Express 9:1728. https://doi.org/10.1364/BOE.9.001728

87. Farber C, Li J, Hager E, Chemelewski R, Mullet J, Rogachev AYu, Kurouski D (2019) Complementarity of Raman and Infrared Spectroscopy for Structural Characterization of Plant Epicuticular Waxes. ACS Omega 4:3700–3707. https://doi.org/10.1021/acsomega.8b03675

88. Khalid M, Bora T, Ghaithi AA, Thukral S, Dutta J (2018) Raman Spectroscopy detects changes in Bone Mineral Quality and Collagen Cross-linkage in Staphylococcus Infected Human Bone. Sci Rep 8:9417. https://doi.org/10.1038/s41598-018-27752-z

89. Da Silva E, Bresson S, Rousseau D (2009) Characterization of the three major polymorphic forms and liquid state of tristearin by Raman spectroscopy. Chem Phys Lipids 157:113–119. https://doi.org/10.1016/j.chemphyslip.2008.11.002

90. Podsednik A, Jacob A, Li LZ, Xu HN (2020) Relationship between Optical Redox Status and Reactive Oxygen Species in Cancer Cells. React Oxyg Species Apex NC 9:95–108

91. Ostrander JH, McMahon CM, Lem S, Millon SR, Brown JQ, Seewaldt VL, Ramanujam N (2010) Optical redox ratio differentiates breast cancer cell lines based on estrogen receptor status. Cancer Res 70:4759–4766. https://doi.org/10.1158/0008-5472.CAN-09-2572

92. Yun Y-H, Bin J, Liu D-L, Xu L, Yan T-L, Cao D-S, Xu Q-S (2019) A hybrid variable selection strategy based on continuous shrinkage of variable space in multivariate calibration. Anal Chim Acta 1058:58–69. https://doi.org/10.1016/j.aca.2019.01.022

93. Hotamisligil GS (2006) Inflammation and metabolic disorders. Nature 444:860–867. https://doi.org/10.1038/nature05485

94. Adams WR, Mehl B, Leiser E, Wang M, Patton S, Throckmorton GA, Jenkins JL, Ford JB, Gautam R, Brooker J, Jansen ED, Mahadevan-Jansen A (2020) Multimodal Nonlinear Optical and Thermal Imaging Platform for Label-Free Characterization of Biological Tissue. Biophysics

95. O'Malley J, Kumar R, Kuzmin A, Pliss A, Yadav N, Balachandar S, Wang J, Attwood K, Prasad PN, Chandra D (2017) Lipid quantification by Raman microspectroscopy as a potential biomarker in prostate cancer. Cancer Lett 397:52–60. https://doi.org/10.1016/j.canlet.2017.03.025

96. Zhang L, Li C, Peng D, Yi X, He S, Liu F, Zheng X, Huang WE, Zhao L, Huang X (2022) Raman spectroscopy and machine learning for the classification of breast cancers.

Spectrochim Acta A Mol Biomol Spectrosc 264:120300. https://doi.org/10.1016/j.saa.2021.120300

97. Kingma DP, Ba J (2017) Adam: A Method for Stochastic Optimization. ArXiv14126980 Cs

98. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830

99. Velioglu SD, Ercioglu E, Temiz HT, Velioglu HM, Topcu A, Boyaci IH (2016) Raman Spectroscopic Barcode Use for Differentiation of Vegetable Oils and Determination of Their Major Fatty Acid Composition. J Am Oil Chem Soc 93:627–635. https://doi.org/10.1007/s11746-016-2808-7

100. Wong MWK, Braidy N, Pickford R, Vafaee F, Crawford J, Muenchhoff J, Schofield P, Attia J, Brodaty H, Sachdev P, Poljak A (2019) Plasma lipidome variation during the second half of the human lifespan is associated with age and sex but minimally with BMI. PloS One 14:e0214141. https://doi.org/10.1371/journal.pone.0214141

101. Han X, Gross RW (2005) Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. Mass Spectrom Rev 24:367–412. https://doi.org/10.1002/mas.20023

102. Jové M, Naudí A, Gambini J, Borras C, Cabré R, Portero-Otín M, Viña J, Pamplona R (2017) A Stress-Resistant Lipidomic Signature Confers Extreme Longevity to Humans. J Gerontol Ser A 72:30–37. https://doi.org/10.1093/gerona/glw048

103. Montoliu I, Scherer M, Beguelin F, DaSilva L, Mari D, Salvioli S, Martin F-PJ, Capri M, Bucci L, Ostan R, Garagnani P, Monti D, Biagi E, Brigidi P, Kussmann M, Rezzi S, Franceschi C, Collino S (2014) Serum profiling of healthy aging identifies phospho- and sphingolipid species as markers of human longevity. Aging 6:9–25. https://doi.org/10.18632/aging.100630

104. Hammond GRV, Schiavo G, Irvine RF (2009) Immunocytochemical techniques reveal multiple, distinct cellular pools of PtdIns4P and PtdIns(4,5)P(2). Biochem J 422:23–35. https://doi.org/10.1042/BJ20090428

105. Baron CL, Malhotra V (2002) Role of diacylglycerol in PKD recruitment to the TGN and protein transport to the plasma membrane. Science 295:325–328. https://doi.org/10.1126/science.1066759

106. Fung AA, Shi L (2020) Mammalian cell and tissue imaging using Raman and coherent Raman microscopy. WIREs Syst Biol Med 12:e1501. https://doi.org/10.1002/wsbm.1501

107. Shi L, Fung AA, Zhou A (2021) Advances in stimulated Raman scattering imaging for tissues and animals. Quant Imaging Med Surg 11:1078101–1071101. https://doi.org/10.21037/qims-20-712

108. Oh S, Lee C, Yang W, Li A, Mukherjee A, Basan M, Ran C, Yin W, Tabin CJ, Fu D, Xie XS, Kirschner MW (2022) Protein and lipid mass concentration measurement in tissues by

stimulated Raman scattering microscopy. Proc Natl Acad Sci 119:e2117938119. https://doi.org/10.1073/pnas.2117938119

109. Zhang J, Zhao J, Lin H, Tan Y, Cheng J-X (2020) High-Speed Chemical Imaging by Dense-Net Learning of Femtosecond Stimulated Raman Scattering. J Phys Chem Lett. https://doi.org/10.1021/acs.jpclett.0c01598

110. Huang K-C, Li J, Zhang C, Tan Y, Cheng J-X (2020) Multiplex Stimulated Raman Scattering Imaging Cytometry Reveals Lipid-Rich Protrusions in Cancer Cells under Stress Condition. iScience 23:100953. https://doi.org/10.1016/j.isci.2020.100953

111. Gaifulina R, Maher AT, Kendall C, Nelson J, Rodriguez-Justo M, Lau K, Thomas GM (2016) Label-free Raman spectroscopic imaging to extract morphological and chemical information from a formalin-fixed, paraffin-embedded rat colon tissue section. Int J Exp Pathol 97:337–350. https://doi.org/10.1111/iep.12194

112. Martin M, Dotti CG, Ledesma MD (2010) Brain cholesterol in normal and pathological aging. Biochim Biophys Acta 1801:934–944. https://doi.org/10.1016/j.bbalip.2010.03.011

113. Ji M, Arbel M, Zhang L, Freudiger CW, Hou SS, Lin D, Yang X, Bacskai BJ, Xie XS (2018) Label-free imaging of amyloid plaques in Alzheimer's disease with stimulated Raman scattering microscopy. Sci Adv 4:eaat7715. https://doi.org/10.1126/sciadv.aat7715

114. Li S, Luo Z, Zhang R, Xu H, Zhou T, Liu L, Qu J (2021) Distinguishing Amyloid β-Protein in a Mouse Model of Alzheimer's Disease by Label-Free Vibrational Imaging. Biosensors 11:365. https://doi.org/10.3390/bios11100365

115. Chau AH, Motz JT, Gardecki JA, Waxman S, Bouma BE, Tearney GJ (2008) Fingerprint and high-wavenumber Raman spectroscopy in a human-swine coronary xenograft in vivo. J Biomed Opt 13:040501. https://doi.org/10.1117/1.2960015

116. Schlame M, Greenberg ML (2017) Biosynthesis, remodeling and turnover of mitochondrial cardiolipin. Biochim Biophys Acta Mol Cell Biol Lipids 1862:3–7. https://doi.org/10.1016/j.bbalip.2016.08.010

117. Acoba MG, Senoo N, Claypool SM (2020) Phospholipid ebb and flow makes mitochondria go. J Cell Biol 219:e202003131. https://doi.org/10.1083/jcb.202003131

118. Kawasaki K, Kuge O, Chang SC, Heacock PN, Rho M, Suzuki K, Nishijima M, Dowhan W (1999) Isolation of a chinese hamster ovary (CHO) cDNA encoding phosphatidylglycerophosphate (PGP) synthase, expression of which corrects the mitochondrial abnormalities of a PGP synthase-defective mutant of CHO-K1 cells. J Biol Chem 274:1828–1834. https://doi.org/10.1074/jbc.274.3.1828

119. Jacobson J, Duchen MR, Heales SJR (2002) Intracellular distribution of the fluorescent dye nonyl acridine orange responds to the mitochondrial membrane potential: implications for assays of cardiolipin and mitochondrial mass. J Neurochem 82:224–233. https://doi.org/10.1046/j.1471-4159.2002.00945.x

120. Lake BB, Menon R, Winfree S, Hu Q, Ferreira RM, Kalhor K, Barwinska D, Otto EA, Ferkowicz M, Diep D, Plongthongkum N, Knoten A, Urata S, Naik AS, Eddy S, Zhang B,

Wu Y, Salamon D, Williams JC, Wang X, Balderrama KS, Hoover P, Murray E, Vijayan A, Chen F, Waikar SS, Rosas S, Wilson FP, Palevsky PM, Kiryluk K, Sedor JR, Toto RD, Parikh C, Kim EH, Macosko EZ, Kharchenko PV, Gaut JP, Hodgin JB, Eadon MT, Dagher PC, El-Achkar TM, Zhang K, Kretzler M, Jain S, Consortium for the K (2021) An atlas of healthy and injured cell states and niches in the human kidney. 2021.07.28.454201

121. Baek J, He C, Afshinnia F, Michailidis G, Pennathur S (2022) Lipidomic approaches to dissect dysregulated lipid metabolism in kidney disease. Nat Rev Nephrol 18:38–55. https://doi.org/10.1038/s41581-021-00488-2

122. Vaziri ND (2016) HDL abnormalities in nephrotic syndrome and chronic kidney disease. Nat Rev Nephrol 12:37–47. https://doi.org/10.1038/nrneph.2015.180

123. Yang H, Fogo AB, Kon V (2016) Kidneys: Key Modulators of HDL Levels and Function. Curr Opin Nephrol Hypertens 25:174–179. https://doi.org/10.1097/MNH.0000000000000217

124. Srivastava SP, Shi S, Koya D, Kanasaki K (2014) Lipid mediators in diabetic nephropathy. Fibrogenesis Tissue Repair 7:12. https://doi.org/10.1186/1755-1536-7-12

125. Reidy K, Kang HM, Hostetter T, Susztak K (2014) Molecular mechanisms of diabetic kidney disease. J Clin Invest 124:2333–2340. https://doi.org/10.1172/JCI72271

126. Gibellini F, Smith TK (2010) The Kennedy pathway-De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. IUBMB Life n/a-n/a. https://doi.org/10.1002/iub.337

127. Haider A, Wei Y-C, Lim K, Barbosa AD, Liu C-H, Weber U, Mlodzik M, Oras K, Collier S, Hussain MM, Dong L, Patel S, Alvarez-Guaita A, Saudek V, Jenkins BJ, Koulman A, Dymond MK, Hardie RC, Siniossoglou S, Savage DB (2018) PCYT1A Regulates Phosphatidylcholine Homeostasis from the Inner Nuclear Membrane in Response to Membrane Stored Curvature Elastic Stress. Dev Cell 45:481-495.e8. https://doi.org/10.1016/j.devcel.2018.04.012

128. Di Pardo A, Amico E, Basit A, Armirotti A, Joshi P, Neely MD, Vuono R, Castaldo S, Digilio AF, Scalabrì F, Pepe G, Elifani F, Madonna M, Jeong SK, Park B-M, D'Esposito M, Bowman AB, Barker RA, Maglione V (2017) Defective Sphingosine-1-phosphate metabolism is a druggable target in Huntington's disease. Sci Rep 7:5280. https://doi.org/10.1038/s41598-017-05709-y

129. Matthäus C, Krafft C, Dietzek B, Brehm BR, Lorkowski S, Popp J (2012) Noninvasive Imaging of Intracellular Lipid Metabolism in Macrophages by Raman Microscopy in Combination with Stable Isotopic Labeling. Anal Chem 84:8549–8556. https://doi.org/10.1021/ac3012347

130. Ledeen RW, Wu G (2006) Sphingolipids of the nucleus and their role in nuclear signaling. Biochim Biophys Acta BBA - Mol Cell Biol Lipids 1761:588–598. https://doi.org/10.1016/j.bbalip.2006.04.010

131. Dominguez G, Maddelein M-L, Pucelle M, Nicaise Y, Maurage C-A, Duyckaerts C, Cuvillier O, Delisle M-B (2018) Neuronal sphingosine kinase 2 subcellular localization is altered in

Alzheimer's disease brain. Acta Neuropathol Commun 6:25. https://doi.org/10.1186/s40478-018-0527-z

132. Fung AA, Hoang K, Zha H, Chen D, Zhang W, Shi L (2022) Imaging Sub-Cellular Methionine and Insulin Interplay in Triple Negative Breast Cancer Lipid Droplet Metabolism. Front Oncol 12:

133. Lin H, Lee HJ, Tague N, Lugagne J-B, Zong C, Deng F, Shin J, Tian L, Wong W, Dunlop MJ, Cheng J-X (2021) Microsecond fingerprint stimulated Raman spectroscopic imaging by ultrafast tuning and spatial-spectral learning. Nat Commun 12:3052. https://doi.org/10.1038/s41467-021-23202-z

134. Wang F, Chen X, Cheng H, Song L, Liu J, Caplan S, Zhu L, Wu JY (2021) MICAL2PV suppresses the formation of tunneling nanotubes and modulates mitochondrial trafficking. EMBO Rep 22:e52006. https://doi.org/10.15252/embr.202052006

135. Bigio EH, Mishra M, Hatanpaa KJ, White CL, Johnson N, Rademaker A, Weitner BB, Deng H-X, Dubner SD, Weintraub S, Mesulam M (2010) TDP-43 pathology in primary progressive aphasia and frontotemporal dementia with pathologic Alzheimer disease. Acta Neuropathol (Berl) 120:43–54. https://doi.org/10.1007/s00401-010-0681-2

136. Wang P, Deng J, Dong J, Liu J, Bigio EH, Mesulam M, Wang T, Sun L, Wang L, Lee AY-L, McGee WA, Chen X, Fushimi K, Zhu L, Wu JY (2019) TDP-43 induces mitochondrial damage and activates the mitochondrial unfolded protein response. PLoS Genet 15:e1007947. https://doi.org/10.1371/journal.pgen.1007947

137. Orringer DA, Pandian B, Niknafs YS, Hollon TC, Boyle J, Lewis S, Garrard M, Hervey-Jumper SL, Garton HJL, Maher CO, Heth JA, Sagher O, Wilkinson DA, Snuderl M, Venneti S, Ramkissoon SH, McFadden KA, Fisher-Hubbard A, Lieberman AP, Johnson TD, Xie XS, Trautman JK, Freudiger CW, Camelo-Piragua S (2017) Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. Nat Biomed Eng 1:1–13. https://doi.org/10.1038/s41551-016-0027

138. Sampaio JL, Gerl MJ, Klose C, Ejsing CS, Beug H, Simons K, Shevchenko A (2011) Membrane lipidome of an epithelial cell line. Proc Natl Acad Sci U S A 108:1903–1907. https://doi.org/10.1073/pnas.1019267108

139. Ejsing CS, Sampaio JL, Surendranath V, Duchoslav E, Ekroos K, Klemm RW, Simons K, Shevchenko A (2009) Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. Proc Natl Acad Sci U S A 106:2136–2141. https://doi.org/10.1073/pnas.0811700106

140. Surma MA, Herzog R, Vasilj A, Klose C, Christinat N, Morin-Rivron D, Simons K, Masoodi M, Sampaio JL (2015) An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. Eur J Lipid Sci Technol 117:1540–1549. https://doi.org/10.1002/ejlt.201500145

141. Liebisch G, Binder M, Schifferer R, Langmann T, Schulz B, Schmitz G (2006) High throughput quantification of cholesterol and cholesteryl ester by electrospray ionization tandem mass spectrometry (ESI-MS/MS). Biochim Biophys Acta 1761:121–128. https://doi.org/10.1016/j.bbalip.2005.12.007

142. Herzog R, Schuhmann K, Schwudke D, Sampaio JL, Bornstein SR, Schroeder M, Shevchenko A (2012) LipidXplorer: A Software for Consensual Cross-Platform Lipidomics. PLOS ONE 7:e29851. https://doi.org/10.1371/journal.pone.0029851

143. Alicic RZ, Rooney MT, Tuttle KR (2017) Diabetic Kidney Disease: Challenges, Progress, and Possibilities. Clin J Am Soc Nephrol CJASN 12:2032–2045. https://doi.org/10.2215/CJN.11491116

144. Reidy K, Kang HM, Hostetter T, Susztak K (2014) Molecular mechanisms of diabetic kidney disease. J Clin Invest 124:2333–2340. https://doi.org/10.1172/JCI72271

145. Herman-Edelstein M, Scherzer P, Tobar A, Levi M, Gafter U (2014) Altered renal lipid metabolism and renal lipid accumulation in human diabetic nephropathy. J Lipid Res 55:561–572. https://doi.org/10.1194/jlr.P040501

146. Tellez D, Balkenhol M, Otte-Holler I, van de Loo R, Vogels R, Bult P, Wauters C, Vreuls W, Mol S, Karssemeijer N, Litjens G, van der Laak J, Ciompi F (2018) Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2018.2820199

147. Mercan C, Mooij GCAM, Tellez D, Lotz J, Weiss N, van Gerven M, Ciompi F (2020) Virtual Staining for Mitosis Detection in Breast Histopathology. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). pp 1770–1774

148. Li Y, Bagheri P, Chang P, Zeng A, Hao J, Fung A, Wu JY, Shi L (2022) Direct Imaging of Lipid Metabolic Changes in Drosophila Ovary During Aging Using DO-SRS Microscopy. https://doi.org/10.3389/fragi.2021.819903

149. Jefferson JA, Shankland SJ, Pichler RH (2008) Proteinuria in diabetic kidney disease: A mechanistic viewpoint. Kidney Int 74:22–36. https://doi.org/10.1038/ki.2008.128

150. Dalla Vestra M, Saller A, Mauer M, Fioretto P (2001) Role of mesangial expansion in the pathogenesis of diabetic nephropathy. J Nephrol 14 Suppl 4:S51-57

151. Tung C-W, Hsu Y-C, Shih Y-H, Chang P-J, Lin C-L (2018) Glomerular mesangial cell and podocyte injuries in diabetic nephropathy. Nephrology 23:32–37. https://doi.org/10.1111/nep.13451

152. Wahl P, Ducasa GM, Fornoni A (2016) Systemic and renal lipids in kidney disease development and progression. Am J Physiol - Ren Physiol 310:F433–F445. https://doi.org/10.1152/ajprenal.00375.2015

153. Chen S, Tseng C-H (2013) Dyslipidemia, Kidney Disease, and Cardiovascular Disease in Diabetic Patients. Rev Diabet Stud RDS 10:88–100. https://doi.org/10.1900/RDS.2013.10.88

154. Abrass CK (2004) Cellular Lipid Metabolism and the Role of Lipids in Progressive Renal Disease. Am J Nephrol 24:46–53. https://doi.org/10.1159/000075925

155. Sasaki T, Tsuboi N, Haruhara K, Okabayashi Y, Kanzaki G, Koike K, Kobayashi A, Yamamoto I, Ogura M, Yokoo T (2018) Bowman Capsule Volume and Related Factors in Adults With Normal Renal Function. Kidney Int Rep 3:314–320. https://doi.org/10.1016/j.ekir.2017.10.007

156. Haruhara K, Tsuboi N, Sasaki T, Amano H, Tanaka M, Koike K, Kanzaki G, Okabayashi Y, Miyazaki Y, Ogura M, Yokoo T (2019) Volume Ratio of Glomerular Tufts to Bowman Capsules and Renal Outcomes in Nephrosclerosis. Am J Hypertens 32:45–53. https://doi.org/10.1093/ajh/hpy147

157. Caramori ML, Kim Y, Huang C, Fish AJ, Rich SS, Miller ME, Russell G, Mauer M (2002) Cellular Basis of Diabetic Nephropathy: 1. Study Design and Renal Structural-Functional Relationships in Patients With Long-Standing Type 1 Diabetes. Diabetes 51:506–513. https://doi.org/10.2337/diabetes.51.2.506

158. Fioretto P, Steffes MW, Mauer M (1994) Glomerular Structure in Nonproteinuric IDDM Patients With Various Levels of Albuminuria. Diabetes 43:1358–1364. https://doi.org/10.2337/diab.43.11.1358

159. Gundersen HJ, Osterby R (1981) Optimizing sampling efficiency of stereological studies in biology: or “do more less well!” J Microsc 121:65–73. https://doi.org/10.1111/j.1365-2818.1981.tb01199.x

160. Lane PH, Steffes MW, Mauer SM (1992) Estimation of glomerular volume: A comparison of four methods. Kidney Int 41:1085–1089. https://doi.org/10.1038/ki.1992.165

161. Torkamani N, Jerums G, Crammer P, Skene A, Power DA, Panagiotopoulos S, Clarke M, MacIsaac RJ, Ekinci EI (2019) Three Dimensional Glomerular Reconstruction: A Novel Approach to Evaluate Renal Microanatomy in Diabetic Kidney Disease. Sci Rep 9:1829. https://doi.org/10.1038/s41598-019-38646-z

162. Thomas MC, Brownlee M, Susztak K, Sharma K, Jandeleit-Dahm KAM, Zoungas S, Rossing P, Groop P-H, Cooper ME (2015) Diabetic kidney disease. Nat Rev Dis Primer 1:1–20. https://doi.org/10.1038/nrdp.2015.18

163. Najafian B, Basgen JM, Mauer M (2002) Estimating Mean Glomerular Volume Using Two Arbitrary Parallel Sections. J Am Soc Nephrol 13:2697–2705. https://doi.org/10.1097/01.ASN.0000033381.53882.25

164. Ligresti G, Nagao RJ, Xue J, Choi YJ, Xu J, Ren S, Aburatani T, Anderson SK, MacDonald JW, Bammler TK, Schwartz SM, Muczynski KA, Duffield JS, Himmelfarb J, Zheng Y (2016) A Novel Three–Dimensional Human Peritubular Microvascular System. J Am Soc Nephrol 27:2370–2381. https://doi.org/10.1681/ASN.2015070747

165. Srivastava SP, Shi S, Koya D, Kanasaki K (2014) Lipid mediators in diabetic nephropathy. Fibrogenesis Tissue Repair 7:12. https://doi.org/10.1186/1755-1536-7-12

166. Vaziri ND (2016) HDL abnormalities in nephrotic syndrome and chronic kidney disease. Nat Rev Nephrol 12:37–47. https://doi.org/10.1038/nrneph.2015.180

167. Fung AA, Hoang K, Zha H, Chen D, Zhang W, Shi L (2022) Imaging Sub-Cellular Methionine and Insulin Interplay in Triple Negative Breast Cancer Lipid Droplet Metabolism. Front Oncol 12:

168. Terasaki M, Brunson JC, Sardi J (2020) Analysis of the three dimensional structure of the kidney glomerulus capillary network. 677864

169. Samuel T, Hoy WE, Douglas-Denton R, Hughson MD, Bertram JF (2007) Applicability of the glomerular size distribution coefficient in assessing human glomerular volume: the Weibel and Gomez method revisited. J Anat 210:578–582. https://doi.org/10.1111/j.1469-7580.2007.00715.x

170. Baues M, Klinkhammer BM, Ehling J, Gremse F, Zandvoort MAMJ van, Reutelingsperger CPM, Daniel C, Amann K, Bábíčková J, Kiessling F, Floege J, Lammers T, Boor P (2020) A collagen-binding protein enables molecular imaging of kidney fibrosis in vivo. Kidney Int 97:609–614. https://doi.org/10.1016/j.kint.2019.08.029

171. Boote C, Palko JR, Sorensen T, Mohammadvali A, Elsheikh A, Komáromy AM, Pan X, Liu J (2016) Changes in posterior scleral collagen microstructure in canine eyes with an ADAMTS10 mutation. Mol Vis 22:503–517

172. Teodoro JS, Rolo AP, Palmeira CM (2013) The NAD ratio redox paradox: why does too much reductive power cause oxidative stress? Toxicol Mech Methods 23:297–302. https://doi.org/10.3109/15376516.2012.759305

173. Lv W, Booz GW, Fan F, Wang Y, Roman RJ (2018) Oxidative Stress and Renal Fibrosis: Recent Insights for the Development of Novel Therapeutic Strategies. Front Physiol 9:

174. Sinha T, Naash MI, Al-Ubaidi MR (2020) Flavins Act as a Critical Liaison Between Metabolic Homeostasis and Oxidative Stress in the Retina. Front Cell Dev Biol 0: https://doi.org/10.3389/fcell.2020.00861

175. (2022) HALO AI

176. Zhang W, Li Y, Fung AA, Li Z, Jang H, Zha H, Chen X, Gao F, Wu JY, Sheng H, Yao J, Skowronska-Krawczyk D, Jain S, Shi L (2022) Multi-Molecular Hyperspectral PRM-SRS Imaging. 2022.07.25.501472

177. Jang H, Li Y, Fung AA, Bagheri P, Hoang K, Skowronska-Krawczyk D, Chen X, Wu JY, Bintu B, Shi L (2023) Super-resolution SRS microscopy with A-PoD. Nat Methods 20:448–458. https://doi.org/10.1038/s41592-023-01779-1

178. Gong L, Lin S, Huang Z (2021) Stimulated Raman Scattering Tomography Enables Label-Free Volumetric Deep Tissue Imaging. Laser Photonics Rev 15:2100069. https://doi.org/10.1002/lpor.202100069