Stimulated Raman scattering: a biophysical perspective for imaging cells and tissues

Thesis by

Kun Miao

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended June 9th, 2023

Kun Miao
ORCID: 0000-0001-6567-3650

ACKNOWLEDGEMENTS

It was a rather overwhelming journey to pursue a PhD degree and to start my scientific career. I couldn't have come even close to this without the help and support of mentors, colleagues, friends, and family. I would like to express my sincere appreciation to several individuals.

I will start by thanking my advisor, Lu Wei. Although I didn't have the most matching background to start with, she put her trust in me and gave me intellectual freedom to explore all the possibilities in science with generous resources. In the early days, Lu taught me how to set up microscopes and needed to bear with my poor optic operations, as I was one of the very first two students she had. Her inspiring scientific vision and ideas always guided me through different projects and difficulties. In addition to science, she cares about me as a person. She gives me advice about my future career and cheers me up when I am in a slump and not productive. I couldn't ask for a better advisor in any facet.

I worked with amazing colleagues in the Wei lab who also contributed tomy science. Chenxi Qian and I worked together a lot and shared similar vision on how the project could be shaped. Li-En Lin is the brave warrior who implemented the machine learning algorithm for our imaging modality with limited related experience. Xiaotian Bi is technically my first mentee yet she became very independent. We often perform trouble shooting together for both her and my projects. She will be taking on the bio-related work that I am leaving behind. Adrain Colazo is the person with whom I worked together more recently on a novel instrumentation for SRS. He will be maintaining our SRS setup as I showed him all the alignment tricks. I would like to thank all of them.

I am thankful for the time I spend with all the Wei Lab members: Jiajun Du, Dongkwan Lee, Haomin Wang, Rahuljeet (RJ) Chadha, Yulu Cao, Ziguang Yang, Noor Naji, Philip Kocheril. You guys make my lab life so much more enjoyable and shape such a supportive environment for me to grow both as a scientist and as a person.

Outside of the Wei Lab, I would like to thank Frances Arnold, Jesse (Jack) Beauchamp, and Mikhail Shapiro for taking time to serve on my committee. Their research guidance and feedback during the committee meetings greatly facilitated my academic growth. I really value my rotation with Frances where I was first exposed to protein engineering. I appreciate Jack attending my presentation for a student-organized seminar right after COVID lockdown. I enjoyed sitting in Mikhail's class and learned about molecular imaging at different levels.

I am very grateful to my undergraduate advisor, Xin Zhang and mentor Yu Liu. They gave me the opportunity to take a peek at scientific research by teaching me all the experiments. I learned not only the experimental techniques but also the approaches they take to tackle different questions. I wouldn't have pursued a PhD degree without their inspiration.

I am fortunate to have made plenty of friends during my time at Caltech: Shumao Zhang, Xinhong Chen, Zhiyang Jin, Yuxing Yao, Zhihao Cui, Guanya Shi, Quan Gan, Weilai Yu, Xiongyi Huang, Xun Wang, Yanzhe Zhu, Yutian Li, Botao Hu, Yukai Liu, Chen Li, Yuelin Shi, Tongxin Li, Xiaozhe Ding, Kai Chen, Zhen Liu, Tian Zeng, Zhiquan Yuan, Yicheng Luo, Yongzhao Guo, Ziyang Qin, Runze Mao Can Li, Alexandra Barth and many more. I really enjoyed chatting with you about research or random gossips in lab hallways, at different restaurants, in the gym, on the basketball court, and at karaoke.

Also want to give a shout out to my old friends who overlapped with me in LA during the past 6 year: Yuan Cui, Zexian Yang, Yipeng Li, Yufei Zhu, Jikai Zhang, Fangzhu Zhao, Changsheng Shen, Yixuan Chen, Yidan Chen, Qing Wang, Jialu Lu, Jing Luo, Ruolin Guo, Jiayi Huang, Yuchen Ye, Yihe Cai. My PhD journey would have been much more challenging without all your support.

Most importantly, I would like to dedicate this thesis to my parents, Yulu Miao and Yinxue Xi. They always gave me unconditional love and supports, yet gave me all the space to make decisions independently. Ten years ago, when they waved to me at the airport, they probably did not expect anything spectacular. Because of COVID related travel restrictions and visa problems, I really owe them plenty of family time.

ABSTRACT

This thesis explores the utilization of Stimulated Raman Scattering (SRS) microscopy as a novel imaging method in the biomedical field, aiming to overcome the limitations associated with traditional fluorescence-based techniques. Given the drawbacks of fluorescence imaging, such as photobleaching, auto-fluorescence, and the complexity of fluorophore labeling, SRS microscopy emerges as a promising solution. The optical imaging contrast in this method originates from bond vibrations of endogenous biomolecules. Grounded in the principle of Raman scattering, SRS amplifies weak spontaneous Raman transitions through stimulated emission, offering a target-specific, high-speed, and label-free imaging modality that can overcome the challenges of traditional bio-imaging techniques.

To tackle the interference from fluorescent proteins when imaging small proteins of interest, we demonstrated a combination of SRS with selective deuterium labeling for visualizing polyQ aggregates in Huntington's disease. We targeted the C-D vibration on deuterated glutamines, which are metabolically enriched in the polyQ sequence. This allowed us to image Huntingtin aggregates without using fluorescent labels. Our method enables, for the first time, the quantification of protein concentrations and compositional analyses of polyQ and non-polyQ proteins within native Huntingtin aggregates. This novel perspective suggests that aggregates have distinct biophysical roles at different stages of aggregation.

In addition to fluorescent proteins, immunofluorescence is the gold standard for visualizing the location and distribution of proteins within cells or tissues. However, the proper delivery of antibodies is slow and labor-intensive. To overcome this issue, we developed a novel method, Vibrational Imaging of Swelled Tissue and Analysis (VISTA), that combines SRS microscopy with sample expansion to enable label-free super-resolution volumetric imaging in tissues. We developed a unique fixation hydrogel chemistry to maximize protein retention, delipidation, and isotropic expansion in tissue samples. By targeting the bond vibrations from endogenous proteins, VISTA bypasses the limitations of antibody labeling and provides an efficient tool for high-throughput imaging that can be scaled to large-volume clinical samples. The addition of image segmentation methods to VISTA equips it with protein-level specificity similar to immunofluorescence. We further used this technique to study protein aggregates, such as amyloid-β plaques in Alzheimer's disease, revealing intricate aggregate structures and polymorphisms absent in conventional fluorescence methods.

Finally, as fluorescent biosensors are indispensable tools for studying intracellular dynamics, we worked on extending the utility of SRS microscopy into the realm of sensing. We employed hydrogen-deuterium exchange on alkyne substrates to develop a Raman-based sensing strategy sensitive to subtle variations in local microenvironments. The rate of hydrogen-deuterium exchange changes under different conditions, and the resulting frequency shift from alkyne to deuterated alkyne is captured by SRS microscopy. This new platform enhances the study of chemical environments in various biological structures, marking a pivotal step in integrating imaging and sensing in biophysical research.

PUBLISHED CONTENT AND CONTRIBUTIONS

Miao, K.; Wei, L. Live-Cell Imaging and Quantification of PolyQ Aggregates by Stimulated Raman Scattering of Selective Deuterium Labeling. *ACS Cent. Sci.* ***6*** (4), 478–486. (2020) doi: 10.1021/acscentsci.9b01196.
K.M. participated in the conceptualization of the project. K.M. designed and performed the experiments and data analysis, and wrote the manuscript.

Qian, C.[#]; **Miao, K**.[#]; Lin, L.-E.; Chen, X.; Du, J.; Wei, L. Super-Resolution Label-Free Volumetric Vibrational Imaging. *Nat. Commun.* ***12*** (1), 3648. (2021) doi: 10.1038/s41467-021-23951-x.
[#] *equal contribution. K.M. participated in the conceptualization of the project. K.M. designed and performed the experiments and data analysis, and participated in writing the manuscript.*

Bi, X.[*]; **Miao, K**.[*]; Wei, L. Alkyne-Tagged Raman Probes for Local Environmental Sensing by Hydrogen–Deuterium Exchange. *J. Am. Chem. Soc.* ***144*** (19), 8504–8514. (2022) doi: 10.1021/jacs.2c01991.
[#] *equal contribution. K.M. participated in the conceptualization of the project. K.M. designed and performed the experiments and data analysis and participated in writing of the manuscript.*

Miao, K.; Lin, L.-E.; Qian, C.; Wei, L. Label-Free Super-Resolution Imaging Enabled by Vibrational Imaging of Swelled Tissue and Analysis. *J. Vis. Exp.* No. 183, e63824. (2022) doi: 10.3791/63824.
K.M. participated in the conceptualization of the project. K.M. performed the experiments, analyzed the data, and wrote the manuscript.

Lin, L.-E.; **Miao, K.**; Qian, C.; Wei, L. High Spatial-Resolution Imaging of Label-Free in vivo Protein Aggregates by VISTA. *Analyst*, **146**, 4135-4145. (2021) doi: 10.1039/D1AN00060H.

K.M. participated in the conceptualization, experiments, and data analysis of the study.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Stimulated Raman scattering: a bio-imaging technology requires no fluorescent labels

In the vast realm of biology and medicine, our understanding continues to be transformed by remarkable advancements in imaging technology. Various imaging methods have been developed to acquire information at both holistic and fine-grained levels under differing circumstances. For instance, magnetic resonance imaging (MRI) allows imaging of upregulated blood flow in the brain during human activities with millimeter resolution[1], while ultrasound imaging is widely employed for detecting tumors and diagnosing conditions related to various organs[2]. While both ultrasound and MRI focus on capturing macroscopic events at the animal and organ levels, optical imaging methods offer unprecedented microscopic perspectives for visualizing molecular events, dynamic processes, and cellular interactions at the tissue and subcellular levels.

Among all optical imaging methods, fluorescence microscopy has garnered the most attention and become widely used due to its exceptional photophysical properties[3]. For example, the Stokes shift between excitation and emission spectra and the high extinction cross sections of fluorescence processes allow high signal-to-noise imaging and hence single-molecule detection of different fluorophores[3]. The development of fluorescent proteins provides a universal approach to tagging fluorophores onto various proteins within living organisms[4]. Moreover, with the advent of novel optical instrumentation and computational methods, super-resolution techniques, such as stimulated emission depletion (STED) microscopy[5] and single-molecule localization microscopy (SMLM)[6], have pushed the boundaries of resolution beyond the diffraction limit, enabling visualization of structures at the nanometer scale[5,6].

While fluorophore labeling is essential for achieving high specificity and sensitivity in fluorescence imaging, it also becomes the pitfall that causes disadvantages. First, fluorescence processes that involve pumping fluorophores into triplet excited states inevitably lead to intersystem crossing and photobleaching[7]. Additionally, the presence of auto-fluorescence from biological samples can overshadow the weakened fluorescence signal. Hence, fluorescence imaging may not be as quantitative and could have deteriorating imaging quality over time. Second, labeling fluorophores to specific targets is not always straightforward. Fluorescent proteins require genetic manipulation, which is not always compatible with histology and tissue samples. On the other hand, immunofluorescence, relying on antibody-epitope recognition for specificity, can bypass the need for genetic modifications. However, it may suffer from inhomogeneous antibody labeling due to the poor permeability of macromolecules[8]. Moreover, specific antibodies for certain protein targets are not always readily available[9].

Considering these challenges, an alternative approach emerges: how can we utilize endogenous molecules as optical imaging contrast to reveal information elusive to fluorescence-based methods? As a chemist, I envision that different chemical bond vibrations can provide target-specific imaging without the need for fluorophores, since chemical bonds are smaller than fluorophores themselves. In this thesis work, my focus will be on harnessing Raman scattering as an imaging contrast to overcome current limitations in bio-imaging.

The discovery of spontaneous Raman scattering, where the scattered photons by molecules reflect bond vibrations, dates back to 1928[10]. While Raman spectroscopy quickly advanced for acquiring molecular information in homogeneous solutions, it was not feasible as an imaging method due to low sensitivity and slow acquisition speeds. Consequently, its application in biomedical studies was hindered, as high temporal and spatial resolutions are crucial for imaging dynamic events in live samples. However, the accidental observation of stimulated Raman scattering (SRS) enables the breakthrough[11]. This coherent two-photon Raman process utilized the principle of stimulated emission amplification, significantly enhancing the weak spontaneous Raman transitions and rendering it suitable for biomedical imaging. With the demonstration of shot-noise-limited high-speed SRS microscopy[12], we can first time image the distribution of bond vibration in biological systems in relevant time scale.

As a result, SRS microscopy offers fast image acquisition based on the concentration of chemical bond vibrations, making it an ideal imaging modality for unveiling molecular information in biological samples. It also holds the potential to overcome limitations associated with conventional fluorescence microscopy. In fact, SRS microscopy has already been implemented in various biomedical applications, including label-free detection of biomolecules[13], real-time SRS imaging on live animals at video-rate speeds[14], and tumor imaging[15], among others. The versatility and capabilities of SRS microscopy have solidified its position as an emerging tool in the field of biomedical research, paving the way for non-invasive and high-resolution imaging of complex biological systems.

1.2 Fluorescent protein labeling influence the biophysical properties of polyQ aggregates

One major hurdle for fluorescence imaging is the size of fluorescent labels. Fluorophores are relatively small, yet they require a self-labeling protein tag of around 30 kDa for protein specificity[16]. Fluorescent proteins, on the other hand, have a size of 27 kDa, which is not negligible when the target protein is small in size. The biophysical properties of fluorescent proteins, such as self-aggregation, could override those of the small proteins of interest simply due to their larger size.

To address this challenge, we have developed an imaging method that combines stimulated Raman scattering (SRS) with selective deuterium labeling to study polyQ aggregates formed in Huntington's disease. The disease-causing Huntingtin protein exon1 fragment, even with a pathogenic polyQ expansion, is almost half the size of fluorescent proteins tag for imaging[17]. In Chapter 2, I demonstrated that native Huntingtin protein aggregates can be imaged without the need for fluorescent proteins by targeting the C-D vibrations in deuterated glutamine, which is preferentially enriched in the polyQ sequence of Huntingtin aggregates[18]. My data revealed different aggregation propensities between Huntingtin aggregates with and without fluorescent protein labeling. The linear dependence of the Raman signal enabled the first quantification of polyQ protein concentration within Huntingtin aggregates in live cells. Additionally, SRS microscopy can also image CH vibrations from endogenous proteins, allowing us to analyze the compositional differences between polyQ proteins and non-polyQ cytosolic proteins in Huntingtin aggregates of different sizes. This analysis suggests that early-stage aggregates (small size) sequester cytosolic proteins, while late-stage aggregates (large size) preferentially contain polyQ proteins.

By combining selective deuterium labeling with SRS microscopy, we have overcome the limitations of fluorescent labeling and gained new biophysical insights into Huntingtin aggregates. We are currently expanding this technique to study Huntingtin aggregates in neuronal cells.

1.3 Label-free super-resolution volumetric tissue imaging bypasses the limitations of antibody delivery

Having developed a novel SRS method to address fluorescent protein limitations in live cell imaging, our focus now shifts to the present challenges of immunofluorescence in tissue studies. The binding of fluorophore-labeled antibodies to target epitopes is crucial for protein-specific staining and imaging. However, the delivery of these macromolecular antibodies into deep tissues is a formidable obstacle[19]. The process of antibody labeling is labor-intensive, requiring weeks or even months, and often results in inhomogeneous labeling, leading to imaging artifacts[19]. Moreover, photobleaching and fluorophore decay limit the applicability of immunofluorescence for repeated imaging of target structures in scarce clinical samples.

In Chapter 3, I describe a method that leverages sample expansion strategy with SRS microscopy to tackle obstacles associated with immunofluorescence in tissues[20]. I envisioned that the bond vibrations from endogenous proteins in tissues could serve as a label-free imaging contrast for SRS microscopy. By eliminating the need for antibodies, our method circumvents the labeling limitations for immunofluorescence. Nevertheless, the intense lipid content in tissues not only worsens the imaging contrast for protein vibrations (protein CH_3 and lipid CH_2 have similar vibrational frequency), but also impedes penetration of photons and limits the imaging depth. Therefore, we have worked out a sample hydrogel chemistry that enables clearing of lipids and retention of proteins onto the hydrogel matrixes. The resulting hydrogels are isotropically expanded and imaged by SRS microscopy targeting endogenous protein vibrations.

We termed this technique VISTA (Vibrational Imaging of Swelled Tissue and Analysis). VISTA enables the first ever label-free super-resolution imaging with a resolution down to 78 nm by physically expands the sample specimens. The volumetric imaging capacity is also a rare case for super-resolution imaging modalities. Furthermore, we implemented a machine-learning based image segmentation model that confers proteins specificity similar immunofluorescence. The segmentations are accomplished based on the features of different proteins in VISTA images. Therefore, VISTA provides a high-throughput label-free imaging method that is scalable to large-volume clinical samples and holds potential for novel biological discoveries.

In neurodegenerative diseases, protein aggregates such as polyQ aggregates, Tau, amyloid-β, and α-synuclein exhibit aggregate polymorphism[21]. These aggregates assume different shapes and sizes, including soluble oligomers, protofibrils, and insoluble fibrils, and they exert distinct toxic effects on cells[22]. This phenomenon also poses challenges when studying different populations of aggregates using fluorescence. The binding sites of fluorophores and the designated epitopes can be absent in certain forms of aggregates. SRS microscopy, targeting endogenous protein vibrations and metabolically incorporated C-D bonds, is well-equipped to provide a comprehensive view of protein aggregates when dyes and antibodies label only subpopulations of them.

In Chapter 4, I detail my efforts to study protein aggregates in cells and tissues using the VISTA technology[23,24]. We first demonstrate high-resolution imaging of intracellular polyQ aggregates associated with Huntington's disease, resolving fibrillary extrusions of 130 nm on the periphery of the aggregates. The volumetric tissue imaging capability of VISTA allows us to investigate extracellular amyloid-β plaques, associated with Alzheimer's disease, in mouse brain tissue. We observed plaques with different fine filament structures and core-shell arrangements, which are indistinguishable in conventional fluorescence imaging. These aggregate polymorphisms are accompanied by astrocytes with varying levels of activation, suggesting that amyloid-β plaques within the brain may have different functional implications.

1.4 Extend SRS microscopy to sensing applications with hydrogen-deuterium exchange on terminal alkynes

In addition to imaging, small-molecule and protein-based fluorescence biosensors serve as powerful tools for real-time monitoring of biological processes[25]. These tools offer insights into the dynamic inner workings of cells, including protein-protein interactions and calcium influxes[26]. Our goal is to equip SRS microscopy, which originates from spectroscopy methods, with sensing capabilities that can probe events orthogonal to those detected by fluorescence biosensors, thus offering new biophysical insights.

In Chapter 5, I detail our exploration of using classical hydrogen-deuterium exchange on novel alkyne substrates as our sensing strategy[27]. This approach is feasible because deuterated alkynes exhibit distinct vibrational frequencies due to altered relative mass. We characterized the hydrogen-deuterium exchange rates for different alkyne probes and established the relationship between exchange rates and Hammett values. We then demonstrated that alkyne-HDX is highly sensitive to various DNA structures and can effectively detect in situ structural changes in DNA resulting from UV-induced damage. Furthermore, this strategy has proven to be an effective solution for distinguishing subtle pD variations in living cells. As a result, we have developed an SRS-based sensing platform that probes local microenvironments, providing an additional dimension of information to imaging.

1.5 Reference

1. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
2. Fenster, A., Downey, D. B. & Cardinal, H. N. Three-dimensional ultrasound imaging. *Phys. Med. Biol.* **46**, R67 (2001).
3. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).
4. Tsien, R. Y. Constructing and exploiting the fluorescent protein paintbox (Nobel Lecture). *Angew. Chem. Int. Ed.* **48**, 5612–5626 (2009).
5. Willig, K. I., Harke, B., Medda, R. & Hell, S. W. STED microscopy with continuous wave beams. *Nat. Methods* **4**, 915–918 (2007).
6. Lelek, M. *et al.* Single-molecule localization microscopy. *Nat. Rev. Methods. Primers* **1**, 1–27 (2021).
7. Yeow, E. K. L., Melnikov, S. M., Bell, T. D. M., De Schryver, F. C. & Hofkens, J. Characterizing the fluorescence intermittency and photobleaching kinetics of dye molecules immobilized on a glass surface. *J. Phys. Chem. A* **110**, 1726–1734 (2006).
8. Murray, E. *et al.* Simple, scalable proteomic imaging for high-dimensional profiling of intact systems. *Cell* **163**, 1500–1514 (2015).
9. Liebmann, T. *et al.* Three-dimensional study of Alzheimer's Disease hallmarks using the iDISCO Clearing Method. *Cell Rep.* **16**, 1138–1152 (2016).
10. Raman, C. V. & Krishnan, K. S. A new type of secondary Radiation. *Nature* **121**, 501-502 (1928).
11. Woodbury, E.J., Ng, W.K. Ruby laser operation in the near IR, *Proc. IRE* **50**, 2347–2348 (1962).
12. Freudiger, C. W. *et al.* Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **322**, 1857–1861 (2008).
13. Wang, P., Liu, B., Zhang, D., Belew, M. Y., Tissenbaum, H. A., & Cheng, J.-X., Imaging lipid metabolism in live Caenorhabditis elegans using fingerprint vibrations. *Angew. Chem., Int. Ed.* **53**, 11787 –11792 (2014)
14. Saar, B. G. et al. Video-rate molecular imaging in vivo with stimulated Raman scattering. *Science* **330**, 1368-1370 (2010).
15. Ji, M. et al. Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci. Transl. Med.* **5**, 201ra119-201ra119 (2013).
16. Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).
17. Li, L. *et al.* Real-time imaging of Huntingtin aggregates diverting target search and gene transcription. *eLife* **5**, e17056 (2016).

18. Miao, K. & Wei, L. Live-cell imaging and quantification of PolyQ aggregates by stimulated Raman scattering of selective deuterium labeling. *ACS Cent. Sci.* **6**, 478–486 (2020).
19. Ku, T. *et al.* Elasticizing tissues for reversible shape transformation and accelerated molecular labeling. *Nat. Methods* **17**, 609–613 (2020).
20. Qian, C. *et al.* Super-resolution label-free volumetric vibrational imaging. *Nat. Commun.* **12**, 3648 (2021).
21. Kodali, R. & Wetzel, R. Polymorphism in the intermediates and products of amyloid assembly. *Curr. Opin. Struct. Biol.* **17**, 48–57 (2007).
22. Morley, J. F., Brignull, H. R., Weyers, J. J. & Morimoto, R. I. The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10417–10422 (2002).
23. Lin, L.-E., Miao, K., Qian, C. & Wei, L. High spatial-resolution imaging of label-free in vivo protein aggregates by VISTA. *Analyst* **146**, 4135–4145 (2021).
24. Miao, K., Lin, L.-E., Qian, C., Wei, L. Label-Free super-resolution imaging enabled by Vibrational Imaging of Swelled Tissue and Analysis. *J. Vis. Exp.* No. 183, e63824. (2022)
25. Barak, L. S., Ferguson, S. S. G., Zhang, J. & Caron, M. G. A β-Arrestin/green fluorescent protein biosensor for detecting G Protein-coupled Receptor activation *. *J. Biol. Chem.* **272**, 27497–27500 (1997).
26. Zhang, Y. *et al.* Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**, 884–891 (2023).
27. Bi, X., Miao, K. & Wei, L. Alkyne-Tagged Raman probes for local environmental sensing by Hydrogen–Deuterium exchange. *J. Am. Chem. Soc.* **144**, 8504–8514 (2022).

Chapter 2

Quantitative imaging of native polyQ aggregates by stimulated Raman scattering coupled with selective deuterium labeling

The content of this chapter has been published in:
Miao, K.; Wei, L. Live-Cell Imaging and Quantification of PolyQ Aggregates by Stimulated Raman Scattering of Selective Deuterium Labeling. *ACS Cent. Sci.* **2020**, *6* (4), 478–486. https://doi.org/10.1021/acscentsci.9b01196.

2.1 Summary

Polyglutamine (polyQ) diseases are a group of neurodegenerative disorders, involving the deposition of aggregation-prone proteins with long polyQ expansions. However, the cytotoxic roles of these aggregates remain highly controversial, largely due to a lack of proper tools for quantitative and non-perturbative interrogations. Common methods including *in vitro* biochemical, spectroscopic assays and live-cell fluorescence imaging all suffer from certain limitations. Here, we propose to couple stimulated Raman scattering microscopy with deuterium-labeled glutamine for live-cell imaging, quantification and spectral analysis of native polyQ aggregates with subcellular resolution. First, through the enrichment of deuterated glutamine in the polyQ sequence of mutant Huntingtin (mHtt) exon1 proteins for Huntington's disease, we achieved sensitive and specific SRS imaging of carbon-deuterium bonds (C-D) from aggregates without GFP labeling, which is commonly employed in fluorescence microscopy. We revealed that these aggregates became 1.8-fold denser compared to those with GFP. Second, we performed ratiometric quantifications, which indicate a surprising dependence of protein compositions on aggregation sizes. Our further calculations, for the first time, reported the absolute concentrations for sequestered mHtt and non-mHtt proteins within the same aggregates. Third, we adopted hyperspectral SRS for Raman spectroscopic studies of aggregate structures. By inducing a cellular heat shock response, shown as a potential therapeutic approach for inhibiting aggregate formation, our results imply a possible aggregate intermediate with changed solvation states. Our method may hence readily unveil new features and mechanistic insight of polyQ aggregates and pave the way for comprehensive *in vivo* investigations.

2.2 Introduction

A hallmark of neurodegenerative disorders is the presence of protein aggregates in peripheral nerves[1-4]. Among these disorders are polyglutamine (polyQ) diseases, such as Huntington's disease (HD), which starts with motor symptoms like chorea and followed by memory deficit and depressions[3,4]. The onset of HD has been linked to abnormally expanded CAG trinucleotide repeats that encode polyQ sequence in mutant Hungtingtin (mHtt) proteins. While Q repeats are typically fewer than 37 in healthy humans, they can range from 40 to 250 in Huntington's patients and are consistently found in the protein depositions of HD brain slices by immunohistology[1,3]. However, the pathological roles of polyQ aggregates still remain elusive[2,4-6]. Recent studies suggest that soluble oligomers are cytotoxic by dynamically interacting with cytosolic proteins and triggering apoptosis while aggregates are cyto-protective by sequestering toxic protein oligomers to form stable inclusion bodies[4,7,8]. In contrast, evidence also indicates that toxicity of aggregates arises from depleting functional (e.g., chaperones and transcription factors) and structural (e.g., actin) proteins and impairing cellular organelles (e.g., ribosomes and endoplasmic reticulum)[9-13].

To understand their molecular roles, extensive efforts have been made to investigate the compositions, structures, and kinetics of mHtt aggregates. Conventional biochemical assays and recent quantitative proteomics offer relative protein compositions of the aggregates in reference to the soluble protein pools. However, these methods rely on extensive post-processing such as aggregation purification and solubilization[9,10]. *In vitro* spectroscopic studies including IR[14], UV-resonance Raman[15,16], NMR spectroscopy[17] and fluorescence[18,19] on model peptides provide crucial information, but they are limited to relatively short expansion lengths because of the difficulty in isolating peptides with long Q repeats[14-19]. More importantly, all these *in vitro* studies cannot recapitulate the native aggregation status in live cells. For live-cell studies, fluorescence imaging offers unprecedented spatial and temporal resolution, by fusing fluorescent proteins[20] or self-labeling tags (e.g., HaloTag)[21] to the C-terminus of a mHtt exon1 (ex1) sequence (**Fig. 1a**). The aggregation-prone ex1 fragment, which comprises of a 17 N-terminal sequence, a polyQ tract followed by a proline-rich domain at the C-terminus (**Fig. 1a**), can effectively induce pathological phenotype of HD in transgenic mouse model and humans[4,22]. Compared to mHtt ex1, however, GFP is much larger in size and has a known tendency to oligomerize[23]. This could perturb the aggregation kinetics and conformations, and may contribute to the controversy of reported toxicity. Moreover, because the dense aggregation environment often causes fluorescence quenching[24], fluorescence imaging is not ideal for quantitative analysis of aggregates. It is therefore highly desirable to have a new modality that combines the advantages from *in vitro* investigations and fluorescence imaging while overcoming their limitations.

Here, we report a novel and general platform for live-cell imaging, quantification and spectral analysis of polyQ aggregates by stimulated Raman scattering (SRS) microscopy (**Fig. S1**) of deuterium-labeled glutamine (Gln-d_5) (**Fig. 1b, c**). We first achieved specific and sensitive SRS imaging of native polyQ aggregates without the need of large fusion-proteins. We then developed a ratiometric strategy for quantitative analysis of protein content in aggregates of varying sizes. Compared to existing methods, our modality, for the first time, reports the absolute concentrations of both mHtt and non-mHtt proteins from aggregates in live cells. Applying the hyperspectral SRS (hSRS), we further probed the aggregate structures and aggregate-environment interactions upon an induced heat shock response.

2.3 Results

The Coupling of SRS Microscopy with Deuterated Glutamine is Ideal for Imaging PolyQ Aggregates. Labeling of aggregates is achieved through replacing regular Gln in the medium with Gln-d_5 (**Fig. 1b, c**), which would be metabolically incorporated and enriched into the long polyQ tail of expressed mHtt proteins (**Fig. 1a, b**). Targeting the vibrational frequency of carbon-deuterium bonds (C-D), SRS imaging obtains subcellular mapping of mHtt aggregates in live cells. Our strategy has following advantages. First, the vibrational frequency of C-D in Gln-d_5 (**Fig. 1c**) is in the desired cell-silent region (1800 – 2600 cm^{-1}), providing high imaging specificity without background from endogenous biomolecules. Second, the non-exchangeable labeling on the C-H side chains of Gln offers reliable signals. Third, Gln enrichment in the polyQ region (**Fig. 1a**) yields both high labeling specificity and superb SRS imaging sensitivity. For example, for a widely used mHtt-97Q ex1 protein, Gln accounts for 68% of the ex1 sequence (**Fig. S2**), while the natural occurrence of Gln is only 4.2% in human proteomes[25]. Compared to label-free SRS[26] and SRS imaging with ^{13}C-Phenyalanine[27] or deuterated all essential amino acids[28], selective Gln-d_5 labeling is significantly more specific for imaging polyQ aggregates. Fourth, compared to using alkyne-tagged unnatural amino acids, which only introduces one tag to one copy of a protein[29], multiple Q labeling (e.g., 103Q for mHtt-97Q ex1) has higher sensitivity and requires less sample manipulation. Fifth, imaging wise, compared to spontaneous Raman[30], SRS provides higher detection sensitivity and faster image acquisition with the stimulated emission quantum amplification principle (**Fig. S1**). Compared to Coherent anti-Stokes Raman scattering (CARS)[31], another nonlinear Raman microscopy technique, SRS offers high-fidelity Raman spectra (**Fig. 1c**) and linear concentration dependence (**Fig. 1d**) without non-resonance background.

We first determined our SRS detection limit on Gln-d_5 solution to be 3 mM (**Fig. 1d**, when signal (S)/noise (N)=1) by targeting the C-D peak at 2147 cm^{-1}. Hence, our detection limit is as low as 29 μM for mHtt proteins with a total of 103 Gln in the mHtt-97Q ex1 protein. This strategy offers higher detection sensitivity compared to previously reported 200 μM by SRS imaging of alkyne tags[32]. In cells, however, the detection limit is set by aggregate-signal to cellular-background ratio (S_{Agg}/B_{Cell}). There are two sources for B_{Cell}. First, the newly synthesized proteome incorporates Gln. Since Gln accounts for 4.2 % of human proteome[25], the estimated detectability for mHtt-97Q ex1 is 86 μM when $S_{Agg}/B_{Cell}=1$ (details in SI). Second, the intracellular free Gln pool is about 8 mM[33], slightly lowering the achievable detectability. To minimize this additional background, we replaced the medium with buffer before imaging. Although a detection sensitivity of 86 μM may be too high for imaging normal proteins, it is highly feasible to detect polyQ aggregates down to very small sizes. Our GFP fluorescence analysis by time-lapse tracking of mHtt-97Q-GFP proteins (**Fig. S3a, b**) and Tetramethylrhodamine (TMR) fluorescence analysis of mHtt-97Q-Halo proteins (**Fig. S3c**) suggest that once the new and small aggregates form in cells, the concentration of polyQ protein is already beyond 86 μM to as high as 400 μM.

Imaging and Quantification of mHtt Aggregates with and without GFP Labeling in Live HeLa Cells. We next validated our SRS imaging by transfecting HeLa cells with mHtt-97Q-GFP plasmid (**Figs. 1a and S2**) and culturing them in Gln-d_5 medium. We first conducted parallel SRS and fluorescence imaging on the same set of live cells. The SRS image of C-D enriched aggregate (**Fig. 2a**, 2167 cm^{-1}) agrees well with the fluorescence image of GFP (**Fig. 2a**, fluorescence). A clear off-resonance image demonstrates high SRS imaging quality (**Fig. 2a**, 2035 cm^{-1}). The C-D peak for Gln-d_5 is shifted from 2147 cm^{-1} in solution to 2167 cm^{-1} after being incorporated into cellular proteins, suggesting a change of microenvironment. Indeed, we found that the shifted C-D spectrum has Raman spectral features from both Gln-d_5 solution and solid (peaked at 2167 cm^{-1}) (**Fig. S4**). As a comparison, label-free SRS images at CH_3 (2940 cm^{-1}) and Amide I (1664 cm^{-1}) channels, widely adopted for imaging total proteins[26], show much decreased detection specificity, particularly pronounced for small-size aggregates (**Fig. 2b**, arrow indicated), which are indistinguishable from the nucleoli. Moreover, the high-contrast of C-D image decreases significantly if Gln-d_5 is replaced by non-enriched Leucine-d_{10} (Leu-d_{10}) for labeling (**Fig. 2c**). Our quantification on the average S_{Agg}/B_{Cell} for each channel clearly demonstrated much higher imaging specificity with Gln-d_5 (**Fig. 2d**).

After establishing the feasibility for SRS imaging of Gln-d_5 labeled mHtt-97Q-GFP aggregates, we aimed to image native mHtt-97Q proteins without GFP (**Fig. 3a**). GFP may perturb aggregation formation of mHtt proteins, because it is 238 amino acid (aa) in length, which is about twice as large as mHtt-97Q ex1 with only 152 aa. We successfully imaged aggregates at the same C-D frequency (**Fig. 3b**, CD on and off). Interestingly, these aggregates (**Fig. 3b**, 2167 cm^{-1}) are about twice as brighter than those of similar sizes with GFP labeling (**Fig. 2a**, 2167 cm^{-1}). Because GFP sequence only contains 8 Gln, the C-D intensity should remain approximately unchanged when deleting GFP. We hence reasoned that the detected intensity increase is due to the formation of denser aggregates. A similar phenomenon was recently reported by Cryo-ET with the same mHtt ex1 sequence[11]. To test our theory, we acquired Amide I images to compare total protein concentrations between aggregates with and without GFP. If the density remains unchanged, mHtt-97Q aggregates would have a much lower amide intensity than that of mHtt-97Q-GFP because the GFP sequence contributes significantly to amide signals. We observed similar levels of amide signals between these two types of aggregates (Fig. S5), confirming that the increase in aggregate density makes up for the signal loss of deleting GFP. Moreover, mHtt-97Q aggregates become barely distinguishable from cellular background in CH_3 channel (**Fig. 3b**, 2940 cm^{-1}), indicating that previous CH_3 signals for aggregates were mainly from GFP sequence (**Fig. 2b**, 2940 cm^{-1} and **Fig. S2**). To prove the high biocompatibility of our method, we also demonstrated time-lapse SRS imaging on the same set of live cells to capture the aggregation formation (**Fig. 3c**). To quantify aggregates from multiple experimental replicates, we plotted the average S_{Agg}/B_{Cell} in both C-D and CH channels for mHtt-97Q and mHtt-97Q-GFP aggregates (**Fig. 3d**). In C-D channel, the S_{Agg}/B_{Cell} for mHtt-97Q aggregates is 1.8-fold higher, confirming the formation of 1.8-fold denser aggregates without GFP labeling, as B_{cell} is not affected by the change of plasmids. In C-H channel, the S_{Agg}/B_{Cell} values remain approximately the same. We further confirmed our high detection specificity of polyQ aggregates by inducing the formation of stress granules in normal HeLa cells and confirmed that our SRS imaging with Gln-d_5 labeling does not detect these non-polyQ-aggregate-type of stress granules (**Fig. S6**). Going beyond mHtt-97Q, we imaged and quantitatively analyzed a shorter Q expansion sequence, mHtt-46Q, for which the Q number is just above the disease-prone aggregation threshold (Fig. S7). Similarly, we found that mHtt-46Q aggregates are about 1.6-time denser compared to those formed by mHtt-46Q-GFP. Moreover, we demonstrated our applicability of imaging polyQ aggregates in a stable embryonic stem cell-line[34] (**Fig. S8**), paving the way for more general future applications.

Ratiometric Quantifications for the Dependence of Protein Compositions on Aggregation Sizes. The feature of parallel C-D and CH_3 SRS imaging enables us to perform ratiometric analysis of CH_3-to-CD ratios (CH/CDs) to investigate polyQ aggregate compositions of non-mHtt and mHtt proteins. This is because the CD signals are from Gln-d_5, which mainly originates from mHtt proteins, while the CH_3 signals represent non-Gln aa, which come from the combined sources of non-Gln aa in mHtt proteins and sequestered non-mHtt proteins. We first applied CH/CD quantifications on both mHtt-97Q-GFP and mHtt-97Q aggregates (**Fig. 3e**). Surprisingly, the average CH/CD remains as high as 5.75 for mHtt-97Q aggregates (**Fig. 3e**). If the aggregates were formed primarily by mHtt proteins, the CH/CD should be about 0.5 because mHtt-97Q sequence contains 49 non-Q aa and 103 Q. Similarly, the CH/CD should have decreased by 5 folds from mHtt-97Q-GFP (289 non-Q aa and 112 Q) to mHtt-97Q aggregates (49 non-Q aa and 103 Q) (**Fig. S2**). Instead, the decrease is only 1.5-fold (**Fig. 3e**, 9.79 vs 5.75). Our control experiments confirmed that Gln-d_5

labeling efficiency is close to 100% (i.e., *de novo* glutamine synthesis for regular Gln is about zero, Fig. S9 and S10), indicating all 103 Q in the mHtt sequence are deuterated. Our results hence imply that these aggregates contain a rather high percentage of non-mHtt proteins. Sequestrations and depletions of cellular functional and structural proteins by aggregates have been suggested to be one underlying mechanism of HD cytotoxicity[9-13]. Consequently, we calculated the relative molar percentages for mHtt and non-mHtt proteins in the aggregates. Assuming that the molar percentage of mHtt and non-mHtt proteins in the aggregates is x and 1-x, respectively, with known protein sequences, we can generate an equation with x and (1-x) for the relative ratio between non-Gln aa and Gln-d_5 to represent the measurable of CH/CDs (**Fig. 3e**) (details in SI). We calculated that the average molar percentage of non-mHtt proteins is indeed as high as 54% for an average CH/CD of 5.75.

Now that we have established CH/CDs could serve as a direct indicator of aggregation compositions, we next examined the relationship between CH/CDs and aggregation sizes. Based on the sequestration theory[12,13], we expect that the relationship might offer insight for understanding the molecular cytotoxicity of polyQ aggregates. Interestingly, we observed a negative correlation (Fig. 4a, Pearson's r=-0.56), which indicates that the percentage of sequestered mHtt proteins increases as the aggregates grow. As a comparison, such distinctive correlation by Gln-d_5 labeling is absent when replaced by Leu-d_{10} labeling (Fig. 4b, Pearson's r=-0.11), because mHtt and non-mHtt proteins have similar Leucine abundance. This underscores the importance of specific polyQ labeling to observe such a size-composition correlation.

In addition to calculating molar percentages from CH/CDs, we can also compute absolute concentrations with aggregate intensities in C-D channel and our reference calibration curves (**Fig. 1d**). Such quantification is otherwise highly challenging for existing methods. We selected three representative aggregates with different sizes for calculation (**Fig. 4a**, color-indicated, details in SI and results listed in Table S1). The mHtt protein concentration in the small aggregate (**Fig. 4a**, green dot) ranges from 1.2 - 2.1 mM (**Fig. 4c**, 6.7 μm^2). It increases to 4.3 - 5.1 mM (and 4d, 19.1 μm^2) for the medium aggregate (**Fig. 4a** red dot). The concentration becomes 5.6 - 6.4 mM for the large aggregate (36.5 μm^2, **Fig. 4a**, magenta dot). The upper and lower limits shown are determined by the relative content of newly synthesized and pre-existing non-mHtt proteins sequestered to the aggregates. Surprisingly, we found that while mHtt concentrations increase with aggregate sizes, the concentrations of non-mHtt proteins remain almost the same for small (4.6 mM), medium (4.7 mM) and large (4.2 mM) aggregates (**Table S1**). Our observations therefore suggest that the formation of small aggregates preferentially sequester non-mHtt cytosolic proteins (to as high as 70%). These proteins are likely functional chaperones, ribonucleoproteins and structural proteins[9-13]. As the aggregates become larger, they then sequester more mHtt proteins. This might indicate a cellular rescue mechanism during aggregate growth by clearing toxic mHtt proteins[4,7,8]. Our data further demonstrate that the total protein concentrations of these aggregates fall in the range of 5-10 mM. To the best of our knowledge, our study is the first to quantify the mHtt protein concentrations and molar percentages for aggregates of different sizes in live cells.

Interrogating Aggregate Structures and Aggregate-environment Interactions by Hyperspectral SRS (hSRS) Imaging. One key advantage for *in vitro* spectroscopic studies over live-cell fluorescence imaging is that they provide structural and conformational information of the aggregates[14-19]. For instance, multiple reports suggested the presence of beta-sheet rich structures in mHtt aggregates. However, there is still no consensus[15,18]. We therefore adopted hSRS to analyze mHtt-97Q aggregates in live cells. We started with Amide I band since its Raman vibration is well documented for probing protein secondary structures by vibrational spectroscopy[16]. A recent report by label-free hSRS imaging of Amide I on amyloid plaques in brain tissues revealed a clear 12 cm^{-1} blue shift (resolvable by our spectral resolution of ~12 cm^{-1}), which corresponds to cross-beta sheet structures[35]. Nonetheless, we observed no such difference comparing the Amide I spectrum of mHtt aggregates to that of cellular background (**Fig. 5a**, 1664 cm^{-1}, blue vs red). Similar to a recent study[18], our data suggest that beta-sheet structures might not be enriched in these aggregates. Note that, we detected a spectral distortion upon chemical fixation (**Fig. 5a**, from 1600 to 1640 cm^{-1}, green vs blue and magenta vs red), highlighting the importance of live-sample analysis[35,36]. Similar to Amide I, CH_3 (2940 cm^{-1}) offers no discernable spectral information (**Fig. 5b**), since the signals mainly come sequestered non-mHtt proteins. The aggregates spectra are almost identical to that of nucleoli with abundant proteins and rRNA and the lower shoulder at 2850 cm^{-1}, attributed to CH_2 vibrations, implies reduced lipids in the aggregates.

We next asked whether C-D spectral region could offer insightful structural information with specific aggregate labeling. Interestingly, we observed a clear dip around 2146 cm^{-1} from aggregates compared to that from cellular background (**Fig. 5c**), suggesting a structural or micro-environmental difference between aggregated and cytosolic mHtt proteins. To understand such spectral changes, we fitted the averaged spectra (**Fig. 5c**) by three Lorentzian peaks (**Fig. S11a-b**). Our results indicate that the observed 2146 cm^{-1} dip is resulted from the narrowing and the slight red shift of the shoulder peak at 2130 cm^{-1} (**Fig. S11c-d**). The narrowing of the peak width is likely due to the formation of a more ordered structure upon aggregation[16]. Our results establish that this shoulder peak could serve as a sensitive indicator for the microenvironment of mHtt aggregates. We note that the weak-intensity cellular background structures, including nucleoli and lipid droplets shown in the C-D images, might interfere with the specific analysis of small aggregates with low signals. In this case, C-D spectra generated by hSRS on aggregates, nucleoli and lipid droplets could help unequivocally differentiate these structures (**Fig. 5d**).

As we have proven our technique as an effective method to analyze the microenvironment of native aggregates, we sought to apply it to understanding molecular interactions between polyQ aggregates and heat shock proteins (HSPs), key chaperones for protein folding and inhibiting protein aggregations[37]. This may offer fundamental insight toward therapeutic development for HD[6]. We induced cellular heat shock responses by a small-molecule drug, Geldanamycin (GA), an inhibitor of HSP 90, which has been shown to effectively induce the expression of HSP 40 and 70 for the clearance of polyQ aggregates[36]. We first validated the function of GA by confirming that the number of large aggregates decreases with the treatment of increased GA concentrations[37] (**Fig. S12**). We next resorted to C-D hSRS on mHtt-97Q aggregates. Interestingly, in GA treated samples, we discovered a subset of aggregates with reduced size and intensity that consistently present a varied spectral feature with a lowered shoulder peak at 2132 cm^{-1} compared to that of larger-size aggregates (**Figs. S13a** and **S13b**). Inspired by this spectroscopic observation, we came back to mHtt-97Q-GFP and adopted a correlative fluorescence and hSRS to understand this phenomenon (**Fig. 6a**). We found that such aggregates, when exhibiting the same SRS spectral feature as found for the mHtt-97Q subset (**Fig. 6b** and **S13c**), always co-exist with surrounding diffusive mHtt protein pools (**Fig. 6a**, Fluorescence). Recently, a liquid-to-solid phase transition is reported for the mHtt-97Q-GFP proteins during aggregation formation[38]. We first speculated that our observed shoulder peak decrease is caused by a reverse phase conversion from solid mHtt aggregates toward surrounding liquid pool by GA triggered up-regulation of HSP 40/70. However, the Fluorescence Loss In Photobleaching (FLIP) experiments on the surrounding fluorescent pool did not reveal any fluorescence exchange between the cytosolic and aggregated mHtt proteins (**Fig. S14**), ruling out our phase-transition hypothesis.

Since C-D vibration is sensitive to microenvironment, similar to that from CH[39], we then asked whether such a decrease at 2132 cm^{-1} shoulder peak corresponds to a change of secondary structures when interacting with HSP 40/70 proteins, as suggested by previous *in vitro* EM experiments[40]. To test the hypothesis, we emerged fixed non-GA-treated cells containing mHtt-97Q-GFP aggregates in solvents with descending hydrogen-bonding capacity (**Fig. 6c**, from DPBS buffer (green), to MeOH (yellow) and DMSO (blue)). Interestingly, the 2132 cm^{-1} shoulder increases accordingly (Fig. 6c, vertical line). Quantitative fittings for the averaged spectra with three Lorentzian peaks (**Fig. S15a-c**) further confirmed an increase of ratios between the shoulder (2130 cm^{-1}, peak1) and the major (2169 cm^{-1}, peak2) peaks (**Fig. S15e-f**) from DPBS treated aggregates (**Fig. S15a**) to DMSO treated aggregates (**Fig. S15c**). This indicates the peak1/peak2 ratios may report the local solvation states (e.g., hydrogen bonding) of aggregates. Similar fitting for the averaged spectrum from GA treated aggregates (**Fig. S15d**) showed that it poses the lowest peak1/peak2 ratio (**Fig. S15e-f**), suggesting that the local solvation environment for our captured GA-treated aggregates may be more hydrogen-bonded compared to that for regular cellular proteomes (**Fig. S15d**). Such likely hyper-hydration status might be caused by a partially folded state, a folding intermediate, for aggregates upon interacting with HSP 40/70 in live cells[41]. We further analyzed the CH/CD ratios on these small aggregates in cells with and without induced heat-shock responses. We postulated that if the heat-shock responses recruit extra HSPs to the aggregates, as our hSRS data suggested above, we might detect an increase in CH/CD ratios since the percentage of the non-mHtt proteins is higher in these aggregates. We indeed found that the CH/CD ratios became slightly higher in small aggregates from cells under induced heat-shock responses compared to those in similar-sized aggregates from non-induced cells (**Fig. S16**).

2.4 Conclusions

In summary, we demonstrated the combination of SRS microscopy with Gln-d_5 labeling to be a general platform for sensitive and specific imaging, ratiometric quantification and hyper-spectral analysis of native polyQ aggregates in HD. Our technique provides new and complementary structural and compositional information to our current knowledge of polyQ aggregates. As shown above, our detection sensitivity and specificity are ensured from three aspects. First, the expanded polyQ sequences yield a Gln-d_5 labeling enrichment in disease proteins over the normal proteins with only low Q abundance. Second, the formation of the aggregates is specifically induced by these disease-form polyQ proteins. Together, these ensure a much higher signal from polyQ aggregates compared to that from cellular background. Third, in the case the aggregates are small with low signals, hSRS on aggregates adds an additional layer of spectral specificity, differentiating the aggregates from other puncta-like structures, including nucleoli and lipid droplets.

Our method is also applicable to polyQ expansions of various lengths. In particular, with linear concentration dependence, it is suited for investigating aggregates of extended polyQ construct (e.g., >200Q), which may form rather dense structures and pose challenges to study by other strategies. Our method is also applicable to other polyQ diseases, including spinocerebellar ataxia and spinobulbar muscular atrophy[42]. Other poly-aa diseases[43,44] with poly-Glycine-Alanine (poly-GA), poly-Proline-Arginine (poly-PR), and poly-Proline-Alanine (poly-PA) aggregates, recently reported in ALS/FTD patient brain, could also be investigated by selective deuteration of corresponding aa. To better guide future SRS work on these aggregates, we further calculated the corresponding SRS detection limits for representative, varying lengths of polyQ and poly-aa sequences (**Table S2**). In addition, since Gln can transport across blood-brain barrier[45] and deuterium labeling is minimally invasive, applications to animal models or even to humans may be possible. Moreover, correlative live-cell SRS imaging with Gln-d_5 labeling with recently demonstrated Cryo-ET[11] or quantitative proteomics[12] may offer comprehensive structure-function relationship for native mHtt-97Q aggregates.

2.5 Figures

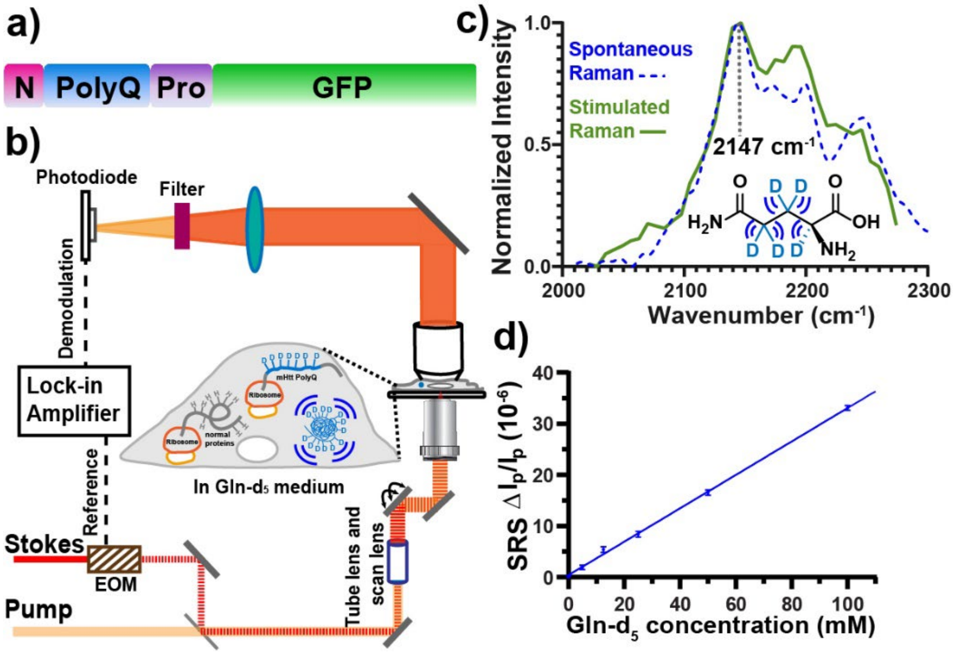


Fig. 1. Experimental scheme for stimulated Raman scattering (SRS) microscopy with deuterated glutamine (Gln or Q) labeling. a) Plasmid construct of a model mutant Huntingtin (mHtt) Exon1 (ex1) protein fused with GFP at C terminus. N: N-terminal 17 aa fragment. PolyQ: poly glutamine region. Pro: proline rich domain. GFP: green fluorescent protein. b) Experimental scheme for SRS imaging of Gln-d_5 labeled polyQ aggregates. c) Spontaneous Raman (blue dashed) and SRS (green) spectra of 60 mM Gln-d_5 solution. d) Linear dependence of SRS signals (at 2,147 cm^{-1}) on Gln-d_5 concentrations under a 50-μs time constant. Error bar: SD.

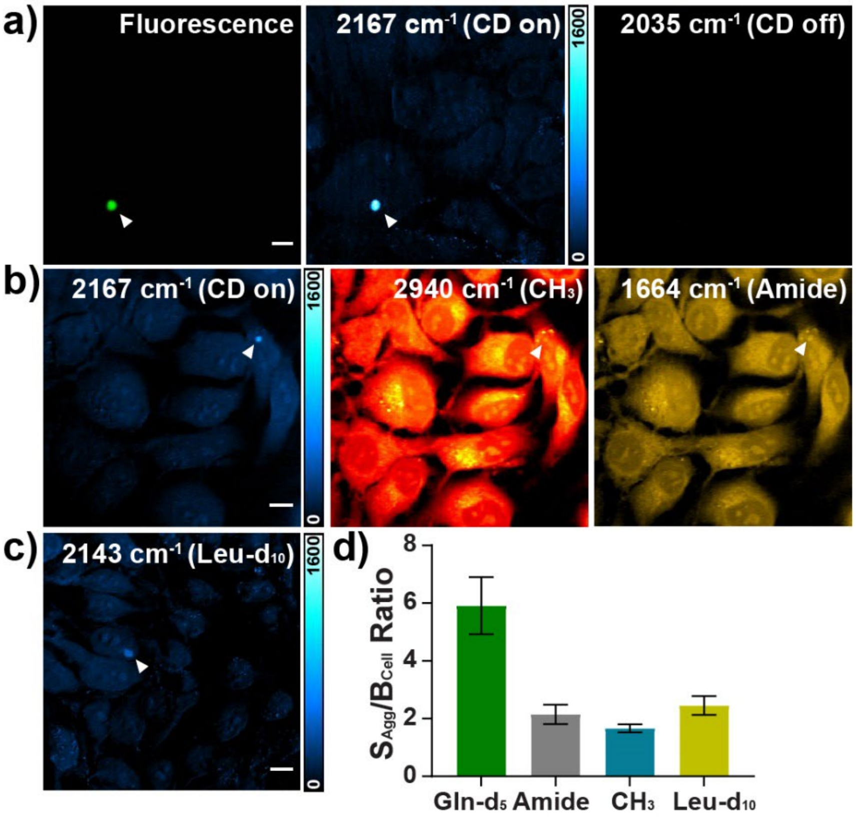


Fig. 2. Live-cell SRS imaging of mHtt-97Q-GFP aggregates with Gln-d$_5$ labeling. a) SRS imaging of mHtt aggregates (arrowheaded, 2167 cm^{-1}, C-D on), validated by fluorescence imaging through GFP (Fluorescence). Off-resonance image at 2035 cm^{-1} shows no signal. b) Live-cell SRS images for an mHtt-97Q-GFP aggregate (arrowheaded) at Gln-d$_5$ (2167 cm^{-1}), CH$_3$ (2940 cm^{-1}) and Amide I (1664 cm^{-1}) channels on the same set of Hela cells. c) SRS imaging of an mHtt-97Q-GFP aggregate at 2143 cm^{-1} by Leucine-d$_{10}$ (Leu-d$_{10}$) labeling. d) Average S_{Agg}/B_{cell} from SRS images of C-D with Gln-d$_5$ labeling (5.75±1.03, n=13); Amide I (2.15±0.34, n=4); CH$_3$ (1.66±0.14, n=10) and C-D with Leu-d$_{10}$ labeling (2.45±0.33, n=10.) Error bar: SD.

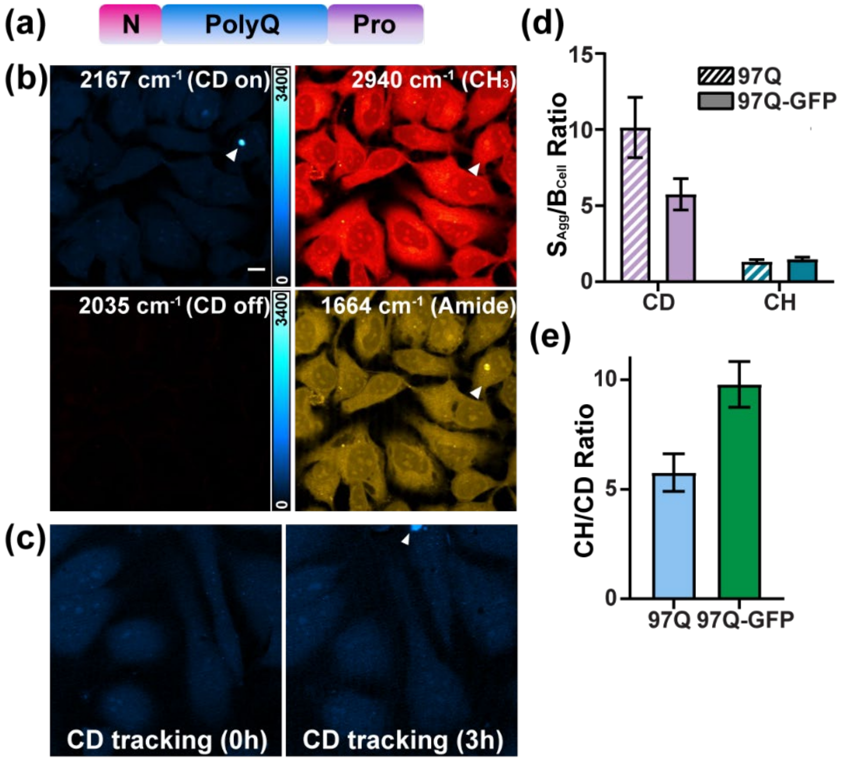


Fig. 3. SRS imaging of mHtt-97Q aggregates without GFP. a) Plasmid construct of mHtt-97Q by deleting GFP sequence. b) SRS images of a Gln-d_5 labeled aggregate at C-D on-resonance (2167 cm^{-1}), C-D off-resonance (2035 cm^{-1}), CH_3 (2940 cm^{-1}) and Amide I (1664 cm^{-1}). c) Time-lapse SRS tracking at C-D channel for capturing aggregation formation on the same set of live cells. d) S_{Agg}/B_{Cell} of mHtt-97Q (97Q) in C-D (CD, 10.14±1.99, n=11) and CH_3 (CH, 1.50±0.12, n=10) channels compared to that for mHtt-97Q-GFP (97Q-GFP, CD: 5.75±1.03, n=13; CH: 1.66±0.14, n=10). e) CH/CDs for 97Q-GFP (9.79±1.05, n=7) and 97Q (5.75±0.86, n=20) aggregates. Scale bar: 10 μm. Error bar: SD.

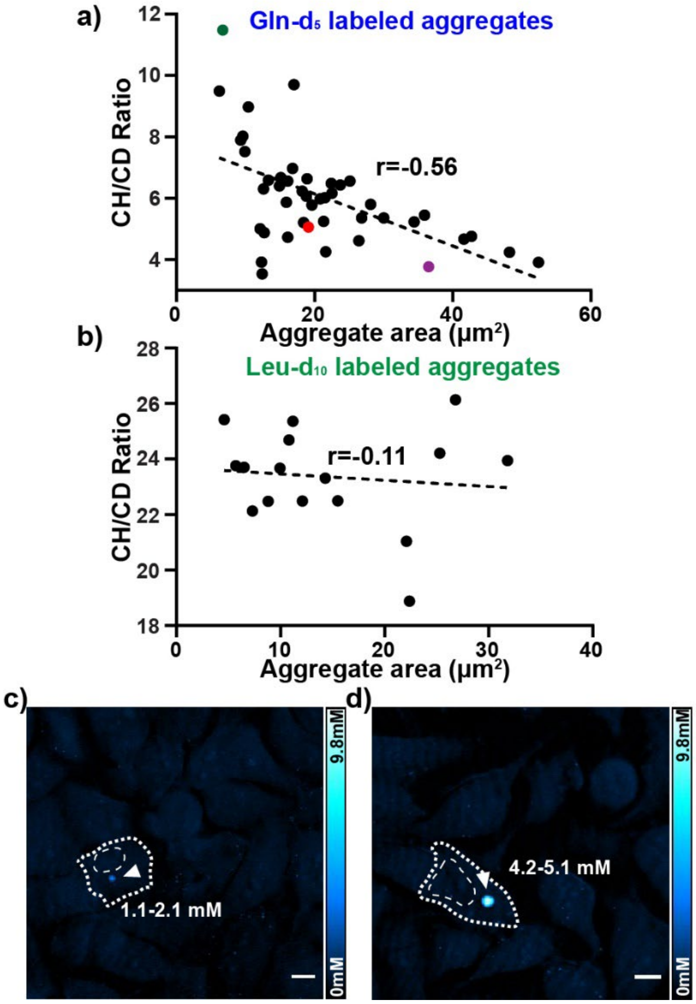


Fig. 4. SRS quantification for mHtt-97Q aggregates of different sizes. a) CH/CDs for Gln-d_5 labeled aggregates present a negative correlation on aggregation areas (The Pearson's coefficient r=-0.56). b) Minimum correlation between CH/CDs for Leu-d_{10} labeled aggregates and aggregation areas (The Pearson's coefficient r=-0.11). c, d) C-D SRS images for representative Gln-d_5 labeled small (c, 6.7 μm^2, green dot in a) and medium (d, 19.1 μm^2, red dot in a) mHtt-97Q aggregates. Calculated mHtt-97Q concentrations for the aggregates are indicated. Cell shapes and nuclei are outlined by white dotted line. Scale bar: 10 μm.

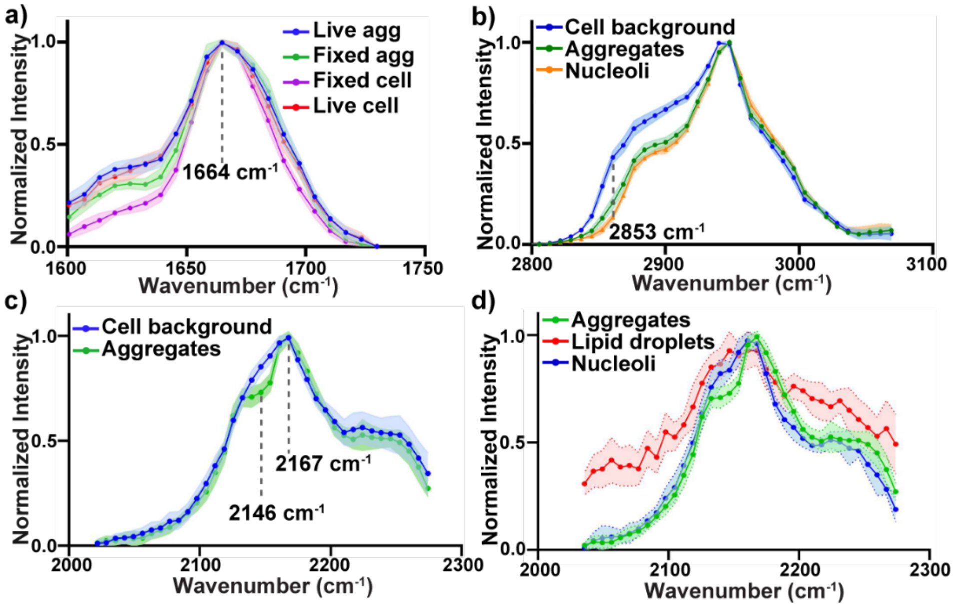


Fig. 5. Hyperspectral SRS analysis on mHtt-97Q aggregates. a) Average Amide I SRS spectra of live aggregates (Live agg, blue, n=14), live cell-background (Live cell, red, n=14), fixed aggregates (Fixed agg green, n=10), and fixed cell-background (Fixed cell, magenta, n=10). b) Live-cell SRS CH_3 spectra of aggregates (green, n=14), cell background (blue, n=16), and nucleoli (orange, n=16). c) Live-cell SRS C-D spectra of Gln-d_5 labeled aggregates (green, n=20) and cell background (blue, n=20). d) Live-cell SRS C-D spectra on small mHtt aggregates (green, n=20), lipid droplets (red, n=10) and nucleoli (blue, n=12) with distinct spectral features. Error bar: SD.

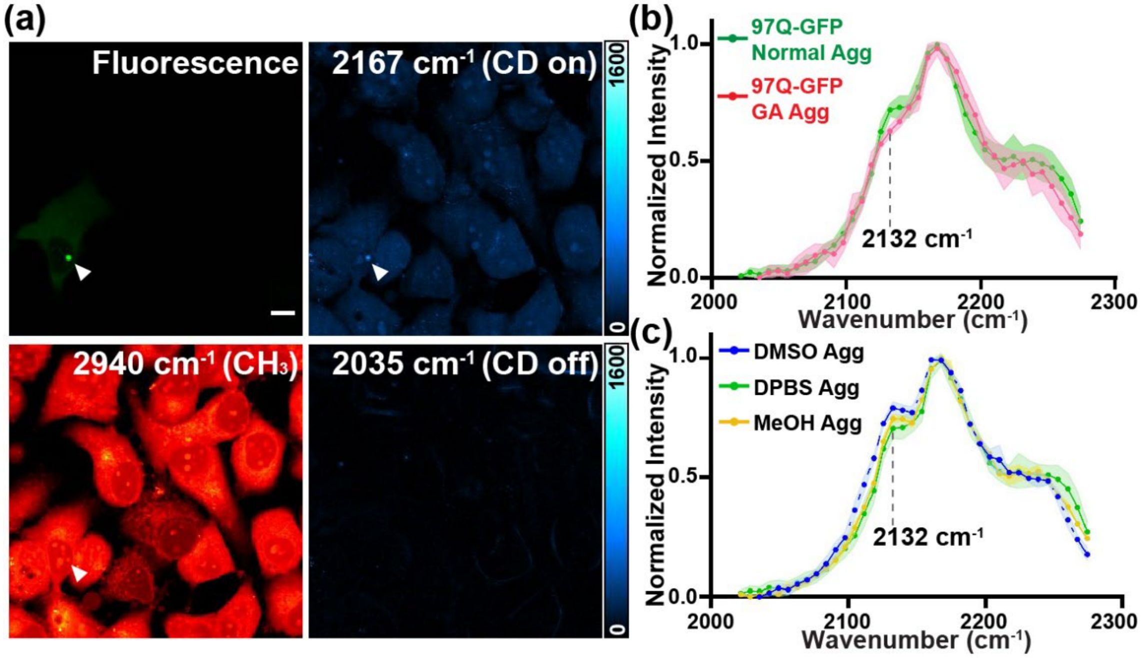


Fig. 6. Hyperspectral SRS study of mHtt aggregations upon cellular heat shock responses. a) Fluorescence (green) and correlative SRS images for mHtt-97Q-GFP aggregates (white arrowheaded) at 2167 cm^{-1} (C-D on-resonance), 2035 cm^{-1} (off-resonance) and 2940 cm^{-1} (CH_3) in live cells after 100 nM Geldanamycin (GA) treatment for 20 h. Scale bar: 10 µm. b) SRS spectra of normal mHtt aggregates (green, n=20) and a subset of GA treated small aggregates (red, n=7). c) SRS spectra of mHtt aggregates in fixed cells in DPBS buffer (green, n=15), MeOH (yellow, n=5) and DMSO (blue, n=6). Error bar: SD.

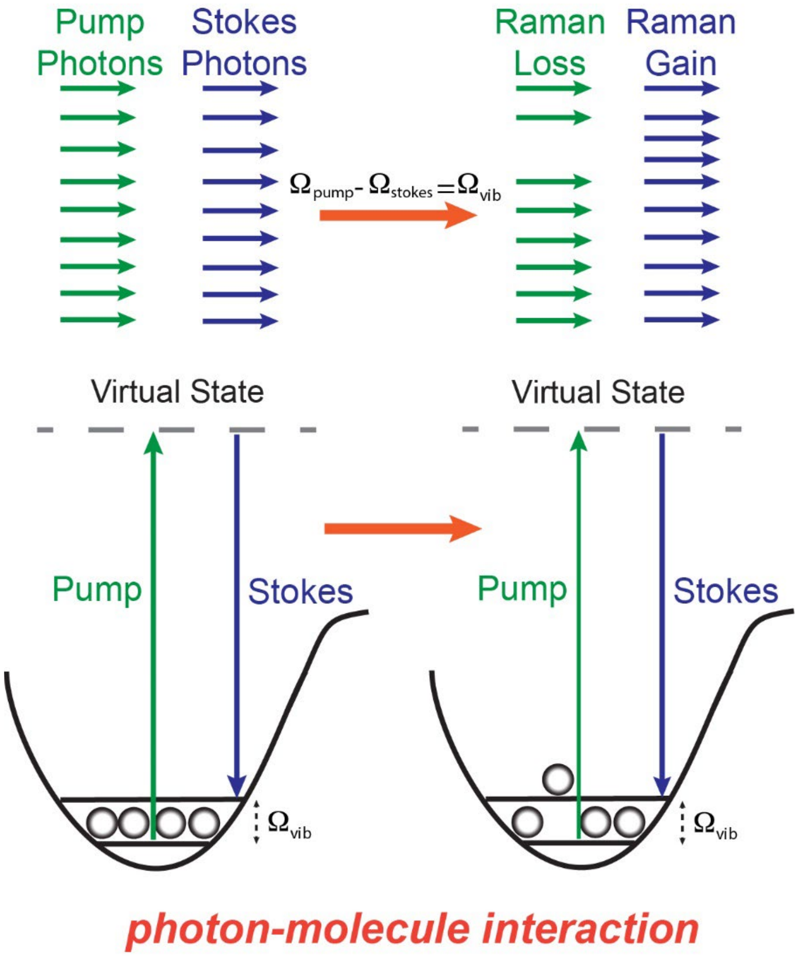
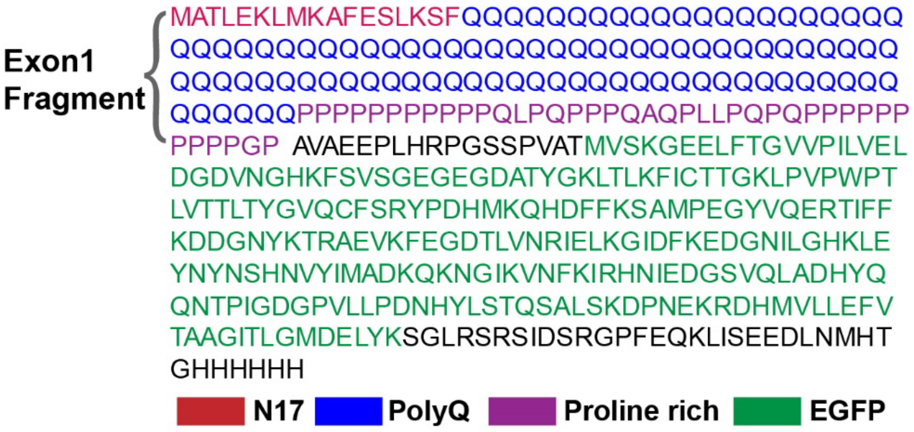


Figure S1. Energy diagram of stimulated Raman scattering (SRS). When the energy difference between the pump (green) and probe (blue) photons is resonant with the vibrational (Raman) transition of a C-D bond. The C-D bond would be excited to the vibrational excited state mediate by a virtual state (i.e., one grey ball, signifying C-D bond, on the left is excited to the upper level shown on the right). For each vibrational excitation of C-D bond, a pump photon is consumed (Raman Loss) and a Stokes photon is created (Raman Gain). The SRS signal is then detected as either Raman Gain or Raman loss signal. In our experiments, we adjust the laser (pump=842.8 nm, stokes=1031.2 nm) to match the vibration frequency of Ω_{vib} =2167 cm^{-1} for C-D bond on Gln-d$_5$ and detect the Raman loss signal at the Pump wavelength.



MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQPPPPPPPPPPPQLPQPPPQAQPLLPQPQPPPPPP
PPPPGP AVAEEPLHRPGSSPVATMVSKGEELFTGVVPILVEL
DGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT
LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFF
KDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLE
YNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ
QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFV
TAAGITLGMDELYKSGLRSRSIDSRGPFEQKLISEEDLNMHT
GHHHHHH

Fig. S2. Detailed amino acid (aa) sequence information of mHtt-97Q-GFP and mHtt-97Q. The mHtt exon1 fragment (highlighted in big parentheses), comprises of N-terminal 17 aa (red), polyQ (blue) and proline rich (purple) domain, is fused with GFP (green) by a linker region (black) followed by C-terminal myc and 6x His-tag (black).

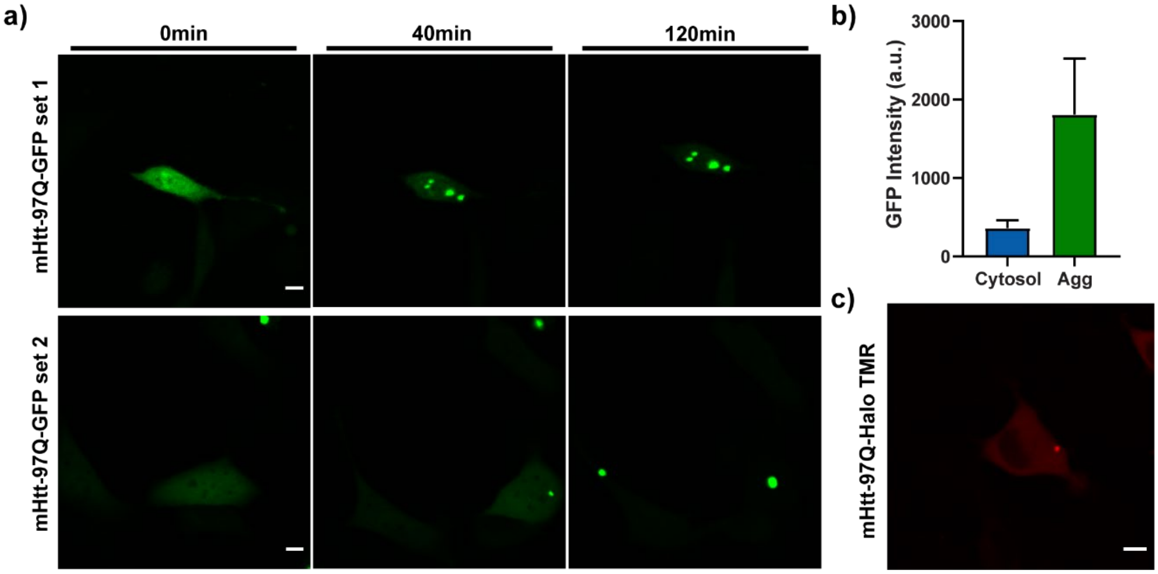


Fig. S3 Time-lapse fluorescence tracking and quantification on aggregates formed by mHtt-97Q-GFP and Tetramethylrhodamine (TMR) labeled mHtt-97Q-Halo proteins. a) Time-lapse fluorescence imaging of mHtt-97Q-GFP captures fast aggregation formation (i.e., at the 40 min) on the same set of cells. b) Quantification of GFP signals for newly formed small aggregates shows a 5-fold fluorescence increase compared to those in cytosol when the aggregates form (e.g., 40 min in a)). c) A representative fluorescence image for 5 μM TMR-ligand co-incubated mHtt-97Q-Halo proteins in cells with a newly formed small aggregate. Comparison of cytosolic TMR fluorescence with known-concentration TMR-ligand solutions suggests an average of 80 μM cytosol polyQ concentration right after the aggregates form. This indicates that the average concentration of polyQ proteins reaches 400 μM in the newly formed small aggregates. Scale bars: 10 μm.

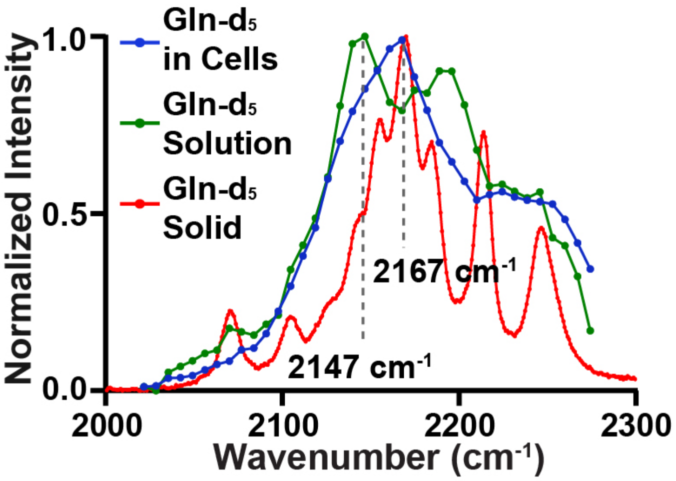


Fig. S4. Spectral change of Gln-d$_5$ after metabolic incorporation into proteins in HeLa cells. The Raman spectrum of Gln-d$_5$ after incorporated into cellular proteins (blue, peaked at 2167 cm^{-1}) presents a change in spectral line-shape compared to that of 100 mM Gln-d$_5$ solution (green, peaked at 2147 cm^{-1}). The shifted major peak of 2167 cm^{-1} (blue) matches with the Raman peak of Gln-d$_5$ solids (red, 2167 cm^{-1}). Overall, the Raman spectrum of Gln-d$_5$ incorporated in cells (blue) shows combinatorial spectral features from both that of Gln-d$_5$ solutions (green) and solids (red).

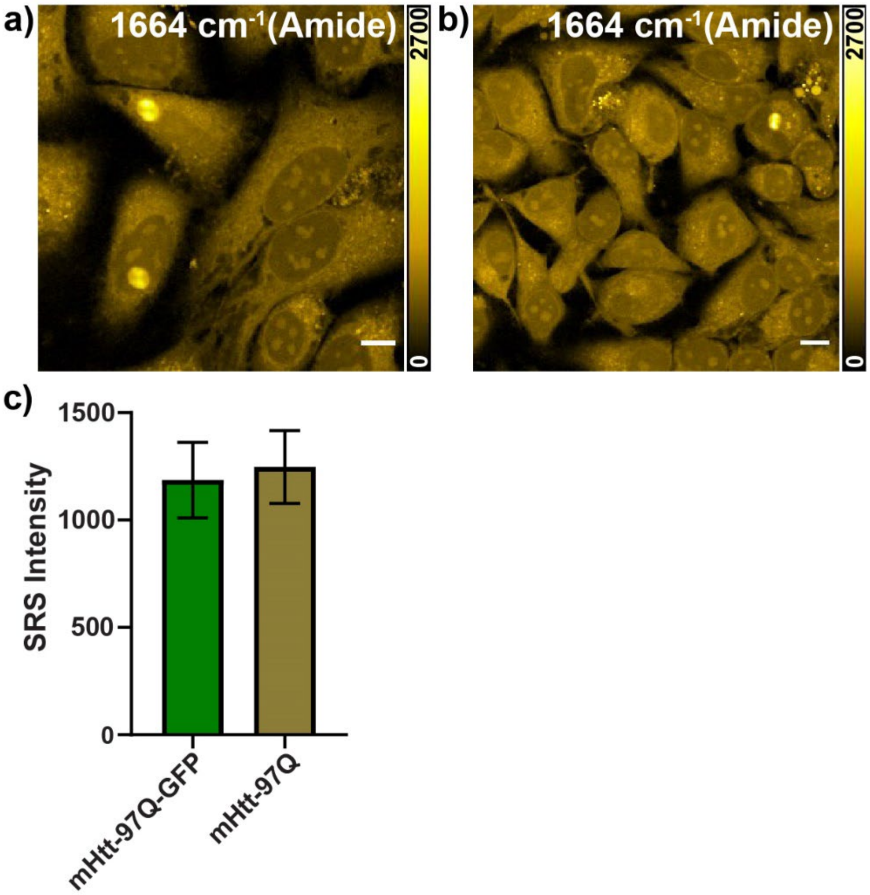


Fig. S5. Similar average Amide I intensities for mHtt-97Q-GFP and mHtt-97Q aggregates were observed. HeLa cells were grown in regular medium till optimal confluence and were transfected by (a) mHtt-97Q-GFP or (b) mHtt-97Q plasmids in Gln-d_5 medium for 24h before imaging by SRS at Amide I frequency (1664 cm^{-1}). (c) Quantification of average Amide I intensities for the two types of aggregates. mHtt-97Q-GFP (1186±175.8, n=10), mHtt-97Q (1248±169.4, n=14). Scale bar: 10 µm. Error bar: SD.

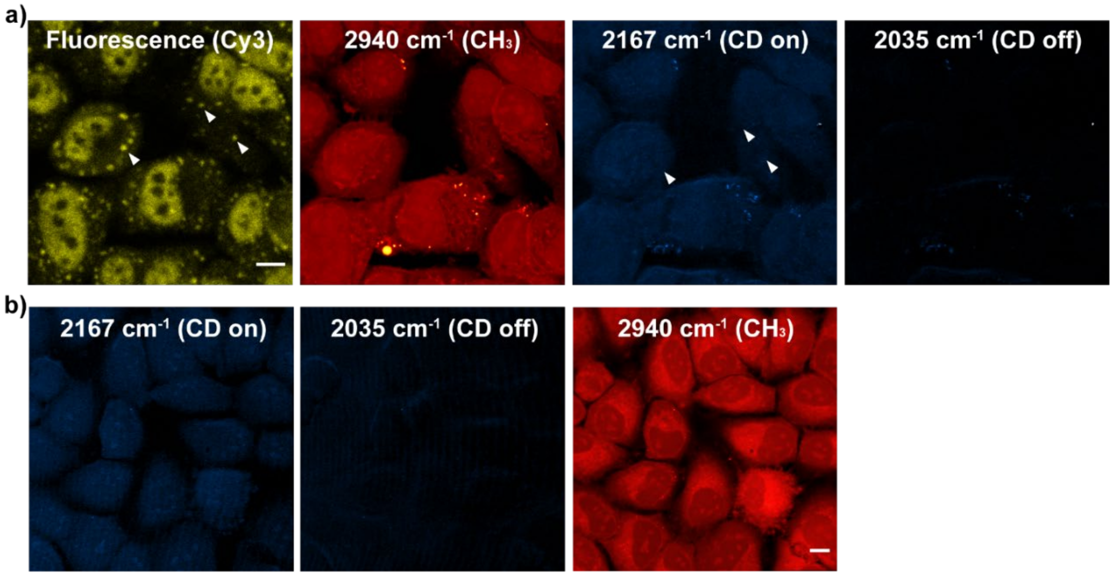


Fig. S6. SRS imaging with Gln-d$_5$ labeling doesn’t detect stress granules. Normal HeLa cells were pre-incubated in Gln-d$_5$ medium for 24h and was then stressed in 0.1 mM NaAsO$_2$ for 1h. a) Cy3-Oligo (dT) was used to visualize and confirm the formations of stress granules in fixed HeLa cells (yellow, fluorescence). Representative stress granules are arrow-headed. These stress granules were not detected in either CH$_3$ (2940 cm^{-1}, red) or C-D (2135 cm^{-1}, cyan) SRS images. b) No obvious puncta-like stress granules are detected by live-cell SRS imaging at the C-D (2135 cm^{-1}, cyan) and the CH$_3$ (2940 cm^{-1}, red) channels. Scale bars: 10 µm.

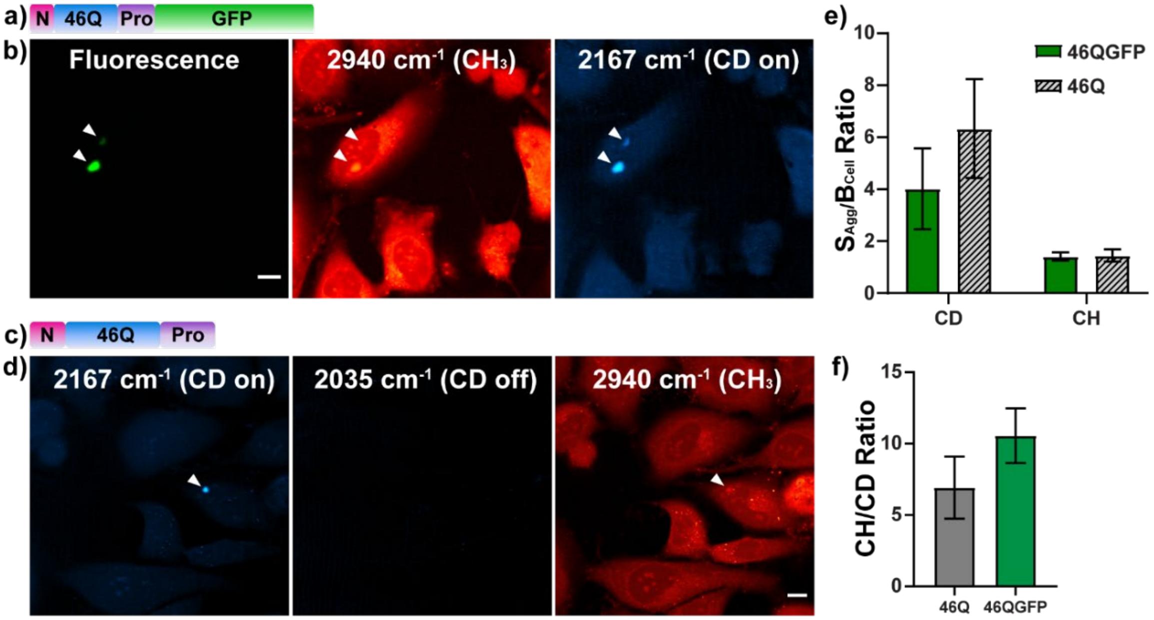


Fig. S7. SRS imaging and quantification of mHtt-46Q with and without GFP. a) Plasmid construct of mHtt-46Q-GFP. b) Fluorescence imaging (green) of mHtt-46Q-GFP aggregates (white arrow-headed). SRS imaging of the same set of cells at CH_3 channel (2940 cm^{-1}, red) and at the C-D channel (2167 cm^{-1}, cyan). The C-D SRS images recapitulate the aggregates (white arrow-headed) of varying sizes and shapes shown in fluorescence. c) Plasmid construct of mHtt-46Q by deleting GFP. d) SRS imaging of mHtt-46Q aggregates at C-D channel (2167 cm^{-1}, cyan, left), off-resonance channel (2035 cm^{-1}, middle) and CH_3 channel (2940 cm^{-1}, red, right). e) Aggregate signal (S_{Agg}) to cellular background (B_{Cell}) ratios (S_{Agg}/B_{Cell}) of mHtt-46Q aggregates at C-D (6.33±1.9, n=12) and CH_3 (1.45±0.24, n=12) channels and of mHtt-46Q-GFP aggregates at C-D (4.01±1.6, n=13) and CH_3 (1.41±0.16, n=13) channels. (f) CH/CD ratios for 46Q (6.9±2.2, n=12) and 46Q-GFP (10.6±1.9, n=13) aggregates. Error bar: SD. Scale bars: 10 μm.

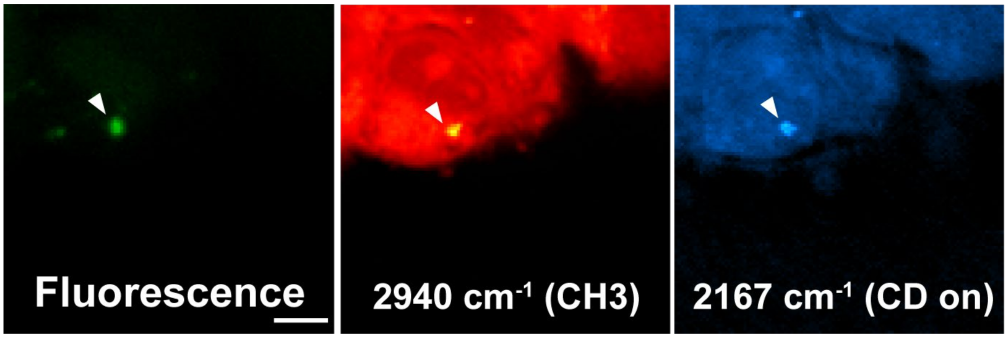


Fig. S8. SRS imaging of polyQ aggregates in stable embryonic stem cell-lines expressing mHtt-94Q-mEos2 proteins. Left, fluorescence imaging from mEos2 (green). Middle, SRS imaging at CH_3 channel (2940 cm^{-1}, red). Right, SRS imaging at C-D channel (2167 cm^{-1}, cyan). Scale bar: 5 µm.

Note: Gln is a non-essential amino acid, which can be *de novo* synthesized by cells. However, the extensive demand makes it an essential supplement in culture medium. It is also reported to become a conditionally essential amino acid for cells under stress.[8,9] In both Figs. S9 and S10 below, we confirmed the Gln-d_5 incorporation efficiency to be near 100%, meaning that Gln from *de novo* synthesis is close to 0%.

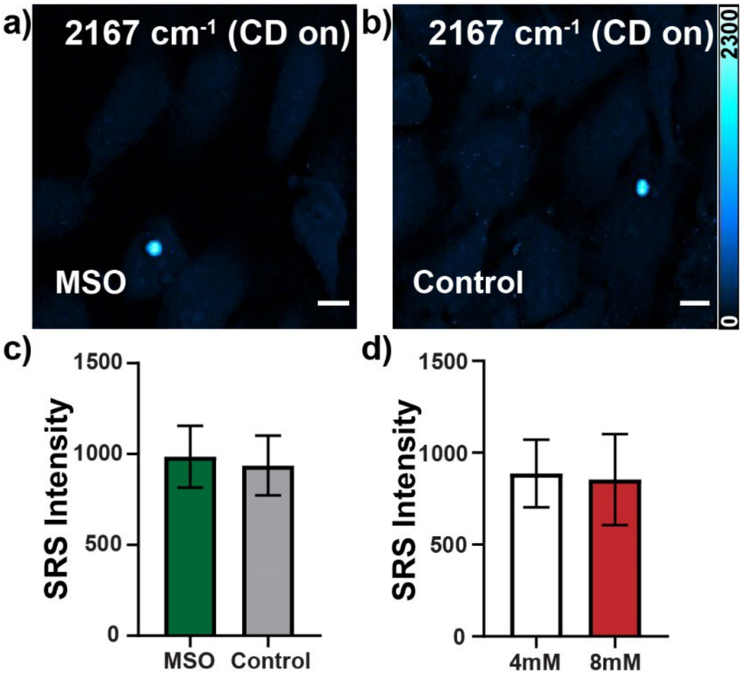


Fig. S9. The labeling efficiency of Gln-d$_5$ for mHtt-97Q aggregates approaches 100%. (a-b) HeLa cells transfected by mHtt-97Q was incubated with (a) or without (b) 3mM L-MSO, the Gln synthetase inhibitor[1] for inhibiting *de novo* Gln synthesis in the Gln-d$_5$ DMEM. (c) The L-MSO treated aggregates have similar intensity (984.1±160, n=9) with the non-treated ones (936.2±147, n=6). (d) Aggregates in HeLa cells, transfected by mHtt-97Q and incubated in normal DMEM (with 4mM Gln-d$_5$) or DMEM with doubled Gln-d$_5$ concentration (8mM), show no intensity difference (887.1±174, n=10 and 854.3±234, n=9, respectively). Our observations demonstrate that both inhibiting *de novo* Gln synthesis (c) and enhancing extracellular Gln concentrations (d) do not increase aggregate intensity. Our control experiments suggest that the labeling efficiency for the aggregates by Gln-d$_5$ from medium is close to 100%. Scale bars: 10 μm. Error bar: SD.

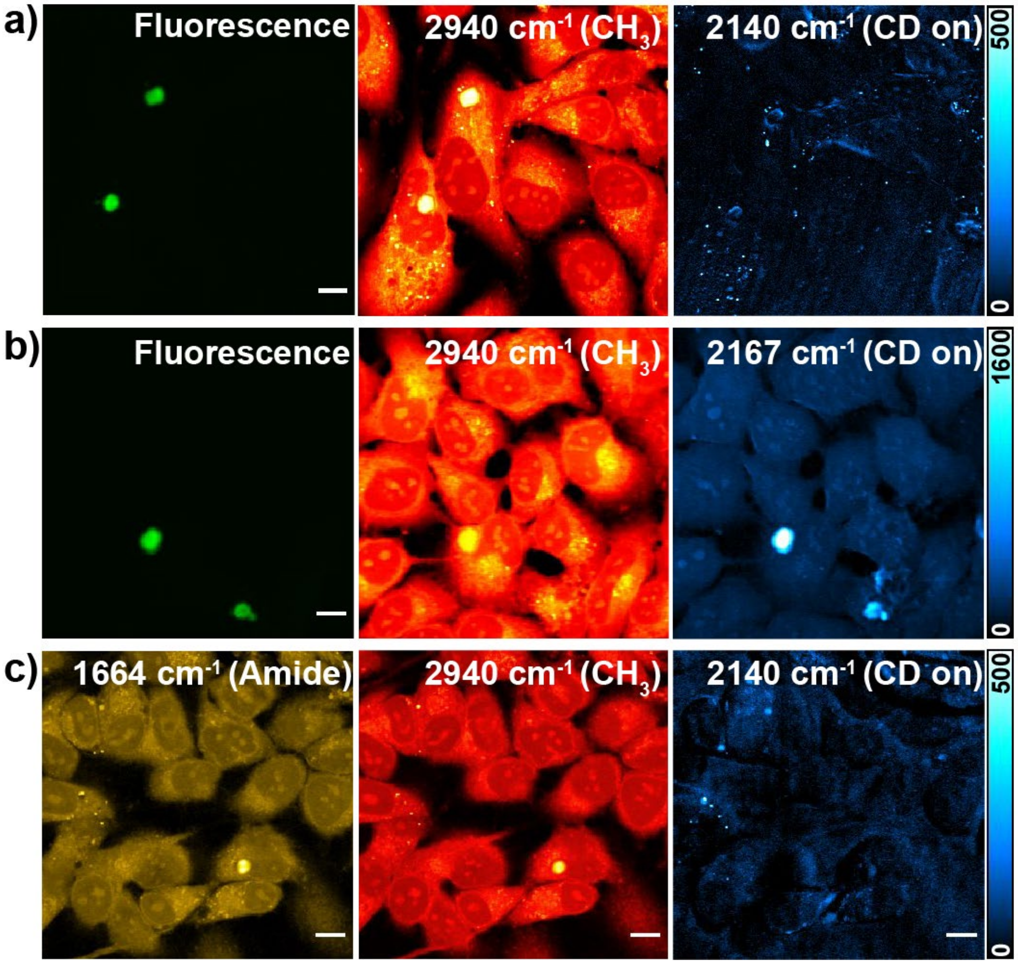
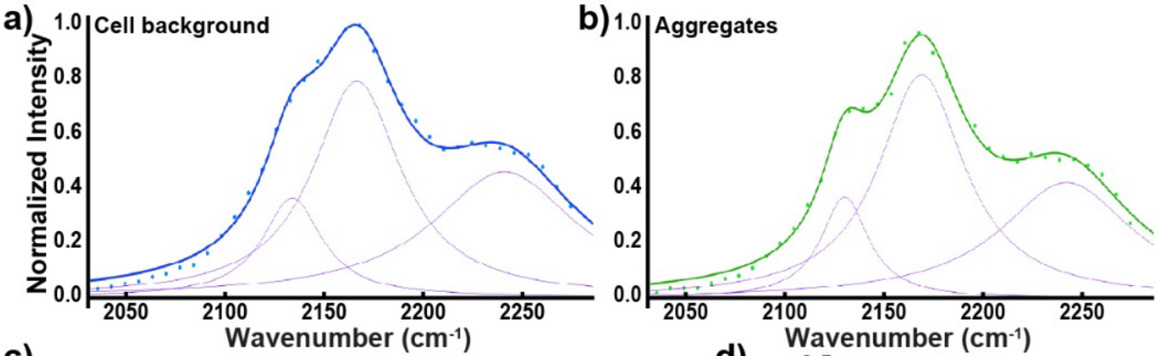


Fig. S10. D_7-glucose incubation has no detectable signals in both mHtt-97Q-GFP and mHtt-97Q aggregations through *de novo* Gln synthesis. (a-b) HeLa cells were transfected by mHtt-97Q-GFP of in (a) d_7-glucose medium for labeling *de novo* synthesized Gln and in (b) Gln-d_5 medium for Gln incorporation for 24h. D_7-glucose incubation shows almost no C-D signals (a, 2140 cm^{-1}, CD on) while Gln-d_5 incubation yields very bright aggregates in C-D channel (b, 2167 cm^{-1}, CD on). GFP is used as a guidance for aggregates due to minimum C-D signals D_7-glucose incubation (c) HeLa cells were first pre-incubated with d_7-glucose DMEM for 24h to pre-label *de novo* synthesized Gln and were next transfected with mHtt-97Q plasmid in d_7-glucose DMEM for another 24 h. No signals were detected for aggregates (c, 2140, CD on), suggesting a minimal level of Gln in aggregates comes from *de novo* synthesis. Scale bar: 10 µm.



c)

Peak Position/ FWHM	Peak 1	Peak 2	Peak 3
Aggregates	2130 cm^{-1}/ 29.0 cm^{-1}	2169 cm^{-1}/ 52.6 cm^{-1}	2242 cm^{-1}/ 81.0 cm^{-1}
Cell Background	2134 cm^{-1}/ 36.5 cm^{-1}	2166 cm^{-1}/ 53.8 cm^{-1}	2241 cm^{-1}/ 88.3 cm^{-1}

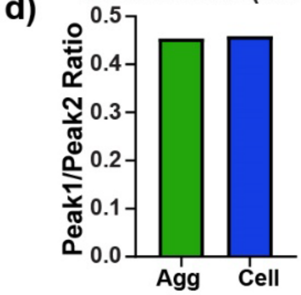


Fig. S11. C-D spectral shift induced by mHtt aggregates is mainly caused by the narrowing of the shoulder peak. (a-b) To understand the C-D spectral difference between the cell background and protein aggregates, we performed spectral fittings on SRS spectra of cell background (blue dots in a) and protein aggregates (green dots in b) by three Lorentzian peaks (purple lines in a and b, sum shown as fitted blue and green lines in a and b) using Fityk 1.3.1. (c) Our fitting results illustrate that the observed dip around 2146 cm^{-1} in aggregates (Fig. 5c) is mainly caused by the narrowing (29.0 cm^{-1} vs 36.5 cm^{-1}) and the slight red-shift (2130 cm^{-1} vs 2134 cm^{-1}) of the shoulder peak 1. (d) The relative intensity ratios of Peak1 to Peak2 remain the same for both aggregates and cell backgrounds.

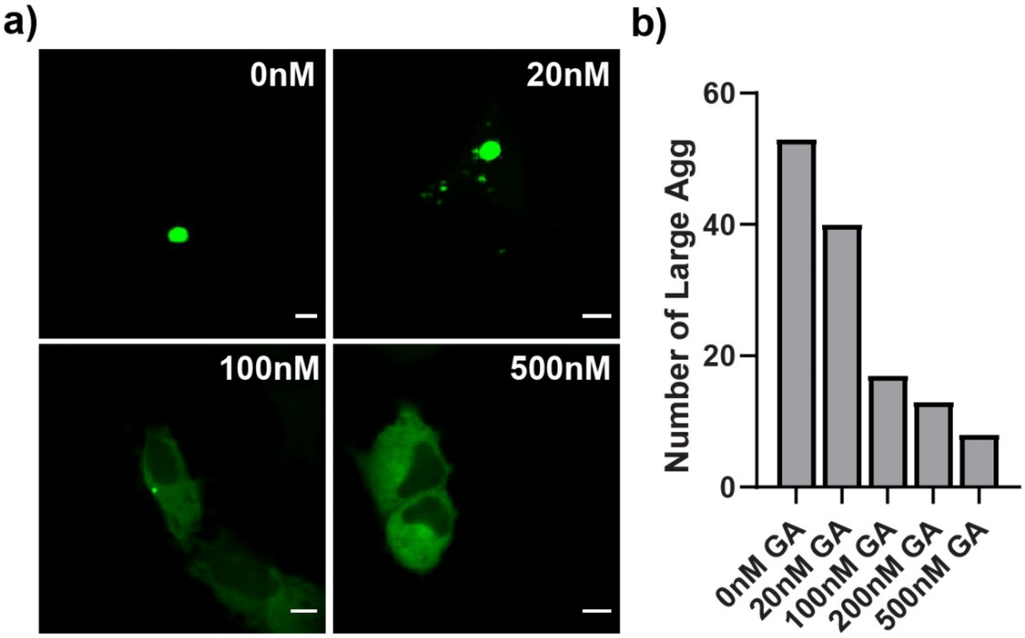


Fig. S12. Geldanamycin (GA) inhibits or solubilizes aggregation formation from both mHtt-97Q-GFP and mHtt-97Q in live HeLa cells. It has been shown that GA inhibits HSP90; induces heat shock response, and inhibits/solubilized EGFP-Htt72Q aggregation in other cell lines in a dose-dependent manner[10]. We confirmed similar GA functions in HeLa cells for mHtt-97Q-GFP aggregations (a) and for mHtt-97Q aggregates (b). We counted the number of large aggregates in one dish (14 mm^2) of HeLa cells transfected with same amount of mHtt-97Q plasmids but incubated with different GA concentration for 24h. Large aggregates were defined as aggregation area larger than 15 μm^2. Scale bar: 10 μm.

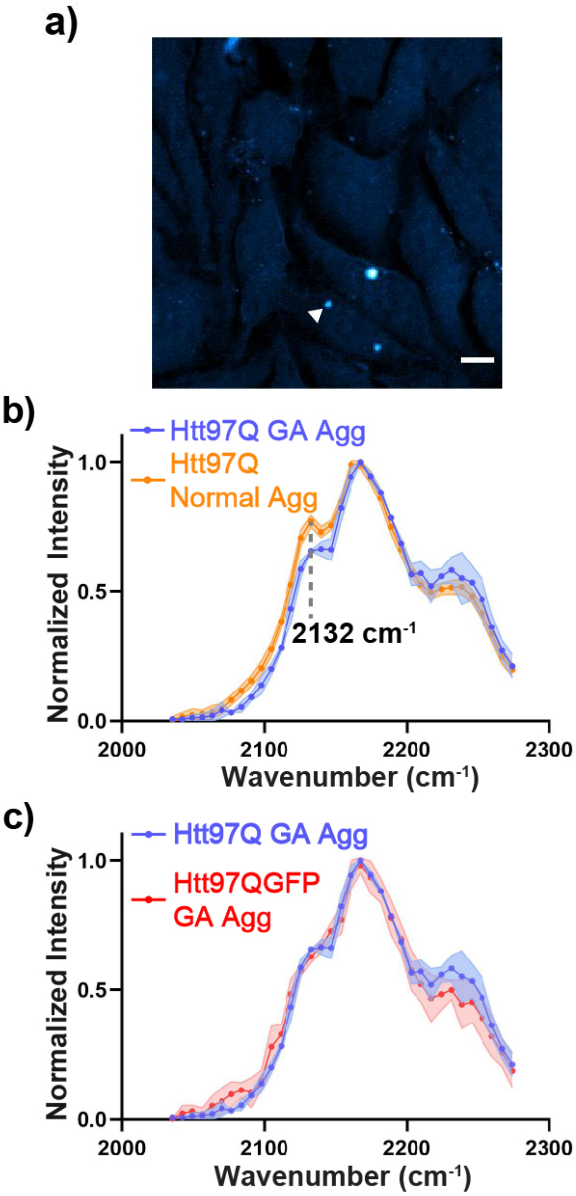


Fig. S13. GA induces a spectral change for a subset of mHtt-97Q aggregates. (a-b) Upon 24 h 100 nM GA treatment, a subset of mHtt-97Q aggregates with smaller sizes and reduced intensity (a, arrow indicated) presents a decrease at the 2132 cm^{-1} shoulder peak (indicated by grey dotted line) from the C-D SRS hyperspectra (b, mHtt-97Q GA agg, n=5) compared to untreated mHtt-97Q aggregates (b, mHtt-97Q normal agg, n=20). (c) Similar hyperspectral dip at 2132 cm^{-1} was also observed for mHtt-97Q-GFP aggregates (mHtt-97Q-GFP GA agg, n=7) compared to that of mHtt-97Q aggregates (mHtt-97Q GA agg) shown in (b). Scale bar: 10 μm. Error bar: SD.

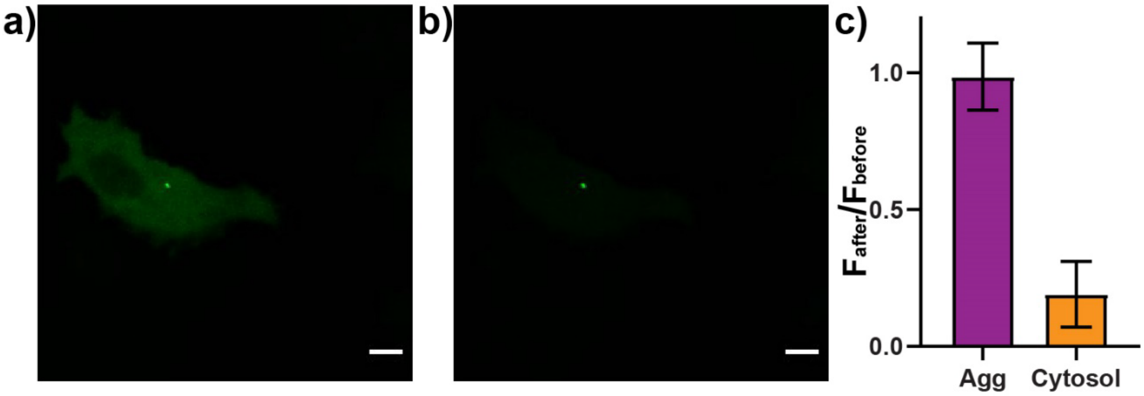
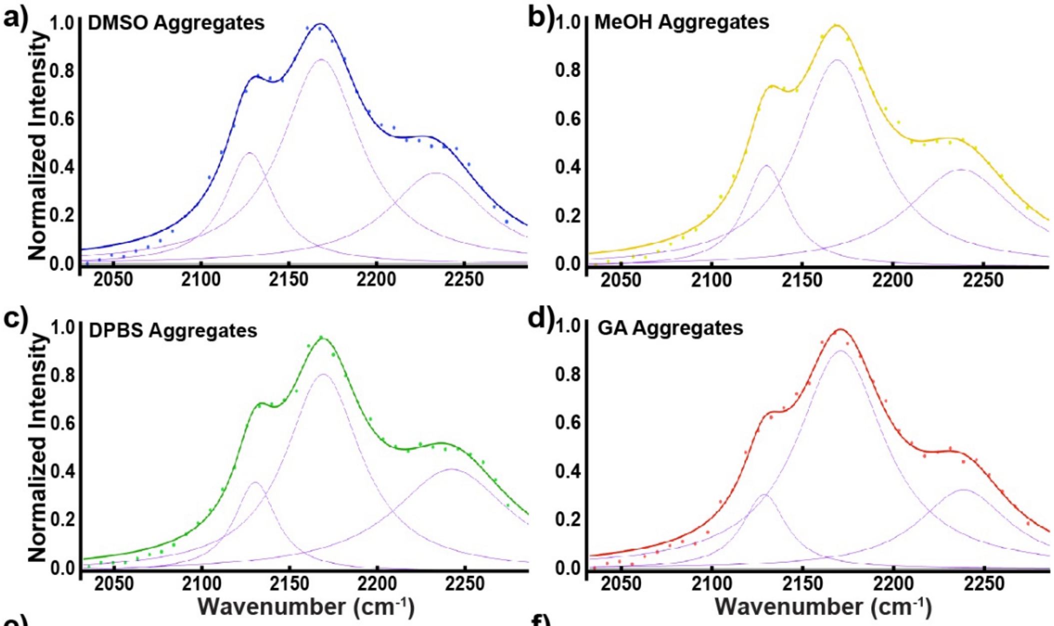


Fig. S14. Fluorescence Loss In Photobleaching (FLIP) on GA treated mHtt-97QGFP aggregates. mHtt-97Q-GFP was transfected into HeLa cells in Gln-d_5 medium with 100 nM GA and incubated for 24h. a) A representative cell with a small aggregate surrounded by diffusive fluorescence was imaged by a 2% 488 nm laser. b) The same cell, imaged by the same 2% 488 nm laser, after its peripheral cytosolic fluorescence being photobleached by 40% 488 nm laser for 150 s. c) Ratiometric quantification of fluorescence signals after and before FLIP (F_{after}/F_{before}) on the aggregates and the cytosolic proteins. After peripheral photobleaching, while the cytosolic fluorescent signals were significantly diminished (0.19±0.12, n=6), the fluorescent signals of aggregates remained unchanged (0.98±0.12, n=6). The FLIP result indicates that there is minimal protein exchange between the aggregates and the cytosolic diffusive mHtt proteins in the presence of GA. Scale bar: 10 μm.



e)

Position/ FWHM	Peak 1	Peak 2	Peak 3
DMSO	2128 cm^{-1}/34.1	2169 cm^{-1}/54.9	2234 cm^{-1}/66.8
MeOH	2130 cm^{-1}/29.5	2169 cm^{-1}/52.9	2238 cm^{-1}/73.2
DPBS	2130 cm^{-1}/29.0	2169 cm^{-1}/52.6	2242 cm^{-1}/81.1
GA	2129 cm^{-1}/29.4	2171 cm^{-1}/58.8	2239 cm^{-1}/61.3

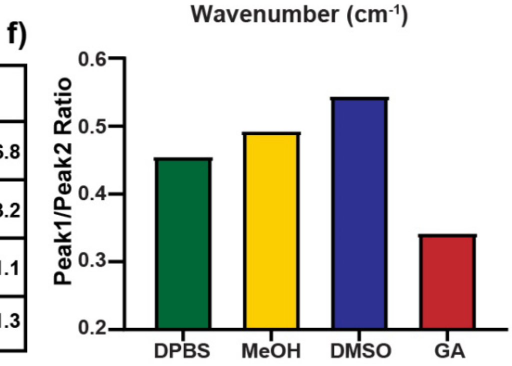


Fig. S15. Fittings of the C-D SRS spectra suggest a hyper-hydrated hydrogen-bonding environment for GA treated mHtt aggregates. To understand the dip induced by GA treatment (Fig. 6b), we performed spectral fitting for aggregates in solvents with descending hydrogen-bonding capacities and GA-treated aggregates. (a-d) The fitting was achieved by fitting the sum (colored lines) of three Lorentzian peaks (purple lines) to the experimental spectra (colored dots) using Fityk 1.3.1. (e-f) While the linewidths of both peak1 and peak2 had no obvious trend in different solvents (e), the peak1-to-peak2 ratios decreased monotonously with the increase of hydrogen-bonding abilities of the solvents (f, from DMSO to DPBS). GA-treated aggregation showed the lowest peak1 to peak2 ratio.

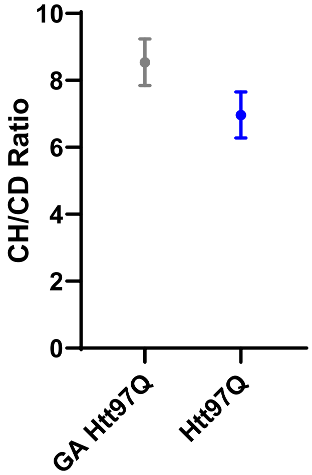


Fig. S16. CH/CD ratios show slight increase from small aggregates (GA Htt97Q, n=7) in cells with GA-induced heat-shock responses to those similar-sized aggregates (Htt97Q, n=12) in normal cells. Error bar: SEM.

2.6 Method and materials

Stimulated Raman Scattering (SRS) Microscopy A picoEmerald laser system (Applied Physics and Electronics) was used as the light source for SRS microscopy. It produces 2 ps pump (tunable from 770 nm – 990 nm, bandwidth 0.5 nm, spectral bandwidth ~ 7 cm^{-1}) and Stokes (1031.2 nm, spectral bandwidth 10 cm^{-1}) beams with 80 MHz repetition rate. Stokes beam is modulated at 20 MHz by an internal electro-optic modulator. The spatially and temporally overlapped Pump and Stokes beams are introduced into an inverted laser-scanning microscope (FV3000, Olympus), and then focused onto the sample by a 25X water objective (XLPLN25XWMP, 1.05 N.A., Olympus). Transmitted Pump and Stokes beams are collected by a high N.A. condenser lens (oil immersion, 1.4 N.A., Olympus) and pass through a bandpass filter (893/209 BrightLine, 25mm, Semrock) to filter out Stokes beam. A large area (10×10 mm) Si photodiode (S3590-09, Hamamatsu) is used to measure the pump beam intensity. A 64 V reverse-biased DC voltage is applied on the photodiode to increase the saturation threshold and reduce response time. The output current is terminated by a 50Ω terminator and pre-filtered by a 19.2-23.6-MHz band-pass filter (BBP-21.4+, Mini-Circuits) to reduce laser and scanning noise. The signal is then demodulated by a lock-in amplifier (SR844, Stanford Research Systems) at the modulation frequency. The in-phase X output is fed back to the Olympus IO interface box (FV30-ANALOG) of the microscope. 30 μs time constant is set for the lock-in amplifier. Correspondingly, 80 μs pixel dwell time is used, which gives a speed of 8.5 s/frame for a 320-by-320-pixel image. For 1664, 2035, 2167 and 2940 cm^{-1}, the wavelengths of pump laser are 880.6, 852.3, 842.8, and 791.3 nm, respectively. On-sample pump and modulated Stokes beam powers are 30 mW and 45 mW, respectively, for 1664 and 2940 cm^{-1}; 30 mW and 170 mW, respectively, for 2035, 2143, 2167 cm^{-1}. Laser powers are monitored throughout image acquisition by an internal power meter and power fluctuations are controlled within 1%. 16-bit greyscale images are acquired by Olympus Fluoview 3000 software. All data presented are from at least three independent experiments.

Spontaneous Raman Spectroscopy Spontaneous Raman spectra were acquired using an upright confocal Raman spectrometer (Horiba Raman microscope; Xplora plus). A 532 nm YAG laser is used to illuminate the sample with a power of 12 mW on sample through a 100×, N.A. 0.9 objective (MPLAN N; Olympus). Data acquisition was performed with 80 s integration by the LabSpec6 software. For Glutamine solution, background was subtracted by measuring signal from non-solution region on the same sample. The spectra are normalized to 2147 cm^{-1}.

Same Raman cross-sections for C-D and C-H are confirmed by Spontaneous Raman measurements on pure DMSO and DMSO-d$_6$. Same system throughput of pump lasers at two designated wavelengths for C-H and C-D SRS imaging are confirmed by SRS acquisitions on pure DMSO and DMSO-d$_6$ at the C-H and C-D channels.

Plasmid construction and molecular cloning. The plasmid pcDNA3.1-N-httQ97-GFP was a generous gift from Prof. Ron Kopito and Prof. F.-U. Hartl. L. Hartl. The mHtt-97Q was obtained by deleting the GFP sequence from the pcDNA3.1-N-httQ97-GFP plasmid. The mHtt-46Q and mHtt-46Q-GFP was obtained by deleting fragment from pcDNA3.1-N-httQ97-GFP and mHtt-97Q. 97Q-Halo was obtained by exchanging GFP with Halotag in pcDNA3.1-N-httQ97-GFP. The subcloning was done by GenScript USA Inc. For plasmid amplification, mHtt97Q-GFP and mHtt97Q were transformed into DH5α *Escherichia coli* cells. The cells were plated onto agar plates with respective antibiotics for selection for 14h. The resulting single colony was picked and grown in LB medium for 24 h. The cultures were collected, and the plasmids were purified by QIAGEN Maxi-prep kit.

Cell culture, transfection, metabolic labeling and imaging Deuterated glutamine DMEM was made by supplying glutamine-d$_5$ (Cambridge Isotope) to glutamine deficient DMEM (Gibco). Deuterated leucine DMEM was made by supplying leucine-d$_{10}$ (Cambridge Isotope) and regular methionine (Sigma-Aldrich) to leucine and methionine deficient DMEM (Thermo Scientific). Deuterated glucose DMEM was made by supplying d$_7$-glucose (Cambridge Isotope) to glucose deficient DMEM (Gibco). The solutions were sterile filtered by 0.22 μm low protein binding filter system (Corning). The filtered solutions were added with 10% FBS and 1% penicillin-streptomycin (Sigma-Aldrich) to make the complete media. Cultured HeLa-CCL2 (ATCC) cells were seeded onto 14 mm glass-bottom microwell dishes (MatTek Corporation) or coverslips (12mm, #1.5, Fisher) for 24 h prior to transfection. Cells were first grown in regular DMEM complete medium until they reached 70-90% confluence. The medium was switched to the designated deuterated medium immediately before transfection. Transfection of 1-μg plasmids encoding mHtt-97Q-GFP or mHtt97Q was performed using Lipofectamine 3000 transfection reagent (Thermo Fisher). After 24h of protein expression, medium was switched to the pre-warmed DPBS buffer before imaging. Coverslips were collected and attached to a microscope slides (1mm thick, VWR) with imaging spacer (Sigma-Aldrich); and the glass-bottom dishes were covered with another coverslip (22x22 mm, #1.5, VWR) on the top. Confocal fluorescence images were obtained by the Olympus FluoView™ FV3000 confocal microscope with SRS setup described above.

Induce and image stress granules HeLa cells. HeLa cells seeded on to glass-bottom microwell dishes were incubated with Gln-d_5 DMEM for 24h after they reach 50% confluence. Before imaging, HeLa cells was treated with 0.1 mM $NaAsO_2$ (Sigma Aldrich S7400) in regular DMEM for 1h at 37 ˚C, 5% CO_2. Live-cell SRS imaging was performed by replacing the $NaAsO_2$ containing medium with DPBS after 1h incubation.

For labeling stress granules with Oligo (dT) probe, Cy3-$(dT)_{30}$ oligonucleotide probe was purchased from IDT. After incubation with 0.1 mM $NaAsO_2$ for 1h, cells were fixed in 4% formaldehyde in PBS for 25 min at room temperature and subsequently permeabilized with 0.1% Triton X-100 in 2x saline sodium citrate (SSC) for 15 min, then washed with 1 M Tris, pH 8, for 5 min. Cells were blocked with 0.0005% BSA diluted in 2x SSC for 20min, washed with 1 M Tris, pH 8, for 5 min, and incubated with hybridization buffer (1.3 ng oligo(dT) probe, 0.005% BSA, 10% dextran sulfate, and 25% formamide, diluted in 2x SSC) for 1 h in a humid chamber. Cells were subsequently washed twice with 4x SSC for 5 min and once with 2x SSC for 5 min. The cells were carried onto perform fluorescence imaging of Cy3 and SRS imaging in C-D and CH_3 channels in DPBS buffer.

Cell culture and imaging of stable mouse embryonic stem (ES) cells continuously expressing mHtt-94Q-mEos2. Mouse D3 ES cells, which stably express mHtt-94Q-mEOS2 (from Prof. Zhe Liu, HHMI), were maintained on 0.1% gelatin coated plates in the absence of feeder cells. The ES cell medium was prepared by supplementing knockout DMEM (Gibco, 10829018) with 15% FBS, 1 mM GlutaMAX (ThermoFisher, 35050061), 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 0.1 mM 2-mercaptoethanol and 10 ng/ml of recombinant LIF (ThermoFisher, PMC9484). 24h before the imaging experiment, ES cells were seeded onto 14 mm glass-bottom microwell dishes (MatTek Corporation) pre-coated with Matrigel (Corning) in culture medium with 2 mM Gln-d_5 in place of 1 mM GlutaMAX for 24 h. ES cell imaging experiments were performed in the ES cell imaging medium, which was prepared by supplementing FluoroBrite medium (Invitrogen) with 10% FBS, 1 mM glutamax, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 10 mM Hepes (pH 7.2~7.5), 0.1 mM 2-mercaptoethanol and 10 ng/ml of recombinant LIF.

Treatment of Glutamine Synthetase Inhibitor in HeLa cells. Glutamine synthetase inhibitor L-Methionine sulfoximine (MSO) (CAS#15985-39-4) was purchased as solid from ACROS Organic. The white solid was dissolved in ddH_2O to obtain 300 mM stock solution. Right before transfection, the stock solution was diluted 100 times with Gln-d_5 DMEM to reach final concentration of 3mM in DMEM[1]. The HeLa cells seeded onto coverslip or imaging dish were switched to MSO medium immediately before transfection and was incubated with the presence of MSO throughout the 24 h expression time. Before imaging, the MSO medium was changed to pre-warmed DPBS buffer.

Hyperspectral SRS (hSRS) on Gln-d_5 Labeled Aggregates. The experimental set-up we used for hSRS imaging was described above (Fig.1 and S1). In brief, two synchronized picosecond laser pulses (bandwidth 0.5nm) from the APE laser described above was used for hSRS imaging. One laser was fixed at the fundamental laser wavelength (1031.2 nm) and the other was tunable from 770 nm to 990 nm. We chose the center wavelength of the tunable laser to be at 791.8 nm, 842.8 nm and 880.8 nm for hsSRS imaging in the 2940 cm^{-1}, the

2147 cm^{-1}, and the 1660 cm^{-1} regions, respectively. The step size of the tunable laser is 0.5 nm and an image is taken as the laser is tuned to each wavelength. 35 steps were collected to cover the full spectrum of 2940 cm^{-1} and 2147 cm^{-1} regions, while 21 steps were used for 1660 cm^{-1} region. The collection of images was acquired as the hyperspectral SRS image and the spectra was acquired by defining specific area of interest on the images and plotting the SRS intensity at corresponding wavenumber. The SRS intensities were background subtracted and normalized to laser power fluctuations. Spectral resolution is about 12 cm^{-1}. A 25× water immersion objective (XLPLN25XWMP, 1.05 N.A., Olympus) was used to focus the lasers onto the sample, with typical optical power at the sample of 30 mW for the pump beam and 160 mW for the Stokes beam. Each SRS image has 320 × 320 pixels and takes 8 seconds to acquire.

Treatment of Geldanamycin (GA) in HeLa cells. Geldanamycin (CAS# 30562-34-6) was purchased as solid from Cayman Chemical. The solid was first dissolved in DMSO (ATCC) to make a 50mM stock solution and further diluted with ddH_2O to 100μM. The aqueous GA solution was diluted by Gln-d_5 DMEM or regular DMEM to reach the 100nM GA concentrations. The HeLa cells seeded onto coverslip or imaging dish were switched to GA medium immediately before transfection and was incubated in the presence of GA throughout the 24h expression time. Further imaging experiments after 24h were performed after changing the GA medium to pre-warmed DPBS buffer.

FLIP (Fluorescence Loss In Photobleaching) on mHtt GA treated aggregates. FLIP experiments were performed by first defining regions of interest (ROI), which were always peripheral cytosolic part at least 5 μm away from the aggregate. An image was first collected at 5% of the bleaching power right before photobleaching takes place. The photobleaching was achieved by continuous illumination of ROI experiment with 40% of the 488nm laser from the Olympus Fluoview FV3000 microscope for 150s. After intense illumination, another image was acquired at the exact same position with same parameter as the image collected before photobleaching. The cytosolic regions and aggregates were analyzed separately.

Image processing and data analysis Color-coding and intensity profile for all images were done by ImageJ. Spectral plotting and signal to background ratio calculations were performed in Graphpad Prism 8.1.1. Figures were assembled in Abode Illustrator.

Reference

1. Paulson, H. Protein fate in neurodegenerative proteinopathies: polyglutamine diseases join the (mis)fold. *Am. J. Hum. Genet.* **64**, 339–345, (1999).
2. Hipp, M. S.; Park, S.-H.; Hartl, F. U. Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol.* **24**(9), 506-514, (2014)

3. Bates, G. Huntingtin aggregation and toxicity in Huntington's disease. *Lancet.* **361**(9369), 1642-1644, (2003)
4. Finkbeiner, S. Huntington's disease. *Cold Spring Harb. Perspect. Biol.* **3**:a007476 (2011).
5. Sahl, S. J.; Lau, L.; Vonk, W. I. M.; Weiss, L. E.; Frydman, J.; Moerner, W. E. Delayed emergence of subdiffraction-sized mutant Huntingtin fibrils following inclusion body formation. *Q Rev. Biophys.* **49**(e2), 1-13, (2016).
6. Eisele, Y. S.; Monteiro, C.; Fearns, C.; Encalada, S. E.; Wiseman, R. L.; Powers, E. T.; Kelly, J. W. Targeting protein aggregation for the treatment of degenerative diseases. *Nat. Rev. Drug Discov.* **14**(11), 759–780, (2015)
7. Ramdzan, Y. M.; Trubetskov, M. M.; Ormsby, A. R.; Newcombe, E. A.; Sui, X.; Tobin, M. J.; Bongiovanni, M. N.; Gras, S. L.; Dewson, G.; Miller, J. M. L.; et al. Huntingtin inclusions trigger cellular quiescence, deactivate apoptosis, and lead to delayed necrosis. *Cell Rep.* **19**(5), 919–927, (2017).
8. Arrasate, M.; Mitra, S.; Schweitzer, E. S.; Segal, M. R.; Finkbeiner, S. Inclusion body formation reduces levels of mutant Huntingtin and the risk of neuronal death. *Nature.* **431**(7010), 805–810, (2004).
9. Hosp, F.; Gutiérrez-Ángel, S.; Schaefer, M. H.; Cox, J.; Meissner, F.; Hipp, M. S.; Hartl, F.-U.; Klein, R.; Dudanova, I.; Mann, M. Spatiotemporal proteomic profiling of Huntington's disease inclusions reveals widespread loss of protein function. *Cell Rep.* **21**(8), 2291–2303, (2017).
10. Suhr, S. T.; Senut, M.-C.; Whitelegge, J. P.; Faull, K. F.; Cuizon, D. B.; Gage, F. H. Identities of sequestered proteins in aggregates from cells with induced polyglutamine expression. *J. Cell Biol.* **153**(2), 283–294, (2001).
11. Bäuerlein, F. J. B.; Saha, I.; Mishra, A.; Kalemanov, M.; Martínez-Sánchez, A.; Klein, R.; Dudanova, I.; Hipp, M. S.; Hartl, F. U.; Baumeister, W.; et al. *In situ* architecture and cellular interactions of polyQ inclusions. *Cell.* **171**, 179–187, (2017).
12. Olzscha, H.; Schermann, S. M.; Woerner, A. C.; Pinkert, S.; Hecht, M. H.; Tartaglia, G. G.; Vendruscolo, M.; Hayer-Hartl, M.; Hartl, F. U.; Vabulas, R. M. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144**, 67–78, (2011).
13. Donaldson, K. M.; Li, W.; Ching, K. A.; Batalov, S.; Tsai, C.-C.; Joazeiro, C. A. P. Ubiquitin-mediated sequestration of normal cellular proteins into polyglutamine aggregates. *Proc. Natl. Acad. Sci. U.S.A.* **100**(15), 8892-8897, (2003).
14. Buchanan, L. E.; Carr, J. K.; Fluitt, A. M.; Hoganson, A. J.; Moran, S. D.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. Structural motif of polyglutamine amyloid fibrils discerned with mixed-isotope infrared spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **111**(16), 5796-5801, (2014).
15. Xiong, K.; Punihaole, D.; Asher, S. A. UV resonance Raman spectroscopy monitors polyglutamine backbone and side chain hydrogen bonding and fibrillization. *Biochemistry.* **51**, 5822−5830, (2012)
16. Devitt, G.; Howard, K.; Mudher, A.; Mahajan, S. Raman Spectroscopy: An emerging tool in neurodegenerative disease research and diagnosis. *ACS Chem. Neurosci.* **9**(3), 404-420, (2018).
17. Matlahov, I.; van der Wel, P. C. Conformational Studies of Pathogenic Expanded Polyglutamine Protein Deposits from Huntington's Disease. *Exp. Biol. Med.* (2019), 153537021985662. https://doi.org/10.1177/1535370219856620.

18. Warner, J. B.; Ruff, K. M.; Tan, P. S.; Lemke, E. A.; Pappu, R. V.; Lashuel, H. A. Monomeric Huntingtin exon 1 has similar overall structural features for wild-type and pathological polyglutamine lengths. *J. Am. Chem. Soc.* **139**(41), 14456-14469, (2017)
19. Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. U.S.A.* **103**(45), 16764-16769, (2006).
20. Iwata, A.; Christianson, J. C.; Bucci, M.; Ellerby, L. M.; Nukina, N.; Forno, L. S.; Kopito, R. R. Increased susceptibility of cytoplasmic over nuclear polyglutamine aggregates to autophagic degradation. *Proc. Natl. Acad. Sci. U.S.A.* **102** (37), 13135–13140, (2005)
21. Liu, Y.; Wolstenholme, C. H.; Carter, G. C.; Liu, H.; Hu, H.; Grainger, L. S.; Miao, K.; Fares, M.; Hoelzel, C. A.; Yennawar, H. P.; et al. Modulation of fluorescent protein chromophores to detect protein aggregation with turn-on fluorescence. *J. Am. Chem. Soc.* **140**, 7381-7384, (2018).
22. Mangiarini, L.; Sathasivam, K.; Seller, M.; Cozens, B.; Harper, A.; Hetherington, C.; Lawton, M.; Trottier, Y.; Lehrach, H.; Davies, S. W.; et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell*, **87**, 493–506, (1996).
23. Tsien, R. Y. The Green Fluorescent Protein. *Annu. Rev. Biochem.* **67**(1), 509–544, (1998).
24. Bhattacharya M., Mukhopadhyay S. Studying protein misfolding and aggregation by fluorescence spectroscopy. In: Geddes C. (eds) *Rev. Fluoresc.* **8**. Springer, Cham (2015)
25. Kozlowski, L. P. Proteome-pI: proteome isoelectric point database. *Nucleic. Acids. Res.* **45**, D1112–D1116, (2017).
26. Freudiger, C. W.; Min, W.; Saar, B. G.; Lu, S.; Holtom, G. R.; He, C.; Tsai, J. C.; Kang, J. X.; Xie, X. S. Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science.* **322**(5909), 1857–1861, (2008).
27. Shen, Y.; Xu, F.; Wei, L.; Hu, F.; Min, W. Live-cell quantitative imaging of proteome degradation by stimulated Raman scattering. *Angew. Chem. Int. Ed.* **53**, 5596 –5599, (2014).
28. Wei, L.; Shen, Y.; Xu, F.; Hu, F.; Harrington, J. K.; Targoff, K. L.; Min, W. Imaging complex protein metabolism in live organisms by stimulated Raman scattering microscopy with isotope labeling. *ACS Chem. Biol.* **10**(3), 901-908, (2015)
29. Zhang, J.; Yan, S.; He, Z.; Ding, C.; Zhai, T.; Chen, Y.; Li, H.; Yang, G.; Zhou, X.; Wang, P. Small unnatural amino acid carried Raman tag for molecular imaging of genetically targeted proteins. *J. Phys. Chem. Lett.* **9**(16), 4679-4685, (2018)
30. Min, W.; Freudiger, C. W.; Lu, S.; Xie, X. S. Coherent nonlinear optical imaging: beyond fluorescence microscopy. *Annu. Rev. Phys. Chem.* **62** (1), 507–530. (2011)
31. Perney, N. M.; Braddick, L.; Jurna, M.; Garbacik, E. T.; Offerhaus, H. L.; Serpell, L. C.; Blanch, E.; Holden-Dye, L.; Brocklesby, W. S.; Melvin, T. Polyglutamine aggregate structure *in vitro* and *in vivo*; new avenues for coherent anti-stokes Raman scattering microscopy. *PLoS ONE*, **7**(7): e40536, (2012)
32. Wei, L.; Hu, F.; Shen, Y.; Chen, Z.; Yu, Y.; Lin, C.-C.; Wang, M. C.; Min, W. Live-cell imaging of alkyne-tagged small biomolecules by stimulated Raman scattering. *Nat. Method* **11**(4), 410–412, (2014).
33. Piez, K. A.; Eagle, H. The free amino acid pool of cultured human cells. *J. Bol. Chem.* **231**, 533-545, (1958).

34. Li, L.; Liu, H.; Dong, P.; Li, D.; Legant, W. R.; Grimm, J. B.; Lavis, L. D.; Betzig, E.; Tjian, R.; Liu, Z. Real-time imaging of Huntingtin aggregates diverting target search and gene transcription. *eLife*. **5**. doi.org/10.7554/eLife.17056. (2016)
35. Ji, M.; Arbel, M.; Zhang, L.; Freudiger, C. W.; Hou, S. S.; Lin, D.; Yang, X.; Bacskai, B. J.; Xie, X. S. Label-free imaging of amyloid plaques in Alzheimer's disease with stimulated Raman scattering microscopy. *Sci. Adv.* **4**: eaat7715, (2018).
36. Hackett, M. J.; McQuillan, J. A.; El-Assaad, F.; Aitken, J. B.; Levina, A.; Cohen, D. D.; Siegele, R.; Carter, E. A.; Grau, G. E.; Hunt, N. H.; et al. Chemical alterations to murine brain tissue induced by formalin fixation: implications for biospectroscopic imaging and mapping studies of disease pathogenesis. *Analyst*. **136**, 2941-2952, (2011)
37. Sittler, A.; Lurz, R.; Lueder, G.; Priller, J.; Hayer-Hartl, M. K.; Hartl, F. U.; Lehrach, H.; Wanker, E. E. Geldanamycin activates a heat shock response and inhibits Huntingtin aggregation in a cell culture model of Huntington's disease. *Hum. Mol. Genet.* **10**(12), 1307-1315, (2001).
38. Peskett, T. R.; Rau, F.; O'Driscoll, J.; Patani, R.; Lowe, A. R.; Saibil, H. R. A liquid to solid phase transition underlying pathological Huntingtin exon1 aggregation. *Mol. Cell*. **70**, 588–601, (2018).
39. Lee, H. J.; Zhang, D.; Jiang, Y.; Wu, X.; Shih, P.-Y.; Liao, C.-S.; Bungart, B.; Xu, X.-M.; Drenan, R.; Bartlett, E.; et al. Label-free vibrational spectroscopic imaging of neuronal membrane potential. *J. Phys. Chem. Lett.* **8**, 1932−1936, (2017).
40. Muchowski, P. J.; Schaffar, G.; Sittler, A.; Wanker, E. E.; Hayer-Hartl, M. K.; Hartl, F. U. Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc. Natl. Acad. Sci. U.S.A*. **97**(14), 7841-7846, (2000).
41. Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature*. **475**, 324–332, (2011).
42. Orr, H. T.; Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30**, 575–621, (2007).
43. Guo, Q.; Lehmer, C.; Martínez-Sánchez, A.; Rudack, T.; Beck, F.; Hartmann, H.; Pérez-Berlanga, M.; Frottin, F.; Hipp, M. S.; Hartl, F. U.; et al. *In situ* structure of neuronal C9orf72 poly-GA aggregates reveals proteasome recruitment. *Cell*. **172**, 696–705. (2018)
44. Ash, P. E. A.; Bieniek, K. F.; Gendron, T. F.; Caulfield, T.; Lin, W.-L.; DeJesus-Hernandez, M.; van Blitterswijk, M. M.; Jansen-West, K.; Paul, J. W.; Rademakers, R.; et al. Unconventional translation of C9orf72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*. **77**, 639–646, (2013).
45. Ennis, S. R.; Kawai, N.; Ren, X.; Abdelkarim, G. E.; Keep, R. F. Glutamine Uptake at the Blood-Brain Barrier Is Mediated by N-System Transport. *J. Neurochem.* **71** (6), 2565–2573. (2002)

Chapter 3

Label-free super-resolution volumetric imaging enabled by combining stimulated Raman scattering with sample expansion

The content of this chapter has been published in:
Qian, C.[#]; Miao, K.[#]; Lin, L.-E.; Chen, X.; Du, J.; Wei, L. Super-Resolution Label-Free Volumetric Vibrational Imaging. *Nat Commun.* 2021, 12 (1), 3648.
https://doi.org/10.1038/s41467-021-23951-x.

3.1 Summary

Innovations in high-resolution optical imaging have allowed visualization of nanoscale biological structures and connections. However, super-resolution fluorescence techniques, including both optics-oriented and sample-expansion based, are limited in quantification and throughput especially in tissues from photobleaching or quenching of the fluorophores, and low-efficiency or non-uniform delivery of the probes. Here, we report a general sample-expansion vibrational imaging strategy, termed VISTA, for scalable label-free high-resolution interrogations of protein-rich biological structures with resolution down to 78 nm. VISTA achieves decent three-dimensional image quality through optimal retention of endogenous proteins, isotropic sample expansion, and deprivation of scattering lipids. Free from probe-labeling associated issues, VISTA offers unbiased and high-throughput tissue investigations. With correlative VISTA and immunofluorescence, we further validated the imaging specificity of VISTA and trained an image-segmentation model for label-free multi-component and volumetric prediction of nucleus, blood vessels, neuronal cells and dendrites in complex mouse brain tissues. VISTA could hence open new avenues for versatile biomedical studies.

3.2 Introduction

Our knowledge of biology is significantly advanced by the development of optical imaging techniques. They reveal multi-dimensional spatial information that is crucial for understanding numerous functions and mechanisms in complex environments. At subcellular levels, super-resolution fluorescence microscopy techniques, including instrument-based stimulated emission depletion (STED) microscopy, photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM)[1,2], have been devised to overcome the diffraction limit barrier and allow optical visualization of previously unresolvable structures with nanometer resolution. Moreover, sample-oriented strategies, by physically expanding specimens embedded in swellable polymer hydrogels, have made a significant impact on high-resolution imaging of versatile samples[3]. For example, expansion microscopy (ExM) typically achieves a four-fold resolution enhancement using conventional fluorescence microscopes with isotropic sample expansion[4,5].

Despite their wide applications for uncovering unknown biological events through fine structural and functional characterizations, the super-resolution fluorescence microscopy has a few fundamental limits that originate from the requirement of fluorophore labeling. First, photobleaching and decay of the fluorophores make these techniques less ideal for repetitive and quantitative examinations of target structures[1-5]. This is especially problematic when the sample specimen is limited (e.g., clinical samples). Second, immunofluorescence, the widely used strategy for visualizing various proteins without genetic manipulation, poses serious issues of prolonged sample preparations and inhomogeneous antibody labeling in intact tissues[6]. This is due to the slow diffusion of large antibodies into the tissues and the depletion of probes on the surface[7]. To circumvent these fluorophore-associated challenges, we seek to have a super-resolution imaging modality that does not require labels.

Complementary to fluorescence, Raman microscopy targets the specific vibrational motions and maps out the distribution of chemical-bond specific structures and molecules inside live biological systems in a label-free or minimum labeling fashion. In particular, nonlinear stimulated Raman scattering (SRS) has been proven to be a highly successful optical imaging strategy for label-free or tiny-label imaging of biological samples with resolution and speed similar to those of fluorescence[8,9]. For example, SRS imaging targeting the methyl group (i.e., CH_3) vibrations from endogenous proteins at 2940 cm^{-1} (**Fig. 1a and** Supplementary Figure **1**) has been demonstrated for visualizing proteins-rich structures with submicron resolution at a speed up to video rate in live animals[10]. In principle, implementation of super-resolution Raman imaging could bypass the need of and hence the issues from fluorophore labeling. Despite extensive efforts, this goal has remained challenging. Strategies including excitation saturation, signal suppression with a donut beam or structural illumination have been reported [11-15]. However, they rely on additional specialized optics and the resolution enhancement is only 1.7 times on biological samples[13-15]. While such optics-based strategies have been heavily explored, there are no known efforts from the perspective of engineering samples for label-free super-resolved Raman imaging.

Here, we report a super-resolution label-free vibrational imaging strategy in cells and tissues that couples the sensitive SRS microscopy with recent sample-treatment innovations. We term this strategy *V*ibrational *I*maging of *S*welled *T*issues and *A*nalysis (VISTA). We embed biological samples in polymer hydrogels, expand the sample-hydrogel hybrid in water, and target the vibrational motion of retained CH_3 groups from endogenous proteins by SRS for visualization. Our devised strategy possesses a few desirable features. First, compared to fluorescence imaging, VISTA avoids any label-associated issues, allowing uniform imaging and a much higher throughput especially in tissues. Second, compared to optics-based Raman imaging, VISTA is easy to implement without any additional instrument and achieves an unprecedented Raman resolution down to 78 nm on biological samples. Importantly, it allows high-resolution imaging deep into the tissues[16], a common limit shared by all instrument-based super-resolution microscopy. Third, with further implementation of a convolutional neural network (CNN) for image segmentation[17], VISTA could offer specific, multi-component, and volumetric imaging in complex tissues with quality similar to that of fluorescence.

3.3 Results

Preservation of Proteins

As a first step to establish VISTA, we asked whether the sample expansion strategy is compatible with SRS microscopy. We embedded HeLa cells in a polymer gel following the widely-used ExM protocol[4,5] (**Fig. 1b, top**), which involves paraformaldehyde (PFA) fixation, gelation and sample homogenization through protease digestion. We then performed SRS imaging to visualize hydrogel-retained endogenous proteins in expanded cells at 2940 cm^{-1} (i.e., the CH_3 channel). However, almost no CH_3 contrast could be detected (**Fig. 1c, ExM and** Supplementary Figure **2a**), indicating an extensive loss of proteins or protein fragments under strong protease digestion (i.e., proteinase K)—a known issue in ExM[5]. Indeed, our quantification in CH_3 channel showed that the protein loss could reach 79% (**Fig. 1f**), consistent with a recent fluorescence analysis[18]. With such a high protein content loss and an approximate 64-fold signal dilution due to 4-fold isotropic sample expansion, SRS signals are therefore diminished. We asked whether reducing digestion time or altering to a milder protease (e.g., Lys-c)[5] would help retain SRS signals. Unfortunately, proteinase K already significantly digests the protein network even within 30 min (Supplementary Figure **2**). Further reducing the digestion time or changing the protease comes at the expense of low expansion ratios and sample distortion after expansion due to incomplete homogenization of PFA-crosslinked protein networks.

The key to VISTA is to preserve the maximum level of proteins for SRS imaging while achieving optimal homogenization for isotropic sample expansion. Since protease digestion is only required when extensive intra- and inter- protein crosslinking arises from PFA fixation[19], we resorted to using magnified analysis of proteome (MAP)[20], an alternative sample-hydrogel hybridization protocol. MAP significantly reduces such PFA-induced protein crosslinking by introducing high concentration of acrylamide (AA) together with PFA fixation so that the excess AA will react with and hence quench the reactive methylols formed by protein-PFA reaction[20] (**Fig. 1b, bottom**). The subsequent sample homogenization is achieved by protein denaturation instead of protease digestion. With optimizations of incubation time, SRS imaging of CH_3 showed clear cellular structures with a decent signal-to-noise ratio in expanded cells (**Fig. 1d, MAP**). Further quantification of the average CH_3 signal from MAP-processed cells compared to that from unprocessed cells indeed confirmed that proteins were largely preserved (**Fig. 1f**). Here, the slightly lowered CH_3 intensity from MAP was likely due to the removal of lipids (Supplementary Figure **3**). Recently, an ultrastructure expansion microscopy (U-ExM) was reported to image ultrastructures with decreased concentration of AA[21]. However, our analysis demonstrated that lowering AA concentration would still lead to a loss of rather significant portion of proteins (**Fig. 1e**, U-ExM and **Fig. 1f**, with an additional loss of 21% in whole cells and 32% in cytoplasm compared to MAP). We therefore concluded that MAP-based sample embedding protocol allows high-quality VISTA imaging of protein-rich subcellular structures that were not clearly identifiable with the normal SRS imaging resolution (**Fig. 1d, MAP vs** Supplementary Figure **1b,** e.g., the sub-structure of nucleoli and the network in cytosol).

Super resolution, three-dimensional imaging

Since the signal of VISTA comes from the CH_3 channel where the spectral crosstalk of other vibrations might exist, we next examined possible background contributions from both CH_2 vibrations of the hydrogel and the O-H stretch of the water. We compared the background sizes and spatial distributions by replacing normal hydrogel monomer and water with their deuterated correspondents, respectively. Our data showed that the background introduced from each component can be largely minimized through the deuteration strategy (Supplementary Figure **4a**). Images of the deuterated-hydrogel-embedded samples also confirmed that the background contribution of the polymer matrix was negligible (Supplementary Figure **4b-c**). After optimizing and confirming the sample processing and imaging conditions, we then aimed to determine the achievable resolution of VISTA. We performed regular SRS imaging on 100 nm polystyrene beads at 3050 cm^{-1} for C-H bonds with a NA=1.05 objective (**Fig. 1g**). We obtained a fitted image cross section with a full width at half maximum (FWHM) of 320 nm, which designates a resolution of 382 nm of the SRS microscope by Rayleigh criterion after deconvolution of the bead object function (**Fig. 1h)**. These data indicate that VISTA offers a 91 nm effective resolution after a 4.2-fold isotropic sample expansion (Supplementary Figure **5**, 4.2±0.1). We also confirmed a similar level of resolution of VISTA on small structural features within HeLa cells (**Fig. 1i-l**). The effective resolution could be further pushed to an unprecedented 78 nm with a higher-NA objective (i.e., NA=1.2) **(**Supplementary Figure **6).**

With isotropic sample expansion, VISTA also provides sharp three-dimensional (3D) views of cellular morphology and subcellular structures (Supplementary Figure **7a-b**). In addition to imaging cells in interphase, we applied VISTA to visualizing structural changes of HeLa cells throughout mitosis in metaphase, anaphase, telophase and interphase (**Fig. 2a-d**). VISTA images clearly resolved fine structures including cytosolic inner networks, small membrane protrusions of filopodia (**Fig. 2a-d, white arrowed**) and protein-rich contractile ring and midbody (**Fig. 2d, blue arrowed**). In addition, we observed an interesting change of protein contents during mitosis. The relative level of chromosome-associated proteins decreases as the nuclear envelopes disintegrate. This is evidenced by the dark region in cells (**Fig. 2a-c)**, which designates chromosome structures, confirmed by DAPI fluorescence stain (Supplementary Figure **8**). The protein level then increases as the nuclear envelopes reform at the end of the mitosis in telophase (**Fig. 2d, green arrowed**). The 3D views of the cells are shown in Supplementary Figure **7c-d**. The quantification of such change for protein abundance (Supplementary Figure **9a**) is further confirmed by regular SRS imaging on fixed, but non-hydrogel processed cells in the similar mitotic phases (Supplementary Figure **9b-d**). These data imply that VISTA is capable of performing quantitative total protein analysis at subcellular compartments.

Volumetric tissue imaging

Apart from visualizing fine subcellular structures, VISTA is well-suited for tissue imaging. We first demonstrated VISTA on the optical transparent zebrafish embryos, in particular the cone and rod photoreceptors in the outer segment (OS), important model systems for understanding visual perception[22,23]. Comparisons of CH_3 images before (**Fig. 2e**) and after (**Fig. 2f**) sample embedding and expansion from the same area presented a clear change of contrast due to lipid removal, which allowed us to unambiguously image protein-rich structures, e.g., the retinal pigment epithelium (**Fig. 2g,** arrow indicated). Due to the change of mechanical properties, our image-registration analysis on different tissue samples confirmed a slightly smaller average expansion ratios (Supplementary Figure **10,** 3.6 x for zebrafish and 3.4x for brain tissues, consistent with previous reports[20]). We then aimed to implement VISTA on the much more scattering and complex mouse brain tissues, especially for the hippocampus where characterizations of intricate structural relationship are important for functional understanding of a series of physiological (e.g., memory formation) and pathological (e.g., neurodegenerative diseases) events[24,25]. Our mosaic VISTA image on hippocampus (**Fig. 2h**) reveals clear and specific contrast from neuronal cell bodies, processes and also likely blood vessel cross-sections at various locations (**Fig. 2i-l, green, blue and white arrows indicated, respectively, in Fig. 2j**). Additional analysis of fine structures in the brain-tissue VISTA images also demonstrated our ability to resolve small features, likely dendritic spine heads or synapses[4], with an effective size smaller than 175 nm (Supplementary Figure **11**). All these features are virtually indistinguishable in regular SRS images due to a much lower resolution and the interference of lipid signals (Supplementary Figure **12**). With the homogenization of sample refractive index after lipid removal, deep VISTA imaging throughout a 1 mm hippocampal tissue (effectively 250 μm in unexpanded tissues) is also achieved (**Fig. 2m**). We note that our current imaging depth is mainly limited by the short working distance of the signal-collecting condenser and could be significantly improved by replacing both the objective and the condenser with long working distance objectives (e.g., 8 mm), specifically designed for tissue-clearing imaging of whole mouse brain hemispheres[26,27].

Validation with immunofluorescence

Since VISTA distinctly delineates the shapes of neuronal cells, processes, and likely blood vessels (**Fig. 2h-l**), we set out to validate the identities of these protein-rich biological components in VISTA images with established immunofluorescence. We were able to nicely correlate almost all structures shown in VISTA with fluorescence targets across various regions in brain tissues, including hippocampus and cortex. First, each vessel-like structure in VISTA is confirmed by lectin-DyLight594 staining, including those in vessel heavy regions within the hippocampus **(Fig. 3a-b,** Supplementary Figure **13a-c)** and larger ones likely of arteries[20,28] **(Fig. 3c-d)**. In addition to capturing vessel structures and distributions, VISTA could also image protein abundant red blood cells inside vessels that are retained in the polymer gel network (**Fig. 3c,** Supplementary Figure **13d-f**). Second, all the nuclei shown in VISTA have one-to-one correspondence to DAPI labeling **(Fig. 3e&g)**. We found that a small portion of these nuclei are from vascular endothelial cells, confirmed by their co-localizations with lectin vessel staining (**Fig. 3e-h,** arrowed) and GLUT1 immunostaining (Supplementary Figure **14**, arrowed). The rest of the nuclei should come from various types of cells including neurons, astrocytes, oligodendrocytes, etc. Third, together with nuclear structures, some cells also present clear contrast of cytoplasm. Our further correlative imaging identified that all these cell bodies captured by VISTA belong to matured neuronal cells, but not to astrocytes or oligodendrocytes. This is confirmed by the co-localizations of VISTA cell bodies with immuno-fluorescence-stained NeuN (matured neuron marker**, Fig. 3i-j**) and MAP2 (marker of neuronal perikarya and dendrites**, Fig. 3k-l**); and the lack of cross-localizations to glial fibrillary acidic protein (GFAP, astrocyte cellular maker, Supplementary Figure **15a-f**) and myelin basic proteins (MBP, oligodendrocyte cellular marker, Supplementary Figure **15g-l**). These results also suggest that maturated neuronal cells are more protein abundant in the cytosols compared to astrocytes and oligodendrocytes, a result difficult to quantify by other methodologies. Fourth, in addition to cell bodies, neuronal processes visualized in VISTA could be assigned to dendrites, which showed decent overlap with those imaged by MAP2 stains (**Fig. 3k-l**). Our characterizations by correlative immunofluorescence imaging hence confirm that VISTA offers holistic mapping of nuclei, vessels, and neuronal cell distributions and dendritic connections in brain tissues, without any labels.

Multiplexed VISTA with machine learning

High-resolution 3D mapping of the intricate cellular, vasculature, and connectivity network in brain has been a long-sought goal for super-resolved fluorescence microscopy. Imaging such multi-cellular interplay with high throughput would significantly maximize the information value and open new avenues for versatile biological investigations, such as in stroke models and neurodegenerative diseases[25,29-31]. As we were able to assign the origins of protein-abundant structures in VISTA to specific protein targets, we then aimed to transform each identified biological structure in **Fig. 3** into individual component for multi-target analysis through image segmentation. Recently, CNN based deep learning has been implemented for SRS imaging, but mostly focused on denoising and diagnosis prediction[32-35].

The imaging segmentation requires hyperspectral SRS and yet offers limited resolution[33]. Here, we adapted a U-Net based architecture[17,36] and trained our model with parallel VISTA and fluorescence images as input datasets for high-resolution image segmentation. VISTA images (**Fig. 4a, d, g**) were then passed through the trained model, and successfully generated predicted structures (**Fig. 4c, f, i**) that correlate well with fluorescence-labeled ground truth (**Fig. 4b, e, h**) for blood vessels (**Fig. 4c, v-lectin**), nuclei (**Fig. 4f, v-DAPI**), MAP2-immunolabeled neuronal cell bodies and dendrites (**Fig. 4i, v-MAP2**), and NeuN-immunolabeled mature neurons (Supplementary Figure **16**). The quality and contrast of these predicted images are close to the corresponding fluorescence images. The prediction performance is quantified by the Pearson correlation coefficient (Supplementary Figure **17a**). We note that the relatively low correlation values for NeuN prediction were mainly due to low fluorescence signals obtained (Supplementary Figure **17b**), likely caused by the loss of NeuN epitope during protein denaturation. With these 4 individual components successfully predicted, 4-color multiplex imaging is readily obtainable in 3D (**Fig. 4j and** Supplementary Figure **18**). For more integrated insights of biological organizations, 6-to-7-component imaging could also be achieved on the same sample with additional 2 to 3 fluorescent colors (**Fig. 4k and** Supplementary Figure **19**). Comparing to conventional label-free imaging, deep learning equipped VISTA offers desired target specificity for multiplex structural analysis. Comparing to sample-expansion fluorescence microscopy, which typically requires week-to-month long sample preparation with immunofluorescence[4,6,20,21], sample-processing steps for VISTA are complete within 48 hours (Supplementary Figure **20**). Such high-throughput nature of label-free VISTA imaging by omitting the multi-round immunostaining processes for multi-component investigations would largely facilitate our understanding of the intricate relationship between these cellular and sub-cellular structures deep inside tissues.

3.4 Discussion

In summary, we established VISTA as a robust and general label-free method for resolving protein-rich cellular and subcellular structures in 3D cells and tissues with an effective imaging resolution down to 78 nm. Targeting the CH_3 vibrational groups from endogenous proteins, VISTA is free from probe bleaching, decay or quenching caused by photo-illumination or gel polymerization and hence suited for repetitive and quantitative interrogations. Implemented with machine learning, VISTA allows specific and multi-component imaging of nuclei, blood vessels, matured neuronal cells and dendrites in brain tissues. Compared with hyperspectral SRS based segmentation methods, VISTA provides higher resolution and the capability of differentiating protein rich structures with similar chemical compositions.[26] Compared to fluorescence-based sample-expansion techniques, VISTA avoids low-efficiency, inhomogeneous delivery, and high cost of fluorescent antibodies, and thus offers fast throughput, uniform imaging throughout tissues, and cost-effective sample preparation, which scales up better with large human brain samples for future clinical investigations.

A few further technical improvements could be explored to bring VISTA a step forward. It could be coupled with all existing instrument-based high-resolution vibrational imaging techniques for further improvement in resolution. In particular, with a recently reported SRS configuration of frequency-doubled (i.e., wavelength-halved) excitation lasers[37], VISTA should further push the obtainable resolution down to 30 nm. As SRS signals scale nonlinearly with the laser frequency, frequency doubling would allow a 16-time improved sensitivity to resolve lower-protein-abundance structures. In addition to imaging proteins, VISTA could be extended to imaging other types of biomolecules including DNA, RNA, lipids with proper development of molecular anchoring chemistry during sample embedding[38]. VISTA should also find applicability in developmental biology for capturing mitotic cells in developing tissues, by employing a similar machine learning algorithm based on new training tailored for the void chromosomal regions, in combination with mitotic cell markers. With all these features, VISTA should find a wide range of applications for mapping subcellular architectures, cell distributions and connectivity across various molecular and resolution scales in complex tissues.

3.5 Figures

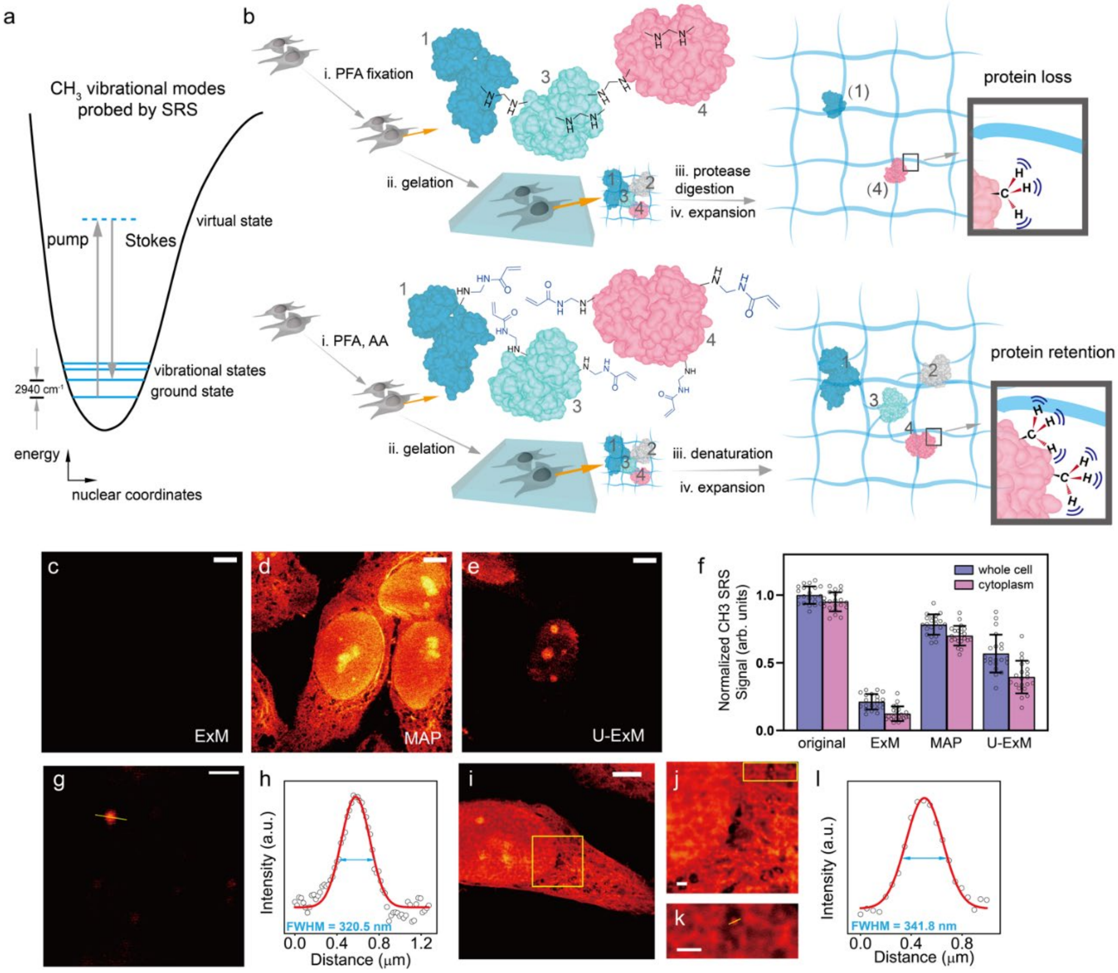


Figure 1. High-resolution label-free vibrational imaging of expanded and protein-retained samples. a, Energy scheme for SRS probing of CH_3 vibrational motion at 2940 cm^{-1}. More details in Fig. S1. **b**, Comparison of protein retention (i.e., the methyl groups, CH_3, from proteins) for SRS imaging between ExM (top) and MAP (bottom) based sample-hydrogel embedding procedures following different fixation, hybridization and homogenization chemistries. **c-e**. SRS imaging of CH_3 at 2940 cm^{-1} for expanded HeLa cells following ExM, MAP and U-ExM sample treatment under the same intensity scale. Scale bars: 20 μm. **f**, Quantification of proteins retention levels by comparing average CH_3 signals in expanded cells after ExM, MAP and U-ExM procedures with that from unprocessed HeLa cells (original). CH_3 signals in expanded cells were scaled back with the average expansion ratios for comparison. n=21 cells examined over 3 independent experiments. Data shown as mean ± SD. **g-h**, Quantification of SRS resolution by imaging the C-H vibration at 3050 cm^{-1} from a

representative 100 nm polystyrene bead (g) and fitting its cross-section profile (h). Scale bar: 1 µm. **i-l**, Fitted VISTA imaging cross-section profile (l) from a small structural feature (k) of expanded HeLa cells (i-k, j, k are zoom-in views from the boxed regions in i, j respectively). Scale bars: 30 µm in (i), 2 µm in j-k. For processed samples, the length scale is in terms of distance after expansion.

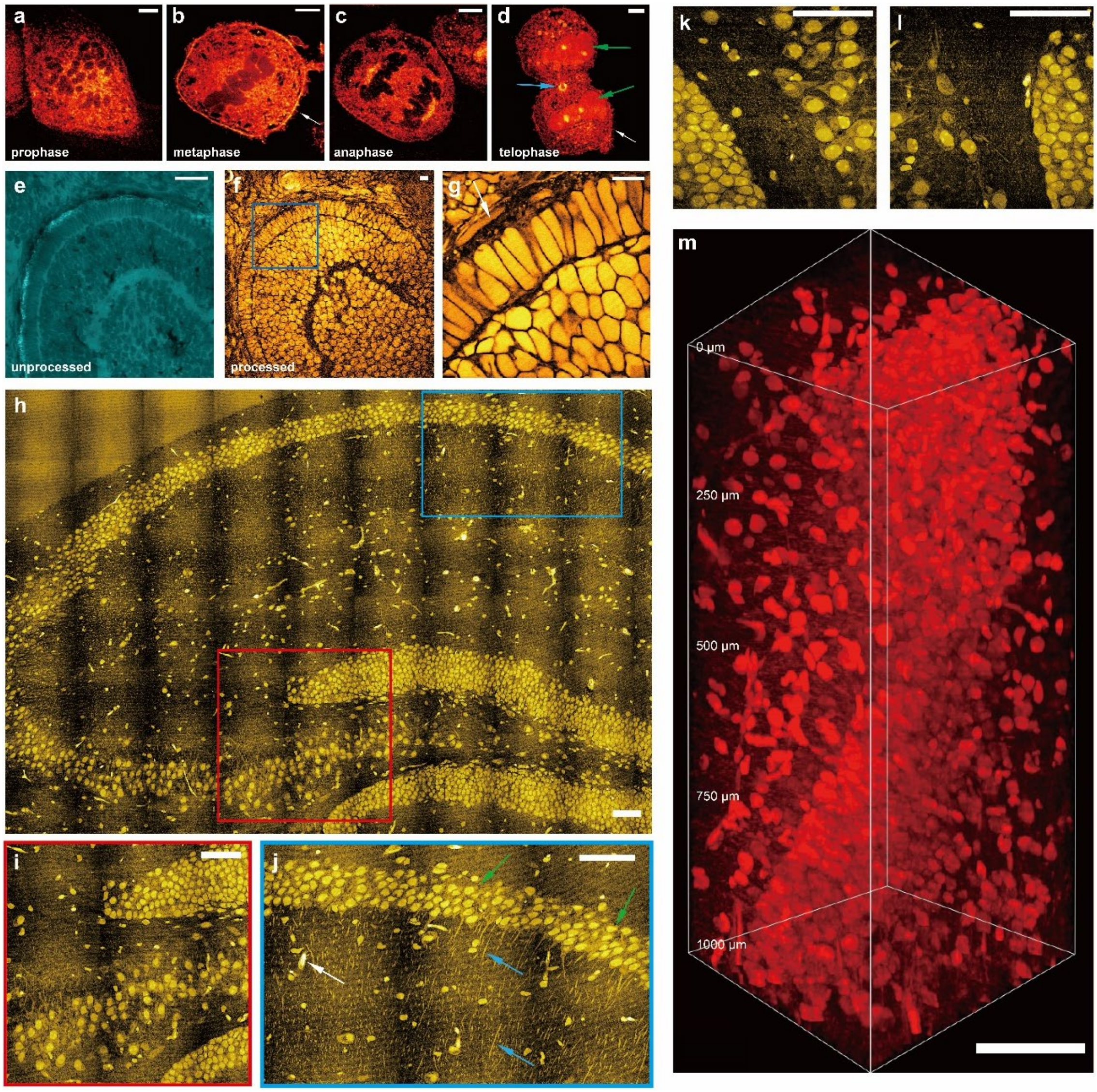


Figure 2. Super-resolution three-dimensional VISTA imaging of cells and tissues. a-d, Mitotic HeLa cells in prophase (a), metaphase (b), anaphase (c) and telophase (d). **e-g**, zebrafish embryonic retina before processing and after expansion: unprocessed (e), expanded (f), and zoom-in view from the boxed region in f (g). **h**, A mosaic VISTA image of a

hippocampal tissue. **i-l**, zoomed-in high-resolution view from color-boxed areas (i, red box; j, blue box) and selected regions in h. Representative neuronal cell bodies, neuronal processes and likely blood vessel cross-sections are indicated by green, blue and white arrows, respectively, in j. **m**, 3D volume VISTA imaging of a hippocampal tissue throughout 1000 μm depth. Scale bars: 20 μm (a-g) and 200 μm (h-m). For processed samples, the length scale is in terms of distance after expansion.

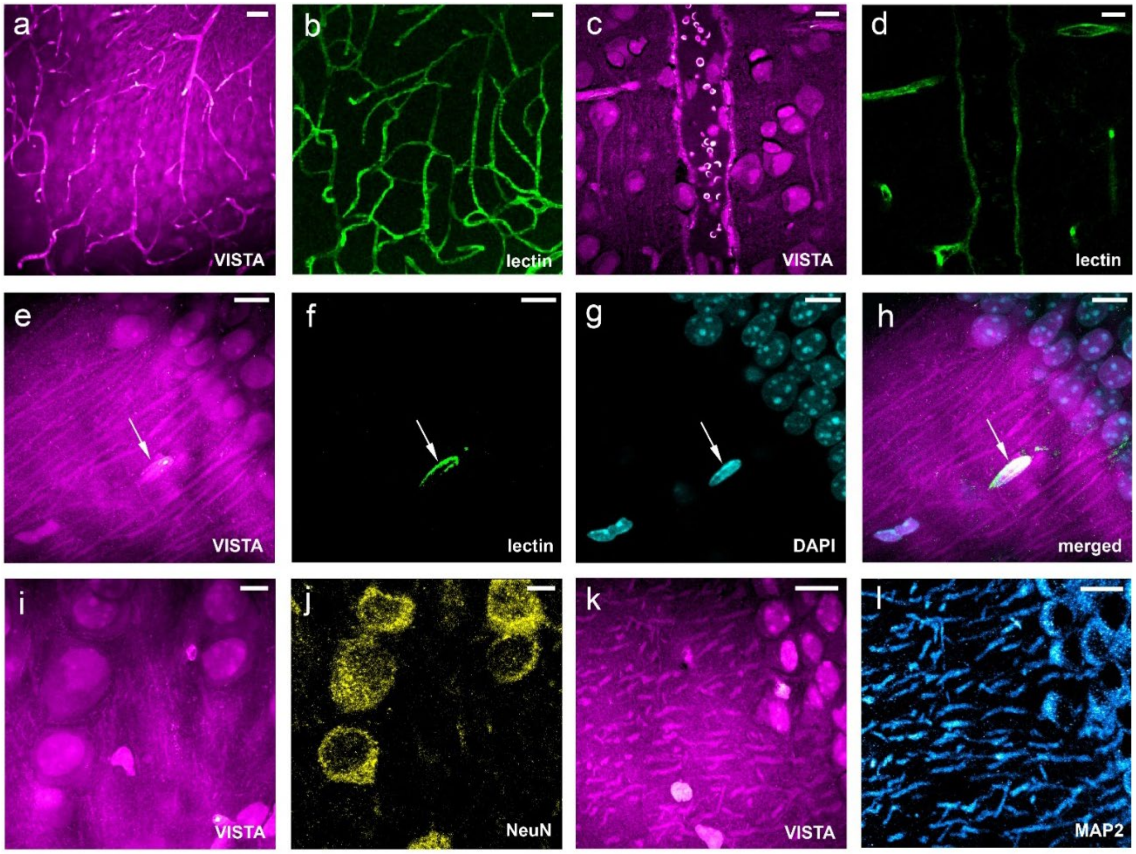


Figure 3. Validation of VISTA imaging features with fluorescent markers on mouse brain tissues. a-d, Parallel images of VISTA (a, c) and fluorescence from lectin-DyLight594 stained blood-vessel (b, d) in vessel-abundant regions. **e-g**, Parallel images of VISTA (e) with two-color fluorescence from lectin-DyLight594 stained vessels (f) and DAPI-stained nuclei (g). **h.** Three-channel merged image from e-g. **i-j**, Parallel images of VISTA (i) and fluorescence from immuno-stained NeuN, the matured neuron marker (j). **k-l**. Parallel images of VISTA (k) and fluorescence from immuno-stained MAP2, the neuronal cell body

and dendrite marker (l). All images are shown as maximum intensity projection from a stack of volume images. Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

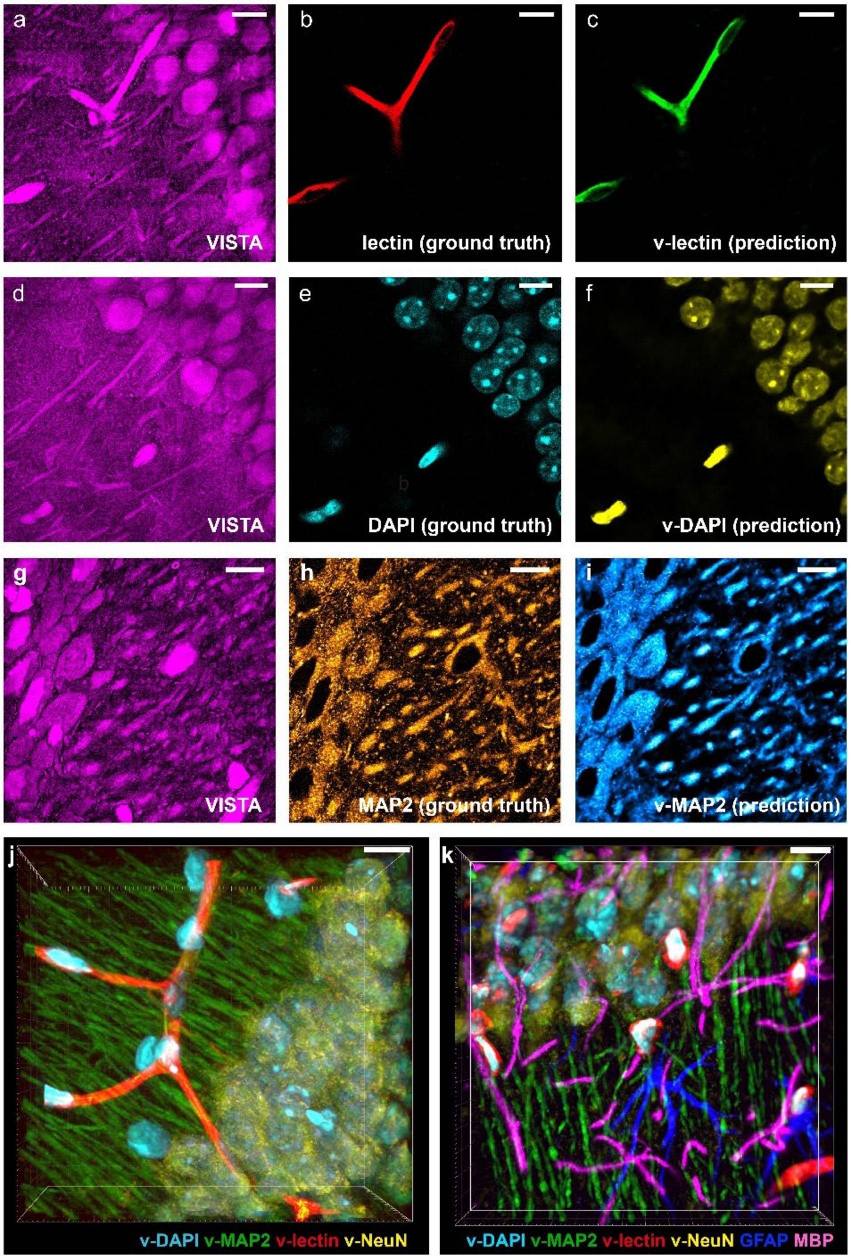


Figure 4. Label-free VISTA prediction for specific and multi-component imaging of brain hippocampal tissues. **a-c**, The input VISTA image (a), the ground truth fluorescence image of lectin-DyLight594 stained blood vessels (b) and the predicted VISTA-lectin (v-lectin) image of blood vessels (c). **d-f**, The input VISTA image (d), the ground truth

fluorescence image of DAPI stained nuclei (e) and the predicted VISTA-DAPI (v-DAPI) image of nuclei (f). **g-i**, The input VISTA image (g), the ground truth immunofluorescence image of MAP2 stained neuronal cell bodies and dendrites (h) and the predicted VISTA-MAP2 (v-MAP2) image of neuronal cells and dendrites (i). **j**, 4-color volume imaging from label-free VISTA prediction for vessels (v-lectin, red), nuclei (v-DAPI, cyan), neuronal cell bodies and dendrites (v-MAP2, green) and matured neuron cell bodies (v-NeuN, yellow). **k**, Tandem 6-color volume imaging from label-free 4-color VISTA prediction and parallel two-color immuno-fluorescence images of GFAP (blue) and MBP (magenta). Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

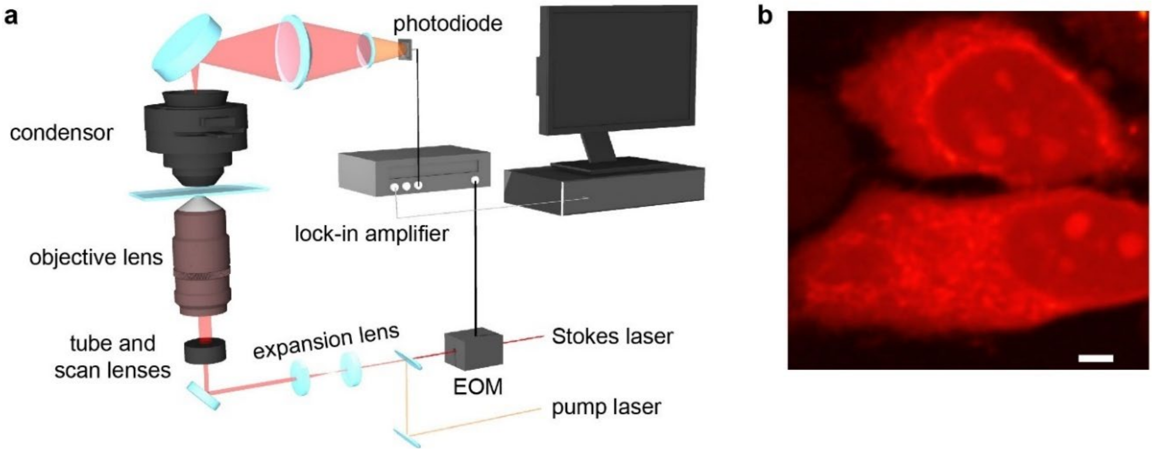


Figure S1. Setup for the stimulated Raman scattering (SRS) microscopy. (a) Instrumental setup of SRS microscope. EOM: Electro-Optic Modulator. When the energy difference between the Pump laser photons and the Stokes laser photons matches with the vibrational frequency of target chemical bonds (e.g., 2940 cm^{-1} for the symmetric vibrational motion of CH_3), the chemical bonds are efficiently driven from the vibrational ground state to the vibrational excited state, creating stimulated Raman loss in pump beam, which is subsequently detected by a photodiode and provides the imaging contrast. (b) A representative regular-resolution SRS image of HeLa cells targeting the protein CH_3 vibration at 2940 cm^{-1}. Scale bar: 3 μm.

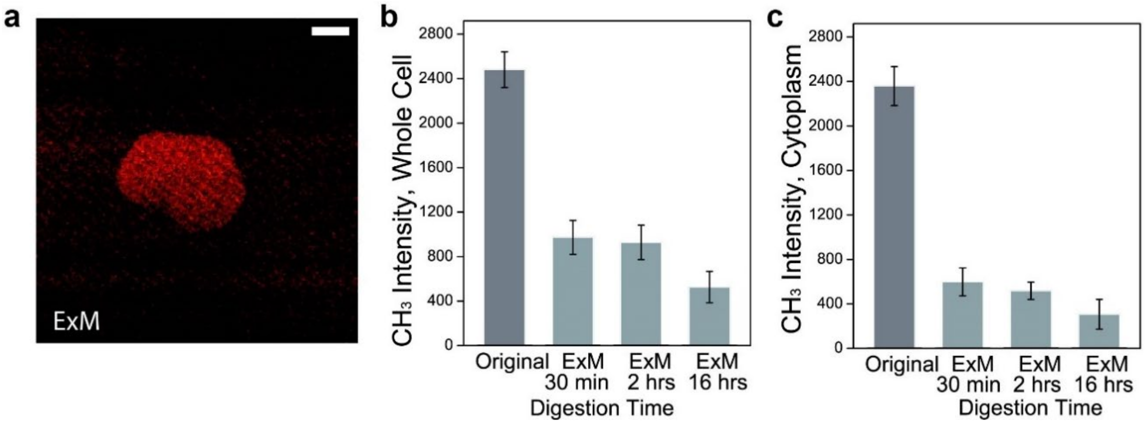


Figure S2. SRS quantification of protein loss following the ExM protocol. (a) Contrast-enhanced SRS image for Fig. 1c with a 2-time enhancement on the intensity scale. Scale bar: 20 µm. The length scale is in terms of distance after sample expansion. (b)-(c) Quantification of protein retention level on different protease digestion time by proteinase K on whole cells (b) and from the cytoplasm (c). SRS images of protein CH_3 from HeLa cells were acquired after designated digestion times. The protein signals decrease rapidly upon protease treatment, confirming that the digestion step in ExM causes significant protein loss. Error bar: SD.

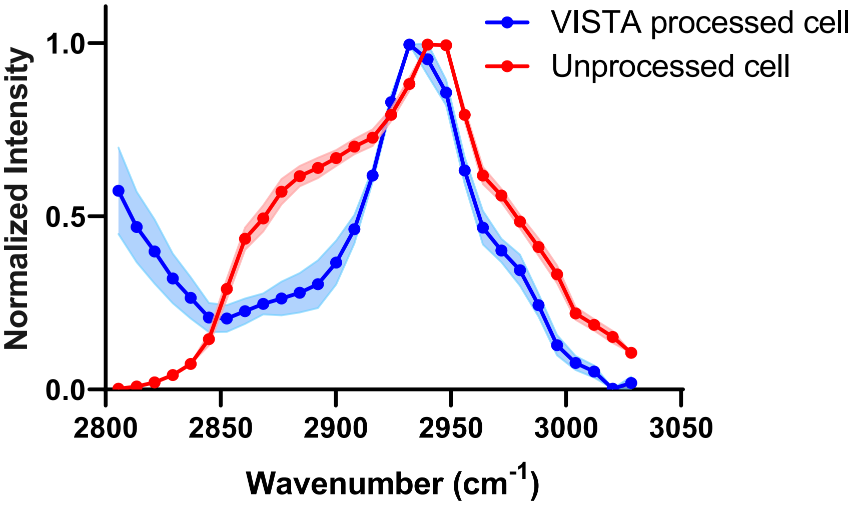


Figure S3. Hyperspectral SRS from HeLa cells with and without VISTA processing. The spectra were acquired by sweeping pump laser wavelength with 0.5 nm step size from 783.3 to 799.8 nm, with Stokes laser fixed at 1031.2 nm. The evident loss of lipid peak at 2855 cm^{-1} between normalized SRS spectra from unprocessed cells (red) and VISTA processed cells (blue) confirms the loss of lipid content from the protein denaturation step, which washes out the lipids. The intact peak shape for CH_3 at 2940 cm^{-1} confirms that VISTA retains proteins. The shoulder around 2800 cm^{-1} for VISTA processed cell spectrum (blue) originates from D_2O. VISTA processed cell, N=10; unprocessed cell, N=15. Error bar: SD.

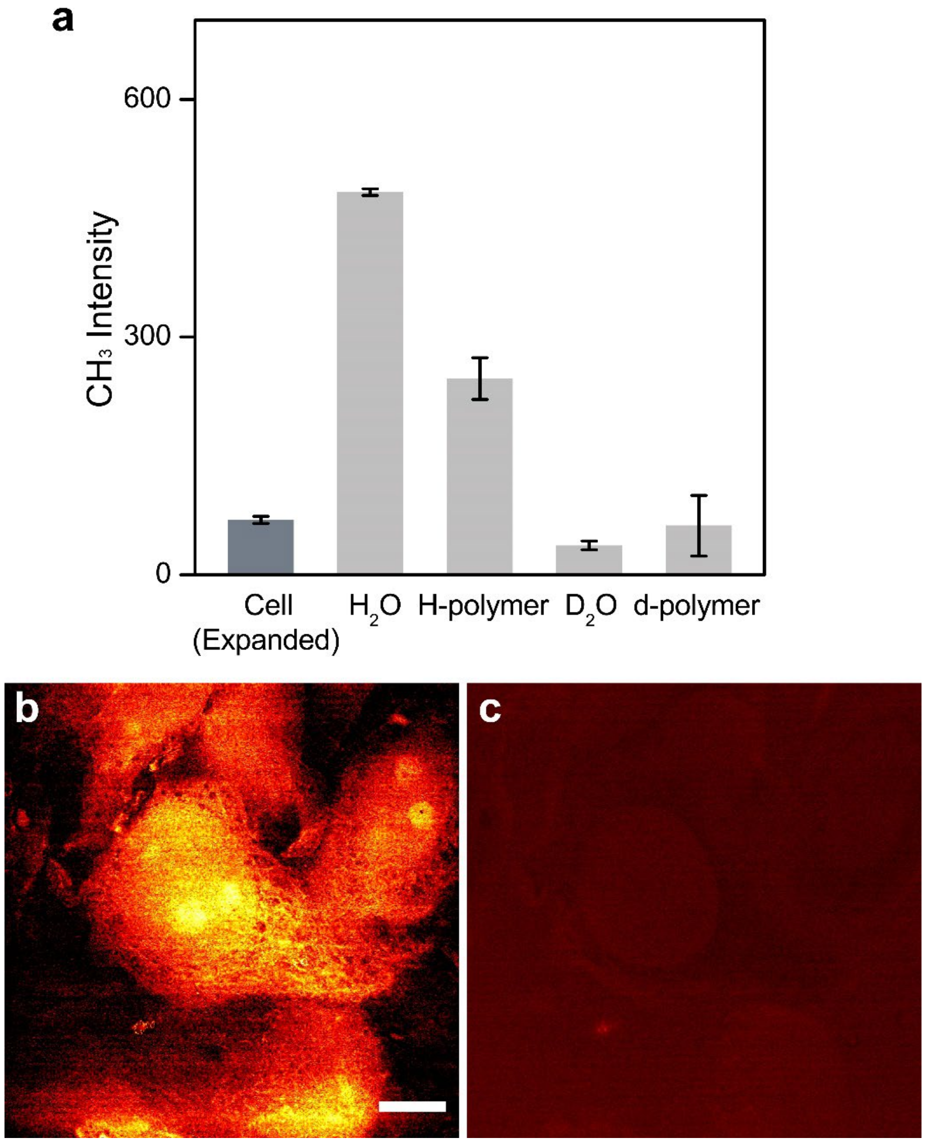


Figure S4. Quantification of backgrounds in VISTA CH_3 images on HeLa cells. (a) Background contributions from each component (i.e., O-H stretch tail from H_2O and C-H vibrations from the polymers composed of sodium acrylate and acrylamide monomers) were obtained by comparing normal VISTA cell images with VISTA images using corresponding deuterated components. The change of the backgrounds is constant across field of views after deuteration. Error bar: SD. (b-c) Representative VISTA images from expanded HeLa cells embedded in the deuterated polymer at the CH_3, 2940 cm^{-1} channel for cellular distributions (b) and the corresponding CD, 2176 cm^{-1} channel for deuterated polymers (c). Scale bar: 20 μm.

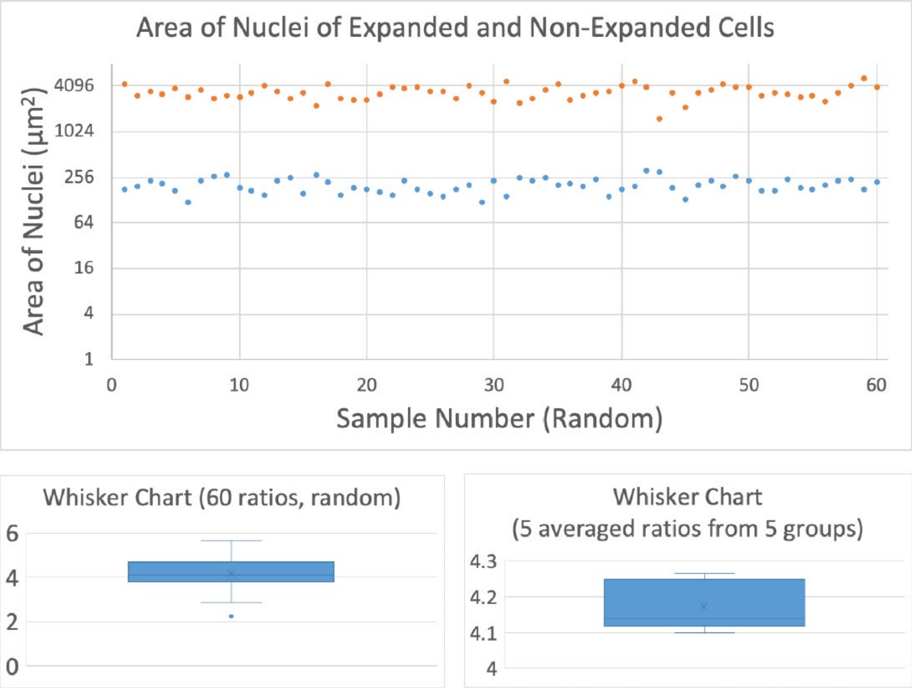


Figure S5. **Statistical quantification of the average expansion ratio for cell samples by VISTA.** Top: Areas of randomly chosen 60 expanded (orange) and non-expanded (blue) cells across 5 individual sample groups. Bottom: statistics with whisker charts for calculation of the average expansion ratios on cells. The left whisker chart (4.2±0.7) is calculated based on total 60 random expanded vs non-expanded pairs, and the right one (4.2±0.1) is calculated based on 5 averaged ratios from the 5 sample groups (12 cells in each group). Data shown as mean±sd.

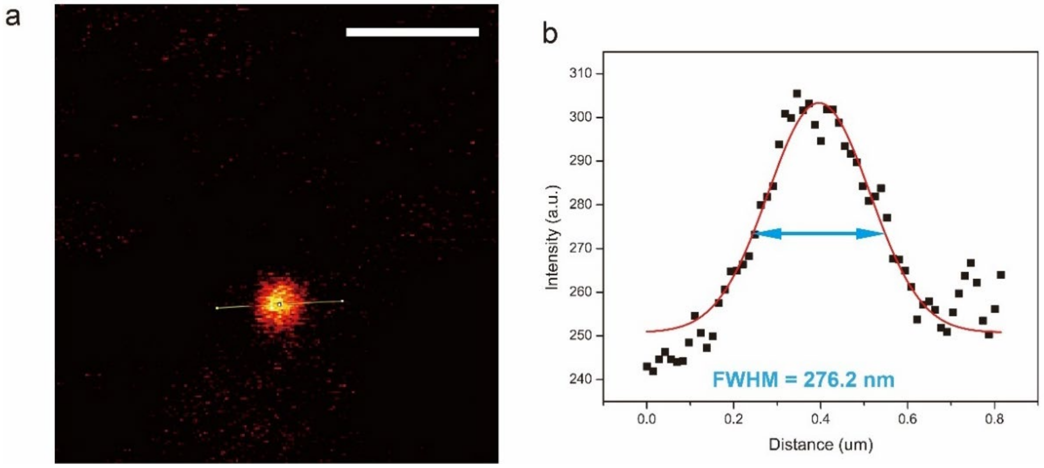


Figure S6. Quantification of SRS resolution under a higher numerical aperture (NA) objective lens. A representative image (a) and the corresponding fitted cross-section profile (FWHM of 276.2nm) (b) from a 100-nm polystyrene bead, by targeting the C-H vibration at 3050 cm^{-1} with a 60X water objective (Olympus, UPLSAPO60XW, 1.2 NA). Scale bar: 2 µm.

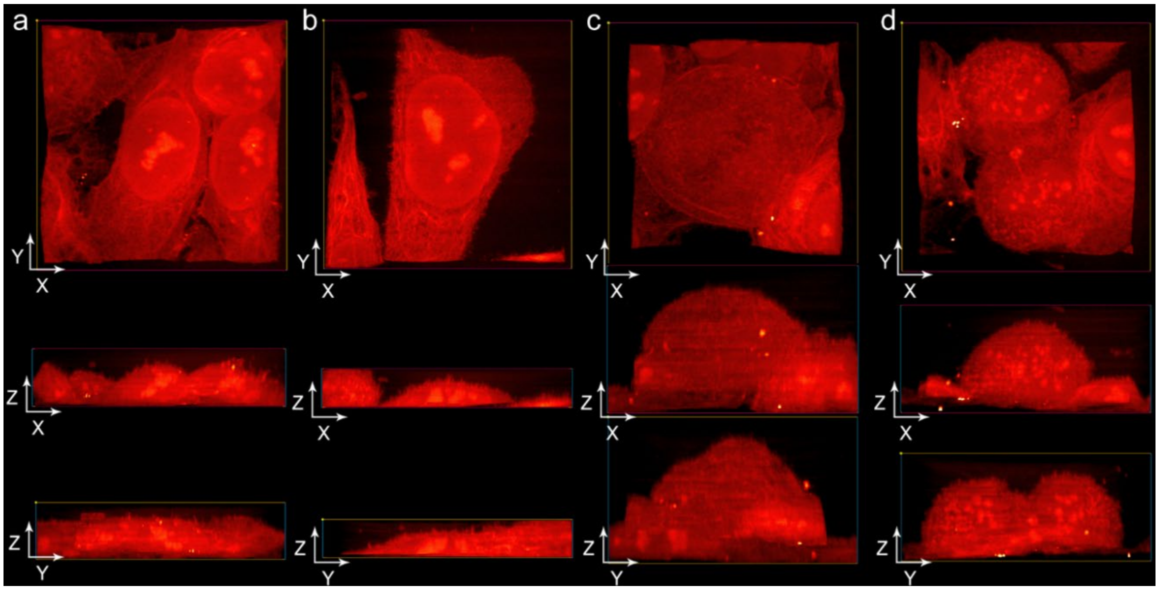


Figure S7. High-resolution three-dimensional VISTA views of cellular morphology and subcellular structures of interphase and mitotic HeLa cells. (a-b) Volume HeLa cells at interphase. The network structure in the cytosol is clear. In the x-z view, minor upward extrusions from cell surface (zx view) were also captured. (c-d) Volume HeLa cells during mitosis for single-z images shown in Fig. 2b and Fig. 2d.

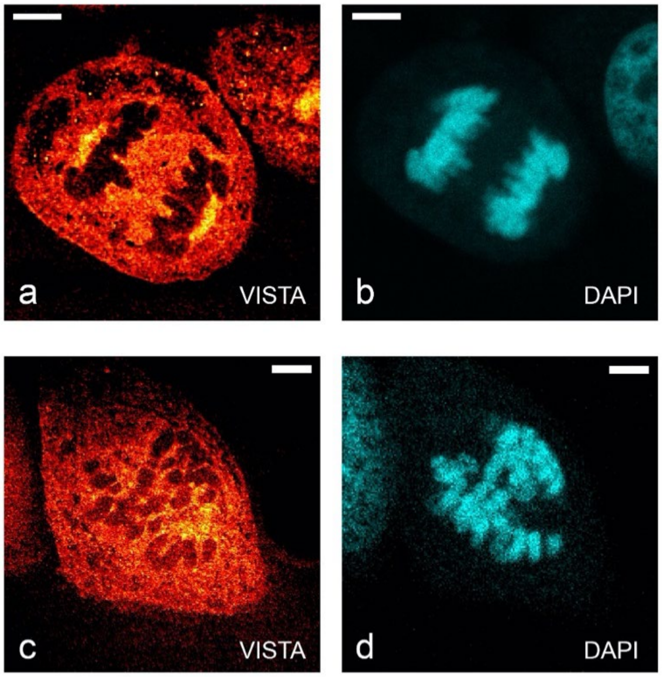


Figure S8. Correlative VISTA images and the fluorescent images with DAPI stain on mitotic HeLa cells. Fluorescent DAPI stains confirms that the dark regions shown in VISTA images on anaphase (a-b) and prophase (c-d) HeLa cells are chromosomes with a low CH_3 signals (i.e., low protein contents). Scale bars: 20 μm. The length scale is in terms of distance after sample expansion.

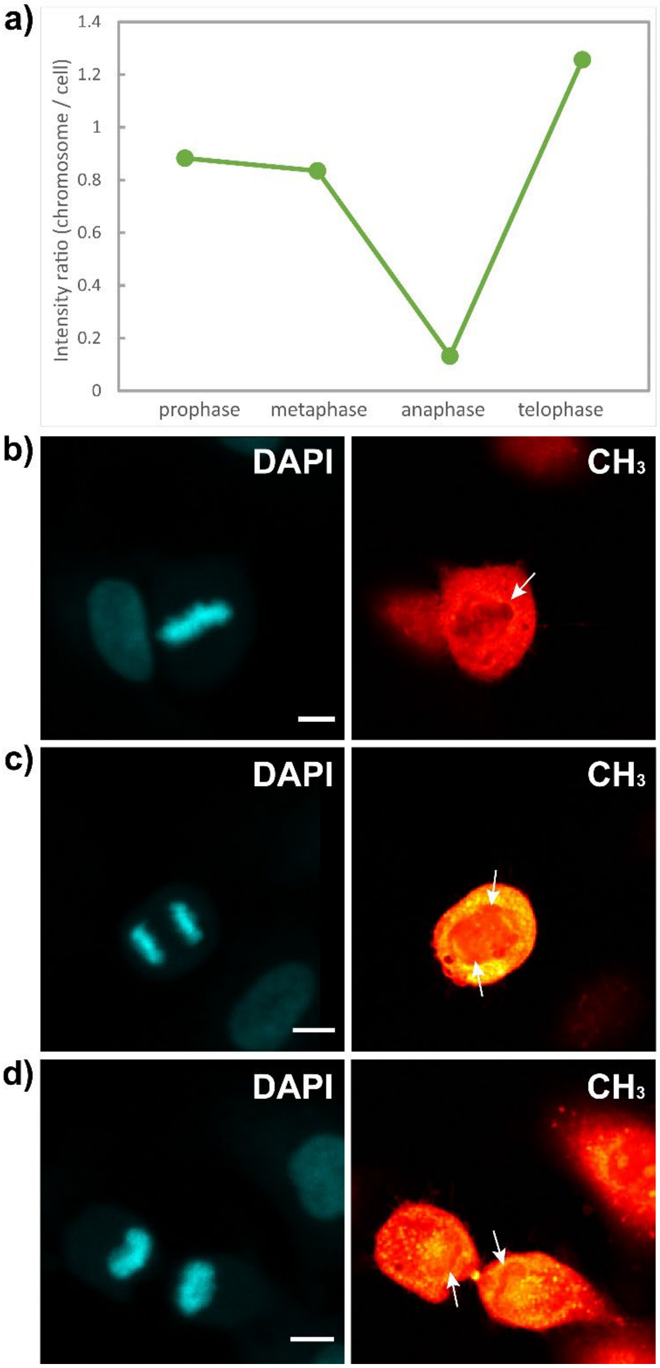


Figure S9. Confirmation of protein abundance change in the chromosomal regions in HeLa cells throughout mitosis. (a) Quantification of relative protein abundance in chromosomal regions in reference to the protein CH_3 signals from the whole cells (i.e., the chromosome/cell ratio) for expanded HeLa cells across different stages of mitosis, shown in Fig. 2a-d. (prophase: 0.88; metaphase: 0.83; anaphase: 0.14; telophase: 1.26). (b-d) SRS images (CH_3, 2940 cm^{-1}) and correlative DAPI-stain fluorescence images of PFA-fixed but unprocessed mitotic HeLa cells in metaphase (b); anaphase (c); and telophase (d). Arrows indicate corresponding chromosome regions in the SRS images guided by DAPI fluorescence stain. Scale bars: 10 μm.

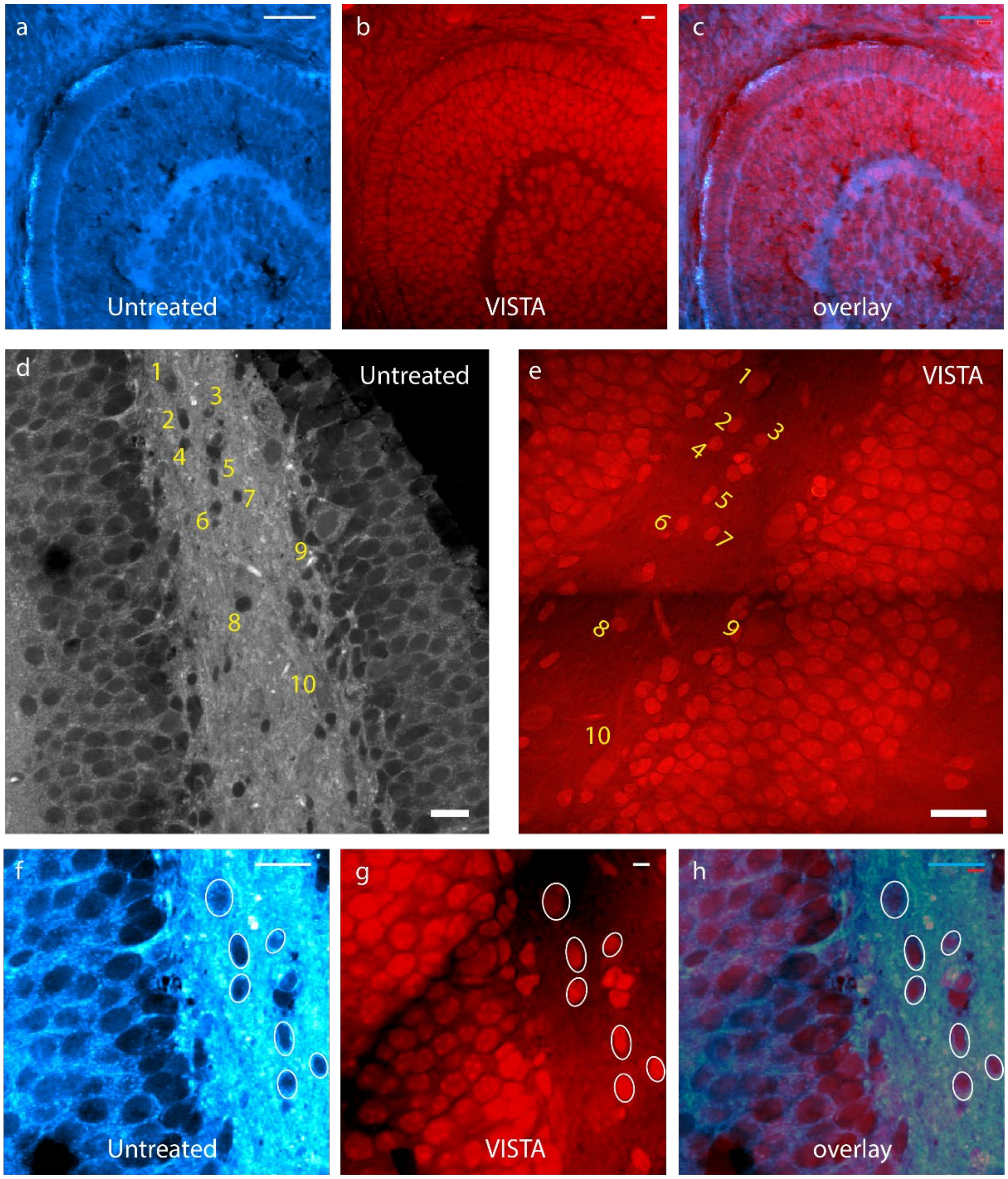


Figure S10. Imaging registrations for tissues before and after expansion. (a-c) SRS imaging of an untreated zebrafish retina tissue (a), VISTA imaging of the same sample (b), and the overlay (e). (d-e) SRS imaging of an untreated mouse brain tissue (a) and VISTA image of the same sample (b), with labels of 9 correlating features for expansion ratio calculation. (f-h) SRS imaging of an untreated brain tissue (a), VISTA imaging of the same sample (b), and the overlay (e). White circles indicate cell bodies to guide the registration. Scale bars: 20 μm (a-d,f-h),80 μm (e).

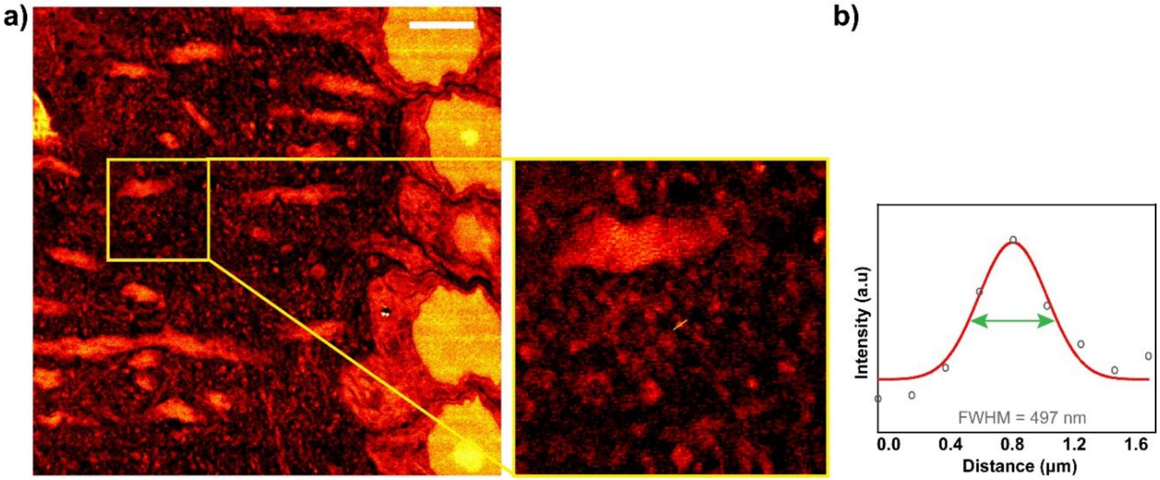


Figure S11. Characterizations of fine dendritic structures by VISTA. (a) A representative VISTA image at the CA1 region of a mouse hippocampus sample with a zoomed-in view of fine structural details in the yellow box. (b) The cross-section profile of a dendritic structure with a FWHM of 497 nm (corresponding to a lateral size of ~175 nm for the unexpanded sample). Scale bars: 20 μm.

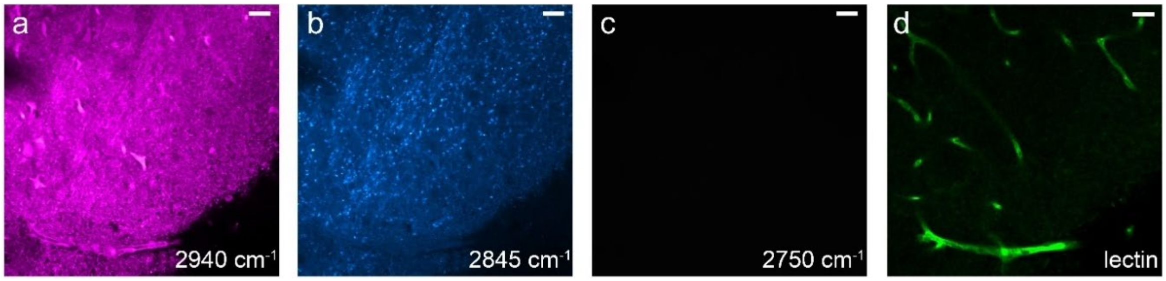


Figure S12. Regular-resolution SRS imaging of a PFA-fixed but unprocessed mouse brain tissue. a-c, SRS imaging at 2940 cm^{-1} (a, CH_3), 2845 cm^{-1} (b, CH_2), and 2750 cm^{-1} (c, off-resonance image). Cells, processes and vessels could not be clearly identified from these

SRS images. d, fluorescence image of lectin-stained blood vessel on the same tissue. Scale bars: 20 μm.

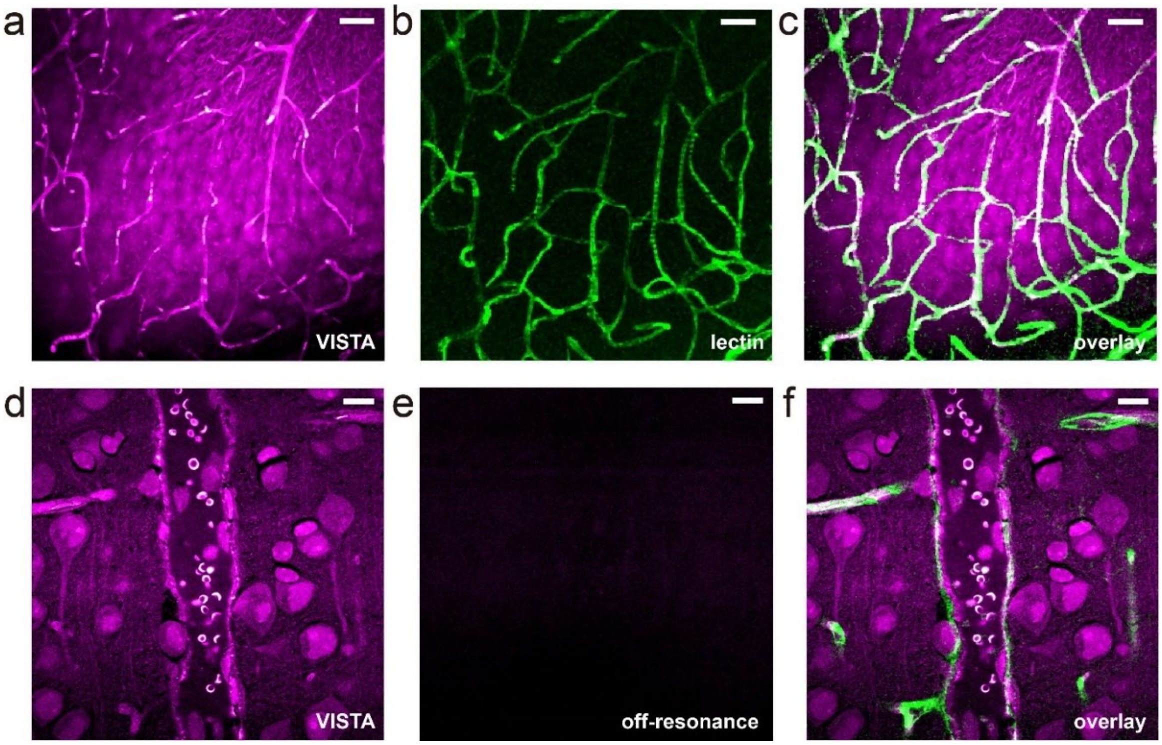


Figure S13. Correlative VISTA images and the fluorescent images with lectin-stained vessels in mouse brain tissues. (a-c) The overlay image (c) for parallel images of VISTA (a, shown in Fig. 3a) and fluorescence from lectin-DyLight594 stained blood-vessels (b, shown in Fig. 3b). (d-e) VISTA image of a larger vessel (likely an artery) showing red blood cells (d, shown in Fig. 3c) and the corresponding off-resonance image at 2810 cm^{-1} confirms that signals in the VISTA image (including that of the red blood cells) are exclusively SRS signals. (f) The overlay image for VISTA (Fig. 3c) and lectin-DyLight594 stained fluorescence (Fig. 3d). Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

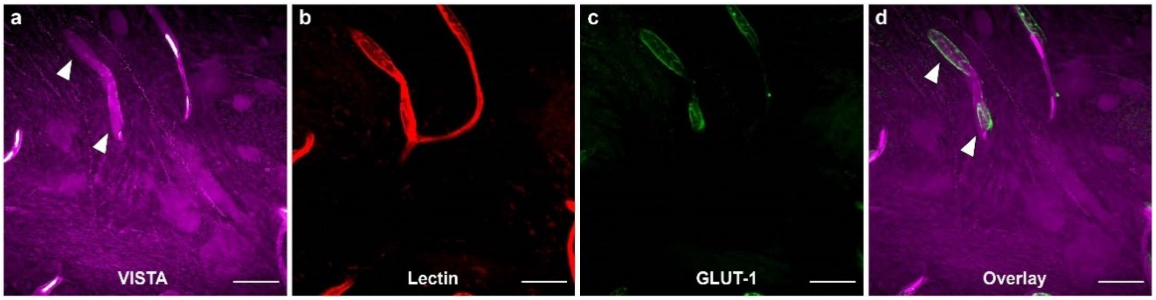


Figure S14. Immuno-fluorescence reveals the cells in VISTA other than neuronal cells to be vascular endothelial cells. Elongated elliptical nuclei shown in VISTA (a, white arrow-headed, VISTA) co-localize (d, Overlay) with both fluorescent lectin-stained vessels (b, Lectin), and GLUT-1 (c, glucose transporter 1, brain endothelial cell marker) immuno-fluorescence-stained brain vascular endothelial cells. Images are shown as maximum Z projection. Scale bars: 50 μm. The length scale is in terms of distance after sample expansion.

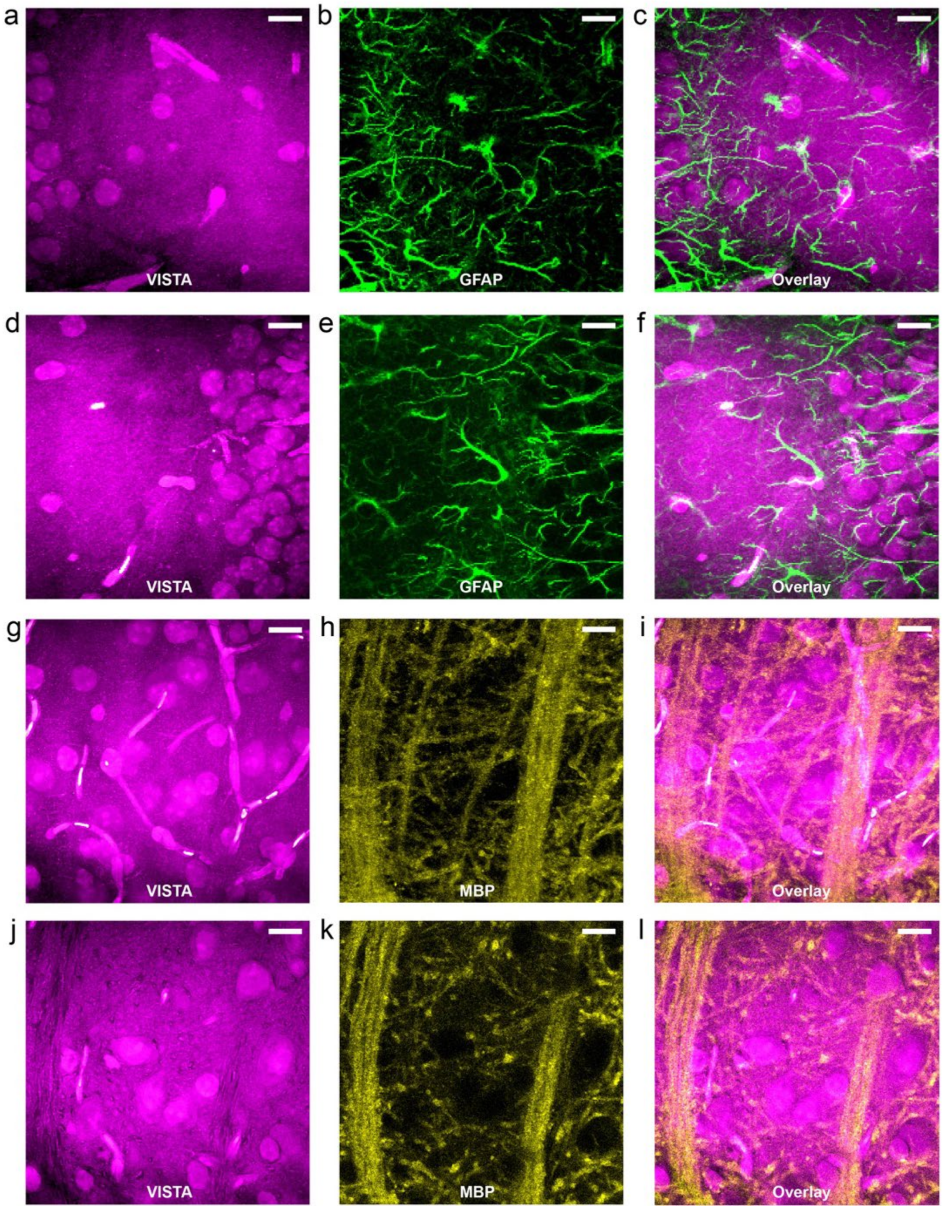


Figure S15. Immunofluorescence confirms the absence of cytoplasmic structures from astrocytes and oligodendrocytes in VISTA. (a-f) Maximum z projection of VISTA image (a, d), immuno-stained fluorescence image with GFAP (Glial Fibrillary Acidic Protein, the astrocyte cellular maker) antibodies (b, e) and the overlay (c, f). (g-i) maximum z projection of VISTA image (g), immuno-stained fluorescence image with MBP (Myelin Basic Protein, the myelin and oligodendrocyte cellular maker) antibodies (h) and the overlay (i). (j-l) single-slice VISTA image (j), immuno-stained fluorescence image with MBP antibodies (k) and the overlay (l). No correlation is identifiable between VISTA revealed cytoplasmic structures and GFAP or MBP stained cellular fluorescence images. Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

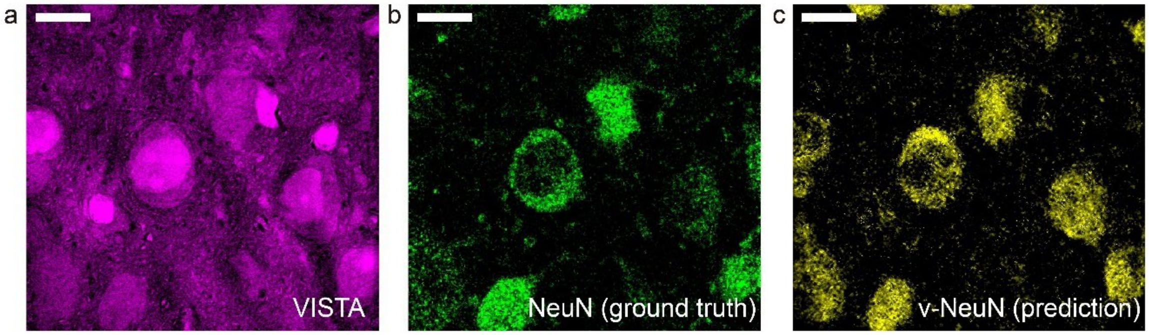


Figure S16. Label-free VISTA prediction for neuronal cell bodies in brain tissues. a-c, The input VISTA image (a), the ground truth fluorescence image of NeuN stained matured neurons (b) and the predicted VISTA-NeuN (v-NeuN) image of matured neurons from a (c). Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

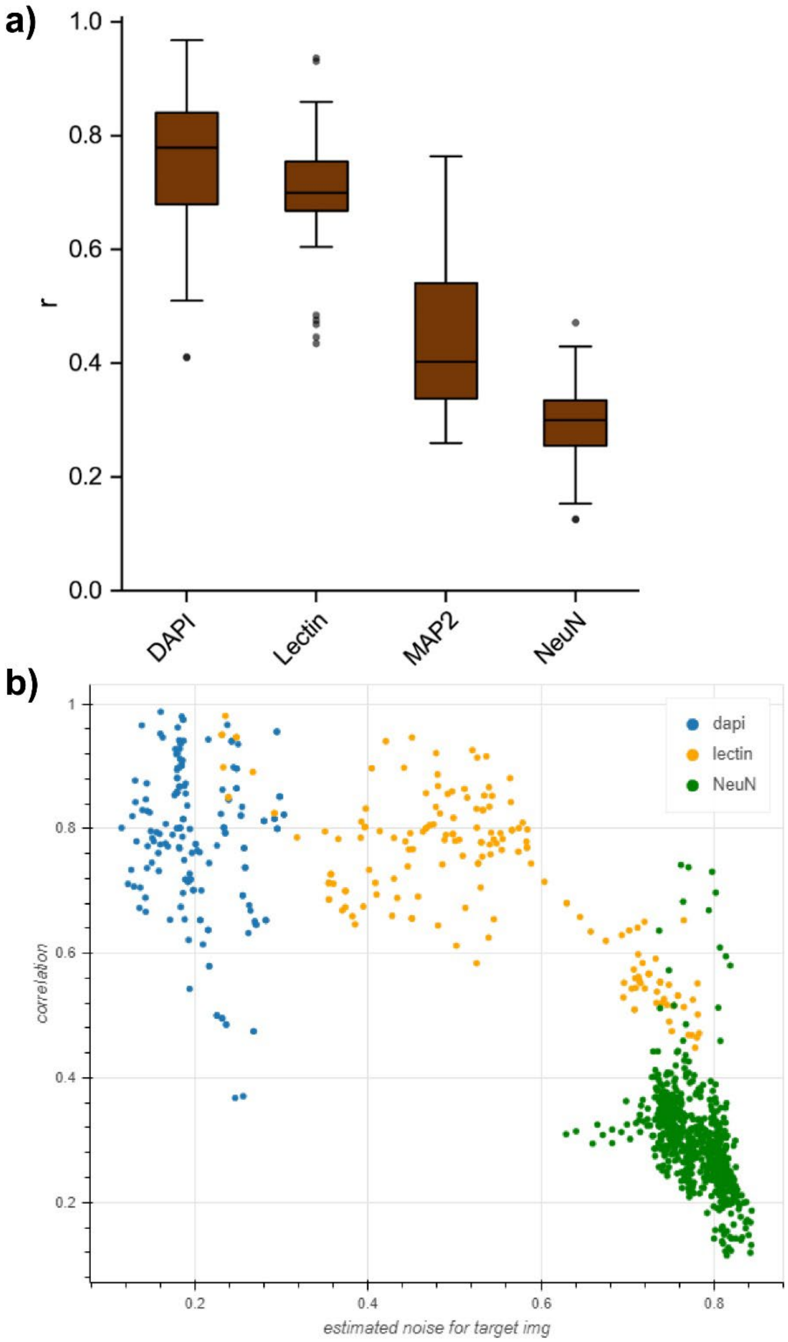


Figure S17. Model performance for the U-Net models, quantified by Pearson correlation coefficient (r). (a) Each point in the plot represents a target/predicted image pair. The boxes indicate the 25th, 50th, and 75th percentile of the Pearson's r for each model, with whiskers with maximum 1.5 interquartile range. Points within the box are not shown. DAPI, N=48; Lectin, N=37; MAP2, N=54; NeuN, N=130. (b) Correlation between the Pearson's r for the prediction results and the estimated standard deviation of the noise for ground-truth fluorescence images for DAPI, Lectin and NeuN.

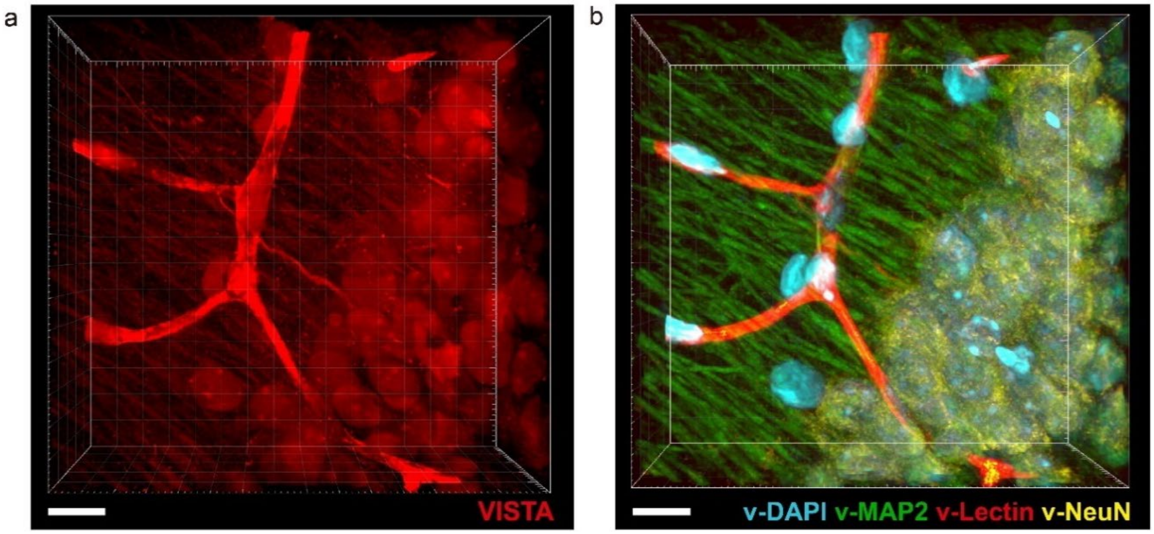


Figure S18. Label-free VISTA prediction for multi-component imaging of mouse hippocampus. (a) The input VISTA image; (b) the predicted multicolor image with predicted v-DAPI (cyan), v-MAP2 (green), v-lectin (red) and v-NeuN (yellow) components, as shown in Fig. 4j. Scale bars: 40 µm. The length scale is in terms of distance after sample expansion.

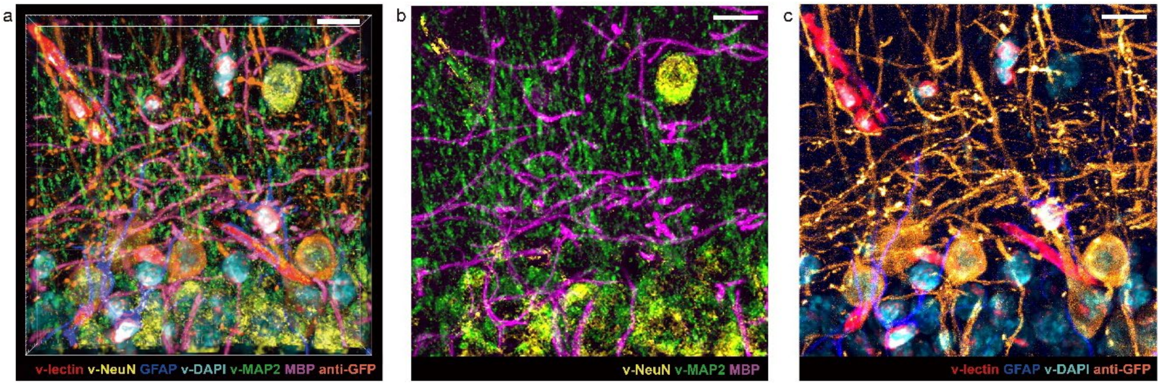


Figure S19. 7-color multiplex imaging combining both label-free VISTA prediction and fluorescence imaging in hippocampus of Thy1-YFP mouse. (a) Volume 3D presentation of 7-color overlay image. (b-c) Maximum z projection view of 7 components in two sets of 3-color (b) and 4-color (c) overlay. VISTA components: v-NeuN (yellow), v-MAP2 (green), v-DAPI (cyan), v-lectin (red); immuno-fluorescence components: GFAP (blue), MBP (magenta) and anti-GFP (orange). Scale bars: 40 μm. The length scale is in terms of distance after sample expansion.

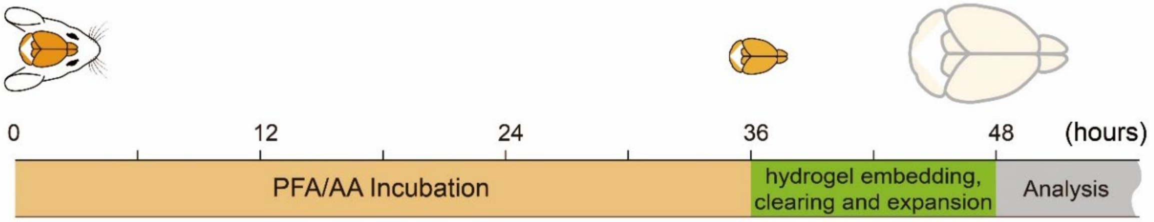


Figure S20. Workflow of the sample processing steps for VISTA. Typical sample processing steps were completed with a 48-hour for mouse brain tissues before VISTA analysis.

3.6 Method and materials

Stimulated Raman scattering (SRS) Microscopy. A picoEmerald laser system (Applied Physics & Electronics) is used as the light source for SRS microscopy. It produces 2 ps pump (tunable from 770 nm – 990 nm, bandwidth 0.5 nm, spectral bandwidth ~ 7 cm^{-1}) and Stokes (1031.2 nm, spectral bandwidth 10 cm^{-1}) beams with 80 MHz repetition rate. Stokes beam is modulated at 20 MHz by an internal electro-optic modulator. The spatially and temporally overlapped pump and Stokes beams are introduced into an inverted multiphoton laser scanning microscope (FV3000, Olympus), and then focused onto the sample by a 25X water objective (XLPLN25XWMP, 1.05 N.A., Olympus) for imaging. Transmitted Pump and Stokes beams are collected by a high N.A. condenser lens (oil immersion, 1.4 N.A., Olympus) and pass through a bandpass filter (893/209 BrightLine, 25 mm, Semrock) to filter out Stokes beam. A large area (10×10 mm) Si photodiode (S3590-09, Hamamatsu) is used to measure the pump beam intensity. A 64 V reverse-bias DC voltage is applied on the photodiode to increase saturation threshold and reduce response time. The output current is terminated by a 50-Ω terminator and pre-filtered by a 19.2-23.6-MHz band-pass filter (BBP-21.4+, Mini-Circuits) to reduce laser and scanning noise.

The signal is then demodulated by a lock-in amplifier (SR844, Stanford Research Systems) at the modulation frequency. The in-phase X output is fed back to the Olympus IO interface box (FV30-ANALOG) of the microscope. Image acquisition speed is limited by 30 μs time constant set for the lock-in amplifier. Correspondingly, we use 80 μs pixel dwell time, which gives a speed of 21s per frame for a 512-by-512-pixel field of view. Pump laser is tuned to 791.3 nm for imaging protein CH_3 vibrational mode at 2940 cm^{-1}. Laser powers on sample are measured to be 30 mW for Pump beam and 200 mW for modulated Stokes beam. 16-bit grey scale images are acquired by Olympus Fluoview 3000 software. Volumetric images were acquired by collecting a z-stack with a step size of 1 micron in z direction.

Reagents and materials. Sodium acrylate (SA, Sigma-Aldrich), acryl amide (AA, Sigma-Aldrich), N,N′-methylenebisacrylamide (BIS, 2%; Sigma-Aldrich), ammonium persulfate (APS, Sigma-Aldrich), tetramethylethylenediamine (TEMED, Sigma-Aldrich), sodium dodecyl sulfate (SDS), Triton X-100, Tween-20, and deuterium oxide were obtained from Sigma-Aldrich, and 1.0 M Tris was obtained from Biosolve. Nuclease-free water was purchased from Ambion–Thermo Fisher. Acrylamide (2,3,3-D3) was obtained from Cambridge Isotope Laboratories. Deuterated sodium acrylate was prepared from acrylic acid (2,3,3-D3, Cambridge Isotope Laboratories) and sodium hydroxide (Sigma-Aldrich). DAPI was purchased from Thermo Fisher (D1306, Thermo Fisher). Primary antibodies: anti-myelin basic protein in rat (Abcam, ab7349); anti-myelin basic protein in rabbit (Abcam, ab40390); anti-GFAP in chicken (Abcam, ab4674); anti-chicken IgY, Alexa 488 (Invitrogen, A-32931); anti-GFP Alexa Fluor 647 (Invitrogen, A-31852); anti-MAP2 (Cell Signaling Technology,8707); anti-NeuN (Cell Signaling Technology, 24307); anti-GLUT-1 (Millipore Sigma, 07-1401); Lycopersicon Esculentum Lectin DyLight®594 (Vector Laboratories, DL-1177-1). Secondary antibodies: goat anti-rat IgG, Alexa Fluor 568 (Invitrogen, A-11077); goat anti-mouse IgG, Alexa Fluor 647 (Invitrogen, A-21236); goat anti-rabbit IgG, Alexa Fluor 488 (Invitrogen, A-11034); goat anti-chicken IgY, Alexa Fluor 647 (Invitrogen, A-21449).

Hydrogel embedding: gelation, denaturation and expansion. Stock solutions include an incubation solution (30% AA in 4% PFA), and a gelation solution (7% SA, 20% AA, 0.1% BIS in 1x PBS) were made and stored in 4 ºC and -20 ºC, respectively. The free-radical initiator APS and accelerator TEMED were dissolved and diluted in nuclease-free water to a concentration of 10% (w/w), and stored at -20 ºC as stocks. Prior to a typical hydrogel embedding step, the cell or tissue samples were incubated in a solution of 30% AA in 4% PFA under different conditions depending on the sample type (detailed in the following sections). The gelation solution, the free-radical initiator and accelerator were thawed and kept at 4 ºC before the gelation step. Coverslips with the cell or tissue samples were placed at the bottom of a pre-fabricated and pre-cooled gelation chamber, with the sample facing upward. After adding sufficient amount of gelation solution (7% SA, 20% AA, 0.1% BIS in 1x PBS) to fully immerse the sample, a layer of flat Parafilm covered coverslip was placed on top of the chamber as the lid. The chamber was kept at 4 ºC for 1 min as the free-radical polymerization proceeded, before it was transferred to a humid incubator for a following incubation at 37 ºC for 1 h. Coverslips with gels were then incubated in sufficient amount of denaturing buffer (200 mM SDS, 200 mM NaCl, and 50 mM Tris in nuclease-free water, pH 8) in petri dishes for 15 min at room temperature. The volume of the denaturing buffer and the size of the petri dishes used were determined by the sample size and thickness after denaturation. In 15 min, gels would be detached from the coverslips. They were transferred into 1.5-ml Eppendorf centrifuge tubes filled with denaturing buffer and were incubated under different conditions depending on the sample type (detailed in the following sections). Initial expansions were carried out immediately after the denaturation step at ambient temperature, in H_2O or D_2O, which was changed twice in 1 h. The expanded gel was then kept in H_2O or D_2O overnight and stores in dark. The gel expanded 4.0±0.22 times in our experiments.

Cultured HeLa cell experiments. In mammalian cell experiments, cultured HeLa-CCL2 (ATCC) were seeded onto coverslips (12mm, #1.5, Fisher) for 24 h. Cells were first grown in regular DMEM medium supplemented with 10% FBS and 1% penicillin-streptomycin antibiotics until they reached 70-90% confluency. Coverslips with HeLa cells were incubated in a solution of 4% PFA with 30% AA in PBS for 7–8 h at 37 °C, without normal fixation with PFA. Hydrogel embedding, including gelation, denaturation and expansion, was processed in above-mentioned steps. The denaturation after transfer into Eppendorf centrifuge tubes was at 95 ºC for 30 min.

Normal brain tissue experiments. Mouse brains tissues harvested from mice were washed once with DPBS on ice and immediately incubated in a solution of 4% PFA with 30% AA in PBS for 24-30 h at 4 °C, then transferred to a shaker and further incubated for 12 h at 37 °C with gentle shaking. After the incubation step the tissue sample was cut into 100 µm to 250 µm thin slices (and thick slices up to 600-650 µm, for thick brain tissue experiments) using a Leica VT 1200S vibratome. Hydrogel embedding, including gelation, denaturation and expansion, was processed in above-mentioned steps. The denaturation after transfer into Eppendorf centrifuge tubes was first at 70 ºC for 3 h and then at 95 ºC for 1 h.

Zebra fish embryo experiments. Fresh zebra fish embryo samples were embedded in gelatin and snap frozen in liquid nitrogen and stored at -80 ºC until ready for sectioning. Immediately before sectioning, allow the whole gelatin embedded sample warm up to -30 ºC for 10 min in the cryostat. Cryosectioning was carried out on the whole frozen block, and sectioned slices were collected onto the glass slides at room temperature. The glass slides with collected sample slices were stored at -20 ºC until ready for de-gelatinization and hydrogel embedding. When ready, the slides were warmed up to room temperature, and then de-gelatinized by incubation in PBS at 42 ºC for 30 min. The de-gelatinized samples were washed in PBS with 0.1% Tween-20 before processing. Hydrogel embedding, including gelation, denaturation and expansion, was processed in above-mentioned steps. For thin (50-µm) slices of the zebra fish embryo we used, the denaturation after transfer into Eppendorf centrifuge tubes was at 95 ºC for 30 min.

Thick brain tissue experiments. Thick brain tissues were cut to slices with thickness of at least 250 µm (1000 µm after expansion). PFA+AA incubation and hydrogel embedding, including gelation, denaturation, and expansion, was processed in above-mentioned steps similarly to the normal tissue processing. The denaturation after transfer into Eppendorf centrifuge tubes was first at 70 ºC for 3 h and then at 95 ºC for 4h. Thicker tissue can also be imaged by VISTA with longer 95 ºC denaturation time.

Immunostaining. In an immuno-labeling process, 60-µm- to 150-µm-thick mouse brain coronal slices were embedded in a hydrogel, denatured, pre-incubated with PBS with 1% (wt/vol) Triton X-100 (PBST) for 15 min, and subsequently incubated with primary antibodies at a typical 1:100 dilution with PBST at 37 ºC for 16h, followed by washing with PBST three times at 37 ºC for 1-2 h. The samples were then incubated with secondary antibodies at a 1:100 dilution with PBST at 37 ºC for 12-16 h, followed by washing with PBST three times at 37 ºC for 1-2 h.

Sample mounting and imaging. Expanded cell or tissue samples were kept in D_2O/H_2O for imaging. Grace Bio-Labs Press-To-Seal silicone isolators with appropriate opening sizes and depths were used as spacers between microscope slides (1 mm, VWR) and coverslips (12 mm, #1.5, Fisher). In particular, the thick brain tissue samples (after expansion) were placed in a 4-mm silicon isolator to avoid any pressure and damage to the sample. For control experiments on normal PFA fixed HeLa cells, 0.5-mm Press-To-Seal silicone isolators or the common Grace Bio-Labs SecureSeal™ spacers were used. Confocal images were obtained by the Olympus FluoView™ FV3000 confocal microscope with SRS setup described above.

SRS imaging of beads. Polystyrene beads (0.1 μm mean particle size, Sigma-Aldrich, Inc.) were resuspended in deionized water by a 1:2000 dilution. The resuspension step required vortexing and sonication for 20 min at room temperature. Before SRS imaging, the diluted beads suspension was sealed between a glass slide and a coverslip, which was then stored in dark and left to settle overnight.

Image processing and data analysis. Images color-coding and intensity profile were done by ImageJ. Intensity normalization of the z-stacks was done in ImageJ. 3D rendering of z-stacks was done in Imaris View. Data plotting and analysis, including spectral plots and Gaussian fitting were performed in OriginLab.

Calculating the resolution of SRS microscopy and the VISTA method. To determine the point spread function (PSF) of the imaging system, deconvolution of the actual size of the beads was done by simulations in MATLAB 2019b. Line profile of the bead image was fitted by Gaussian approximation and bead object was modelled as a circle.

Fluorescence imaging. The fluorescence images of processed samples with fluorescent labels were obtained with a 25×, 1.05 NA water-immersion objective with the Olympus Fluoview system. Single-photon confocal laser scanning imaging was performed with 405-, 488-, 561-, and 640-nm lasers (Coherent OBIS). The images were visualized and analyzed with Fiji or Imaris Viewer.

U-Net construction, training and prediction for label-free imaging. The prediction of subcellular structures from SRS images was based on a U-Net CNN convolutional neural network demonstrated by Ounkomol et al.[17] Training data were collected by sequentially acquiring respective fluorescence targets and protein SRS images on the same field of view. Image sets were generated by performing a z scan with a z-direction step size of 1 micron. Such a step size is larger than the axial resolutions of both fluorescence and SRS. Before training, the fluorescence and SRS images were background subtracted in imageJ with a rolling ball radius of 50 pixels (0.497 micron/pixel) before training. The image sets (minimum 200 sets for each channel) were separated randomly at a 1:3 ratio for testing and training set, respectively. The models were trained by batches of 128 pixels x 128 pixels patches subsampled from the original images. The training was performed using the Adam optimizer to optimize the mean squared error between the fluorescence image and the predicted image. The learning rate set at 0.001 and trained for 50,000 epochs with batch size of 32 images. All the trainings and predictions were run on a node of High Performance Computing Center at Caltech equipped with Nvidia P100 GPU containing 16 GB of memory. The trained algorithm was applied across different samples with no further adjustments.

Model accuracy. Model accuracy was quantified by the Pearson correlation coefficient:

$$r = \frac{\sum(x-\overline{x})\,(y-\overline{y})}{\sqrt{\sum(x-\overline{x})^2\sum(y-\overline{y})^2}}$$

between the pixel intensities of the model's output, y, and independent ground-truth test images, x, for all the images of the test sets except for lectin. The Pearson correlation coefficient for lectin was calculated for images with lectin signal, since the random noise in the background could not be predicted by the model.

We also quantified the contribution of the noise from the ground-truth fluorescence images to the prediction performance (i.e., the Pearson's r) by our model.[39] We establish that our model accuracy (i.e., the Pearson's r) decreases as the noise of training-set ground truth increases.

Data and code availability. The authors declare that all data supporting the findings of the present study are available in the article and its supplementary figures and tables, or from the corresponding author upon request. MATLAB code used for PSF determination and Python code for U-Net training and prediction in this paper is available at https://github.com/Li-En-Good/VISTA. (10.5281/zenodo.4717251)

Animals. All animal procedures performed in this study were approved by the California Institute of Technology Institutional Animal Care and Use Committee (IACUC), and we have complied with all relevant ethical regulations. The C57BL/6J (000664), Thy1-YFP (003709) mouse lines used in this study were purchased from the Jackson Laboratory (JAX) and bred in our animal facility. 6- to 8-week-old, male and female C57BL/6J, Thy1-YFP mice were used for tissue collection. At the day of collection, the mice were anesthetized with Euthasol (pentobarbital sodium and phenytoin sodium solution, Virbac AH) and transcardially perfused with 30–50 mL of 0.1 M phosphate buffered saline (PBS) (pH 7.4). After this procedure, the brains were harvested and proceeded to VISTA processing.

Statistics and Reproducibility. For all imaging experiments yielding the micrographs reported herein, at least three independent experiments were repeated with similar results.

3.7 Reference

1. Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. *Annu. Rev. Biochem.* **78**, 993–1016 (2009).
2. Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nat. Rev. Mol. Cell Biol.* **18**, 685–701 (2017).
3. Wassie, A. T., Zhao, Y. & Boyden, E. S. Expansion microscopy: principles and uses in biological research. *Nat. Methods* **16**, 33–41 (2019).
4. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
5. Tillberg, P. W. *et al.* Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016).
6. Murray, E. *et al.* Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell* **163**, 1500–1514 (2015).
7. Kim, S.-Y. *et al.* Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proc. Natl. Acad. Sci. U.S.A* **112**, E6274–E6283 (2015).
8. Shen, Y., Hu, F. & Min, W. Raman imaging of small biomolecules. *Annu. Rev. Biophys* **48**, 347–369 (2019).
9. Cheng, J.-X. & Xie, X. S. Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine. *Science* **350**, (2015).
10. Saar, B. G. *et al.* Video-Rate Molecular imaging in vivo with stimulated Raman scattering. *Science* **330**, 1368–1370 (2010).
11. Gasecka, A., Daradich, A., Dehez, H., Piché, M. & Côté, D. Resolution and contrast enhancement in coherent anti-Stokes Raman-scattering microscopy. *Opt. Lett.* **38**, 4510–4513 (2013).
12. Silva, W. R., Graefe, C. T. & Frontiera, R. R. Toward label-free super-resolution microscopy. *ACS Photonics* **3**, 79–86 (2016).

13. Gong, L., Zheng, W., Ma, Y. & Huang, Z. Higher-order coherent anti-Stokes Raman scattering microscopy realizes label-free super-resolution vibrational imaging. *Nat. Photonics* **14**, 115–122 (2020).
14. Gong, L., Zheng, W., Ma, Y. & Huang, Z. Saturated stimulated-Raman-scattering microscopy for far-field superresolution vibrational imaging. *Phys. Rev. Appl.* **11**, (2019).
15. Watanabe, K. *et al.* Structured line illumination Raman microscopy. *Nat. Comm.* **6**, 10095 (2015).
16. Gradinaru, V., Treweek, J., Overton, K. & Deisseroth, K. Hydrogel-tissue chemistry: principles and applications. *Annu. Rev. Biophys.* **47**, 355–376 (2018).
17. Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).
18. Shi, X. et al. Label-retention expansion microscopy. *J. Cell Biol.* **220**, e202105067 (2021).
19. Puchtler, H. & Meloan, S. N. On the chemistry of formaldehyde fixation and its effects on immunohistochemical reactions. *Histochemistry* **82**, 201–204 (1985).
20. Ku, T. *et al.* Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* **34**, 973–981 (2016).
21. Gambarotto, D. *et al.* Imaging cellular ultrastructures using expansion microscopy (U-ExM). *Nat. Methods* **16**, 71–74 (2019).
22. Angueyra, J. M. & Kindt, K. S. Leveraging zebrafish to study retinal degenerations. *Front. Cell Dev. Biol.* **6**, (2018).
23. Dvoriantchikova, G., Seemungal, R. J. & Ivanov, D. The epigenetic basis for the impaired ability of adult murine retinal pigment epithelium cells to regenerate retinal tissue. *Sci. Rep.* **9**, 3860 (2019).
24. Kandel, E. R. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294**, 1030–1038 (2001).
25. Liebmann, T. *et al.* Three-dimensional study of Alzheimer's Disease hallmarks using the iDISCO clearing method. *Cell Rep.* **16**, 1138–1152 (2016).
26. Wei, M. *et al.* Volumetric chemical imaging by clearing-enhanced stimulated Raman scattering microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6608–6617 (2019).
27. Hama, H. *et al.* ScaleS: an optical clearing palette for biological imaging. *Nat. Neurosci.* **18**, 1518–1529 (2015).
28. Kirst, C. *et al.* Mapping the fine-Scale organization and plasticity of the brain vasculature. *Cell* **180**, 780-795.e25 (2020).
29. Helmstaedter, M. Cellular-resolution connectomics: challenges of dense neural circuit reconstruction. *Nat. Methods* **10**, 501–507 (2013).
30. Ueda, H. R. *et al.* Tissue clearing and its applications in neuroscience. *Nat. Rev. Neurosci."* **21**, 61–79 (2020).
31. Todorov, M. I. *et al.* Machine learning analysis of whole mouse brain vasculature. *Nat. Methods* **17**, 442–449 (2020).
32. Hollon, T. C. *et al.* Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat. Med.* **26**, 52–58 (2020).

33. Zhang, J., Zhao, J., Lin, H., Tan, Y. & Cheng, J.-X. High-speed chemical imaging by dense-net learning of femtosecond stimulated Raman scattering. *J. Phys. Chem. Lett.* **11**, 8573–8578 (2020).
34. Lin, H. *et al.* Fingerprint spectroscopic SRS imaging of single living cells and whole brain by ultrafast tuning and spatial-spectral learning. *arXiv* preprint arXiv:2003.02224. (2020).
35. Manifold, B., Thomas, E., Francis, A. T., Hill, A. H. & Fu, D. Denoising of stimulated Raman scattering microscopy images via deep learning. *Biomed. Opt. Express* **10**, 3860–3874 (2019).
36. Falk, T. *et al.* U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).
37. Bi, Y. *et al.* Near-resonance enhanced label-free stimulated Raman scattering microscopy with spatial resolution near 130 nm. *Light Sci. Appl.* **7**, (2018).
38. Sun, D. et al. Click-ExM enables expansion microscopy for all biomolecules. *Nat. Methods* **18**, 107–113 (2021).
39. Immerkær, J. “Fast Noise Variance Estimation,” *Comput. Vis. Image Underst*. 64, 300–302 (1996).

Chapter 4

Label-free super-resolution imaging of protein aggregates by VISTA (Vibrational Imaging of Swelled Tissue and Analysis)

The content of this chapter has been published in:
Miao, K.; Lin, L.-E.; Qian, C.; Wei, L. Label-Free Super-Resolution Imaging Enabled by Vibrational Imaging of Swelled Tissue and Analysis. *JoVE (Journal of Visualized Experiments).* **2022**, No. 183, e63824. https://doi.org/10.3791/63824.

4.1 Summary

The universal utilization of fluorescence microscopy, especially super-resolution microscopy, has greatly advanced our knowledge about modern biology. Conversely, the requirement of fluorophore labeling in fluorescent techniques poses significant challenges, such as photobleaching and non-uniform labeling of fluorescent probes and prolonged sample processing. In this protocol, we report the detailed working procedures, termed VISTA, to circumvent obstacles associated with fluorophores and achieve label-free super-resolution volumetric imaging in biological samples with spatial resolution down to 78 nm. VISTA is established by embedding cells and tissues in hydrogel, isotropically expanding the hydrogel-sample hybrid, and visualizing endogenous protein distributions by vibrational imaging with stimulated Raman scattering microscopy. We demonstrated VISTA on both cells and mouse brain tissues and observed highly correlative VISTA and immunofluorescence images, validating the protein origin of imaging specificities. We exploited such correlation and trained a machine-learning based image-segmentation algorithm to achieve multi-component prediction of nuclei, blood vessels, neuronal cells and dendrites from label-free mouse brain images. We further adapted VISTA to investigate pathologic poly-glutamine (polyQ) aggregates in cells and amyloid-beta (Aβ) plaques in brain tissues with high throughput, justifying its potential for large-scale clinical samples.

4.2 Introduction

The development of optical imaging methods has revolutionized our understanding of modern biology because they provide unprecedented spatial and temporal information of targets across different scales, from subcellular proteins to whole organs[1]. Among them, fluorescence microscopy is the most well-established: with a large palette of organic dyes with high extinction coefficients and quantum yields[2], easy-to-use genetic encoded fluorescent proteins[3], and super-resolution methods such as STED, PALM and STORM for imaging nanometer-scale structures[4,5]. In addition, recent advancement in sample engineering and preservation chemistry, which expands specimens embedded in swellable polymer hydrogels[6-8], enables sub-diffraction limited resolution on conventional fluorescence microscopes. For instance, typical expansion microscopy (ExM) effectively enhances the image resolution by 4 times with fourfold isotropic sample expansion[7].

Despite its advantages, super-resolution fluorescence microscopy shares limitations that originate from fluorophore labeling. First, photobleaching and inactivation of fluorophores compromises the capacity of repetitive and quantitative fluorescence evaluations. Photobleaching is an inevitable event when light keeps pumping electrons into electronic excited states[9]. Second, labeling the fluorophores to the desired targets is not always a straightforward task. For instance, immunostaining demands a long and laborious sample preparation process and hinders imaging throughput[10]. It could also introduce artifacts due to inhomogeneous antibody-labeling, especially deep inside tissues[11]. Moreover, proper labeling strategies that targets fluorophores to desired proteins might be underdeveloped. For example, extensive screenings were required to find effective antibodies for Aβ plaques[12]. Smaller organic dyes, such as Congo red, often have limited specificity, only staining the core of the Aβ plaque. We aim to develop a label-free super-resolution modality that circumvents the drawbacks from fluorophore-labeling and provides complementary high-resolution imaging from cells to tissues, and even to large-scale human samples.

Raman microscopy provides label-free contrast for chemical-specific structures and maps out the distribution of otherwise invisible chemical bonds by looking at the excited vibrational transitions[13]. In particular, stimulated Raman scattering (SRS) imaging on label-free or tiny-labeled samples has been demonstrated to have similar speed and resolution to fluorescence microscopy[14,15]. For example, healthy brain region has been readily differentiated from tumor infiltrated region in human and mouse tissues[16,17]. Aβ plaques was also clearly imaged by targeting protein CH_3 vibration (2940 cm^{-1}) and amide I (1660 cm^{-1}) on a fresh-frozen brain slice without any labeling[18]. Raman scattering therefore offers robust label-free contrast that overcomes the limitations of fluorophores. The question then became how we can accomplish super-resolution capacity using Raman scattering, which could reveal nanoscale structural details and functional implications in biological samples.

Although extensive efforts have been made to achieve super-resolution for Raman microscopy with elegant optic instrumentations, the resolution enhancement on biological samples have been rather limited[19-21]. Here, based on our recent works[22,23], we present a protocol that combines a sample-expansion strategy with stimulated Raman scattering for super-resolution label-free vibrational imaging, named *V*ibrational *I*maging of *S*welled *T*issues and *A*nalysis (VISTA). We embedded cells and tissues in hydrogel matrixes through an optimized protein-hydrogel hybridization protocol. We then incubated hydrogel tissue hybrids in the detergent-rich solutions for delipidation followed by expansion in water. The expanded samples were then imaged by a regular SRS microscope by targeting CH_3 vibrations from retained endogenous proteins. VISTA, owing its label-free imaging feature, bypasses photobleaching and inhomogeneous labeling arising from fluorophore labeling with much higher sample processing throughput. VISTA also for the first time enabled sub-100 nm (down to 78 nm) label-free imaging with no change on SRS instrumental setup[22], making itself readily applicable. With correlative VISTA and immunofluorescence images, we trained an established machine-learning image-segmentation algorithm[24,25] that generates protein-specific multiplex images from single-channel VISTA images. We further applied VISTA to investigate Aβ plaques in mouse brain tissues, providing a holistic image suited for sub-phenotyping based on the fine views of the plaque core and peripheral filaments surrounded by cell nuclei and blood vessels.

4.3 Detailed protocol

1. Preparation of stock solutions for fixation and sample expansion

1.1. Prepare 40 ml of fixation solution by first dissolving 12g of acryl amide (30% w/v) solid in 26 ml of nuclease-free water. Then add 10 ml of 16% PFA stock solution in the mixture. Finally, add 4 ml of 10x phosphate-buffered saline (PBS, pH 7.4). Made solution can be stored in 4 ˚C for 2 weeks.

Note: Acrylamide is hazardous so the step of dissolving acrylamide solid in water should be handled in fume hood.

1.2. Prepare of gelation solution (stock X) by dissolving 70 mg of sodium acrylate (7% w/v), 200 mg of acryl amide (20% w/v), 50 µl N,N′-methylenebisacrylamide (0.1% w/v) in 420 µl ultrapure water, add 57 µl sterile-filtered 10x PBS (pH 7.4) in the end . Solution can be stored in -20 ˚C up to 1 week.

Note: Avoid sodium acrylate solid that forms clumps, make sure sodium acrylate used is dispersive powders. Stock X made will be a colorless liquid; if the liquid looks light yellow, obtain a new sodium acrylate source.

1.3. Prepare the polymerization initiator solution by dissolving 1 g of ammonium persulfate (APS, 10% w/w) in 9 ml nuclease-free water. Similarly, prepare polymerization accelerator solution by dissolving 1 g of tetramethylethylenediamine (TEMED, 10% w/w) in 9 ml nuclease-free water. The resulting solutions should be aliquot into small portions and stored in -20 ˚C.

1.4. Denaturing buffer (50 ml) was prepared by dissolving 2.88 g of sodium dodecyl sulfate (SDS, 200 mM), 0.584 g of sodium chloride (200 mM), and 2.5 ml of 1M Tris-HCl buffer (50 mM, pH 8) in nuclease-free water.

Note: The concentration of SDS is close to its saturation condition. Some crystallization in the buffer is normal. Slight warm up can make the solution clear and ready to use.

2. Preparation of mammalian cell samples

2.1 Cultured HeLa cells were first seeded on to borosilicate coverslip (#1.5) and were cultured in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum and 1% penicillin–streptomycin antibiotics (complete DMEM) under humidified atmosphere with 5% CO_2 at 37 °C.

2.2 To obtain cells in different mitotic stage, incubate the cells in DMEM without fetal bovine serum for 20 h to synchronize the cell stages once they reached 60% confluency. Switch the medium for cells to complete DMEM and incubate for another 2-6 hour to target various stages of mitosis. Wash the HeLa cells on coverslips with sterile PBS and incubated them with fixation solution (4% PFA 30% AA in PBS) at 37 °C for 7-8 h before further processing.

2.3 For cells with polyQ aggregates: Cells were first grown in complete DMEM until they reached 70–90% confluence. After changing to new complete DMEM, 1 μg plasmids encoding mHtt97Q-GFP was transfected into the cells using transfection agent (details in Table of materials). 24–28 h post-transfection, the coverslips were harvested by fixing the cells in fixation solution (4% PFA, 30% AA) at 37 °C for 6–8 h. The resulting cell samples were ready for further processing

3. Preparation of mouse brain samples

3.1 The normal mice and Alzheimer mice were purchased from commercial sources (details in Table of materials). Rodent euthanasia via carbon dioxide narcosis was performed according to standard protocol[26]. In brief, place mice in a chamber and fill the chamber with a flow of 100% CO_2 in the order of 30–70% of the volume of the chamber per minute and maintain the flow for at least one minute after clinical death. After confirming the humane euthanasia, post-mortem tissue collection of the brain was performed by trained veterinary technical staff.

3.2 First, wash the fresh mouse brain with ice cold PBS and transfer the brain into fixation buffer (4% PFA, 30% AA). Incubate the mouse brain tin fixation buffer first at 4 °C for 24-48h and then at 37 °C for overnight. After washing it with sterile PBS, cut the mouse brain into 150 μm sections using the vibratome and stored in PBS at 4 °C.

4. Hydrogel embedding, denaturation, and expansion of cell and tissue samples

4.1 Make sure the samples were incubated with fixation buffer for proper amount of time, as indicated above. Thaw Stock X, free-radical initiator, and accelerator stock solutions and keep them on ice during the whole process.

4.2 Pick up the coverslip with cells and place it on a glass slide by a tweezer. The brain slices (150 μm thick) can be picked up and placed onto glass slide (make sure they are flat) by a soft-wool paint brush. Get rid of excess liquid that come with samples. Stack 2 coverslips (# 1.5) on both sides of the sample on the glass slide to make a gelation chamber. Place the glass slide with sample facing upward onto an ice-cold heat block and cool it to 4 °C for 1 min.

4.3 First, add 10% (w/w) TEMED to Stock X and then 10% APS solution. Quickly vortex and slowly drop the resulting solution into the chamber onto samples to cover up the surface of the sample, avoid air bubbles. Once the sample is fully immersed by the solution, place a flat parafilm-covered coverslip on top of sample as a lid for the chamber. Leave the chamber on ice or ice-cold heat block for another 1 min before incubating it in a humidify incubator at 37 °C for 30 min.

Note: Sample hydrogel embedding must be done within 1 min after mixing stock X and APS. The whole process should be kept on ice (or ice-cold heat block) as much as possible to avoid gel formation (high degree of polymerization) before adding to the sample.

4.4 Take out the gelation chamber on the glass slide from the 37 °C incubator. A non-transparent gel should be observed after removing the lid. Retrieve the gel by cutting with razorblade or alternatively put the glass slide into denaturing buffer at room temperature for 15 min. The gel will separate itself from the glass slide. Incubate the isolated gel in denaturing buffer under heat condition to achieve further denaturation and delipidation. For cell samples, denaturing at 95 °C for 1 hour would be sufficient. For 150 µm thick brains slices, denaturing needs to be done at 70 °C for 3 h and 95 °C for 1 h.

Note: The gel could become curly when first incubated with denaturing buffer at room temperature. It is normal and will become flat after high temperature denaturation. When working with thicker tissue samples, longer denaturation time is needed.

4.5 After the heat treatment, wash the denatured sample 3 times with PBS for 10 min. At this point, the gel should have expanded to about 1.5 times of the original size. Incubate the washed gel with ultrapure water in a large container (at least 20 times the volume of the gel) to achieve higher expansion ratio. Change the water every 1 hour 3 times and leave the sample in darkness overnight. The resulting gel is ready to image.

Note: The expansion ratio largely depends on the mechanical properties of the original sample. We obtained 4.2 times expansion for our cell samples and 3.4 times expansion for brain tissue sample.

5. Label-free imaging of endogenous protein distribution in expanded cell and tissue samples

5.1 Keep the expanded gel samples in H_2O throughout entire imaging process. Place the expanded sample-bearing gel onto a microscope slide (1 mm thick) with a microscope spacer filled with water. Cover the spacer with a coverslip (#1.5) and ensure its properly sealed to avoid sample movement. The spacer with appropriate opening sizes and depths was used to hold the sample hydrogel hybrid and water.

Note: The gel is fragile after expansion. It needs to be handled with caution.

5.2 Place the sample with either the expanded cell or tissue onto the motorized stage with the coverslip facing the objective. Use the bright field to find z and adjust the condenser to proper position for Köhler illumination. Tune the pump wavelength to 791.3 nm on laser panel control to target protein CH_3 vibration at 2940 cm^{-1}. Switch the light of the microscope from bright-field (eye piece) to laser scanning mode and open the SRS laser shutter. Find the proper focus by changing the z while looking at the real-time SRS image under short pixel-dwell time (12.5 µs/pixel) and low image resolution (256x256).

5.3 Once the ideal z-position is found, perform SRS imaging at higher resolution (1024x1024) with a longer pixel-dwell time (80 µs/pixel) that matches the time constant of the lock-in amplifier. Acquire volumetric images by collecting a z-stack with a step-size of 1 µm in z-direction. Process and analyze the saved OIR file from the Olympus software by ImageJ.

Note: In our SRS setup, a tunable picosecond laser with an 80-MHz repetition rate provides the pump (770–990 nm) and Stokes (1031.2 nm) excitation lasers into a laser scanning confocal microscope (detailed in Table of materials). Temporal and spatial overlapping of the two beams were optimized for signals with pure D_2O. It takes some efforts to find the right z for the expanded sample under bright field because refractive index is very homogenous throughout the gel.

5.4 We determined the resolution of VISTA by taking an SRS image of polystyrene beads (100 nm) using both a 25x objective and a 60x objective. The pump laser was set to 784.5 nm, corresponding to Raman shift of 3050 cm^{-1}, characteristic of the aromatic C-H stretch vibrations of polystyrene. The pump laser wavelength used here is also close to what we use in VISTA for proteins. With the 60x objective, the experimental FWHM of the bead image was 276.17 nm. We modeled the function of the bead object as a half circle; when the PSF Gaussian function has a c (σ) = 269 nm/2.35, the convoluted bead image would have a measured FWHM of 276.17 nm. As a result, the resolution of our SRS system is 269 nm * 1.22 = 328 nm by the Rayleigh Criterion. As VISTA has an average of 4.2 times expansion on cell samples, the effective resolution of VISTA is down to 328 nm / 4.2 = 78 nm.

6. Correlative VISTA and Fluorescent imaging of immuno-labeled and expanded tissue samples

6.1 After denaturation, pre-incubate the hydrogel embedded samples (e.g., 150 µm brain coronal section) in 1% (v/v) Triton X-100 (PBST) for 15min. Then switch the incubation buffer to PBST with primary antibody at 1:100 dilution. If multiple protein targets are needed for multiplex imaging, dilute respective primary antibodies for different targets to proper concentrations in the same cocktail and incubated with the samples simultaneously. Incubate the gels with diluted primary antibodies at 37 °C with gentle shaking (RPM 80) for 16-18 h, followed by extensive washing (3 times) with PBST for 1-2 h at 37 °C.

6.2 Incubate the samples with secondary antibodies of corresponding species targets at 1:100 dilution with PBST at 37 °C for 12-16 h, protected from the light. Wash labeled hydrogel samples 3 times by PBST for 1-2 h at 37 °C with gentle shaking. Expand the immuno-labeled gel samples by incubating with a large volume of double deionized H_2O. Change the water every 1 h 3 times and incubate samples in H_2O overnight, protected from light.

Note: During labeling, the gel should be in 1.5 times expansion as it is in PBS buffer. Flipping over the gel during the incubation would help prevent inhomogeneous antibody labeling. Double-check antibody cross-reactivities to avoid crosstalk between fluorescence channels. Cell nuclei are stained by DAPI.

6.3 Prepare the imaging sample as described section 5.1. Place the imaging sample onto our laser scanning confocal microscope with the 25x, 1.05 NA, water immersion objective for fluorescent imaging. Change the laser (405-, 488-, 561-, and 640-nm) and PMT pair to the proper wavelength according to the targeted antibodies. Adjust focus position, laser power, image acquisition time, and PMT gain according to real-time fluorescence signal to avoid dim signal or over saturation

6.4 Acquire correlative SRS and fluorescent images by first performing volumetric fluorescence imaging on immuno-labeled samples. Switch the microscope light path to IR transparent condition. Open the shutter for SRS laser and perform SRS volumetric imaging on the field of view of same sample with the same range of z.

Note: There is slight chromatic aberration that causes shift in z position between the fluorescent images and the SRS images because of the wavelength difference in the excitation lasers. Manual side-by-side comparison between the SRS and fluorescence z-stack images is needed. We seek the exact matching features from both SRS and fluorescence channel to match the z position precisely.

7. Construction, training, and validation of U-Net architecture

Note: Installing on Linux is recommended. An graphics card with >10GB of ram is required.

7.1 Setting up environment

- 7.1.1. Install Anaconda or Miniconda (we use 3-5.3.0-linux-x86_64).
- 7.1.2. Clone or download https://github.com/Li-En-Good/VISTA.
- 7.1.3. Create a Conda environment for the platform.

  Command line: conda env create -f environment.yml

7.2. Training the prediction model

- 7.2.1. Pair the directories of corresponding SRS images with ground truth images in a csv file. Place the directories of SRS images under path_signal and ground truth images under path_target columns.

- 7.2.2. Put the csv file in the folder data/csvs

- 7.2.3. Modify the configuration in scripts/train_model_2d.sh if needed

- 7.2.4. Activate the environment

Command line: `conda activate fnet`

7.2.5. Initiate model training
Command line: `./scripts/train_model_2d.sh <file name of the csv file> 0`
The training will then start. The losses for each iteration will be shown in the command line and saved with the model in the folder save_models/<file name for the csv file>.

7.3. Validate the images in the training set and test set.
Command line: `./scripts/train_model_2d.sh <file name of the csv file> 0`
The prediction results will be saved in the folder results/<file name of the csv file>.

8. VISTA combined with U-Net predictions enables protein specific multiplexity in label-free images.

8.1 Modify the csv file in data/csvs/<file name of the csv file>/test.csv. Replace the directories of both path_signal and path_target with the directory of new SRS images.

8.2 Remove the prediction results from the training, which is the folder results/<file name of the csv file>.

8.3 Run predictions
Command line: `./scripts/train_model_2d.sh <file name of the csv file> 0`
The prediction results will be saved in the folder results/<file name of the csv file>/test.

4.4 Result

After establishing the working principle of VISTA, we employed image registration to evaluate the expansion ratio of VISTA and to ensure isotropic expansion during the sample processing (**Fig. 1a, b**). We imaged both untreated and VISTA samples targeting bond vibration at 2940 cm^{-1}, which originates from CH_3 of endogenous proteins. In untreated sample, the protein rich structures like nuclei are dark due to the overwhelming lipid content from surrounding tissues[22] (**Fig. 1a**). After VISTA processing that includes the delipidation treatment, the resulting image showed exactly the same feature with an inversed contrast (**Fig. 1b**). The shapes and the relative positions of nuclei and vessels are completely unaltered (**Fig. 1a**, **b**; numbered structures), confirming VISTA treatment is an isotropic process. By comparing the sizes of the corresponding nuclei, we concluded that VISTA achieves 3.4 times expansion in brain tissue samples[22,23].

Knowing the expansion ratio in brain tissue, we can now resolve new features in our label-free SRS images that are previously unresolvable. First, we showed that we could capture features from healthy mouse cortexes down to 150 nm (**Fig. 1c**, **d**). Based on the dispersive patterns around neuronal dendrites, the observed small structures are likely dendritic spine heads[7], which have a size of 146 nm (**Fig. 1d**). In addition, we applied VISTA to image the fibrillar structures in Aβ plaques which are believed to have thickness around 100 nm[27,28]. Indeed, we demonstrate that we could resolve ~ 130 nm fibrillar structures in a representative diffusive Aβ plaque by VISTA (**Fig. 1e**, **f**).

As VISTA enables effective protein retention and protein imaging[22], we can clearly distinguish protein rich nucleoli in nuclei and the ribbon-like cytoskeleton structures in the cytosols of cultured HeLa cells (**Fig. 2a**, arrowheaded). We further applied VISTA to study poly-glutamine (polyQ) aggregates that are transiently expressed in mammalian cell (**Fig. 2b**, **c**). We confirmed that the aggregates, as an expectedly densely packed structure, were expanded isotropically by comparing the same aggregate structures before and after expansion across multiple replicate samples[23]. We then indeed obtained high-resolution structures that are absent/blurred in the normal-resolution SRS images. Our VISTA-aggregates images revealed fibril-like protrusions on the peripheral of the polyQ aggregates and with a hollow structure in the center (**Fig. 2b**, arrowheaded). The observation that protrusions seamlessly attach to cytosolic contents might suggest that aggregates engage functional proteins in cytosol. With hindsight, the capacity to expand dense aggregates also becomes plausible because fixation reagents formaldehyde and hydrogel monomers acryl amide, sodium acrylates are all small molecules that can diffuse in and out of protein aggregates. Once the aggregate was co-polymerized with monomers into hydrogel, the expansion process should proceed as normal.

We then applied VISTA to mouse brain tissue to further extend the scope of our method. Although tissue samples pose challenges like reduced permeability, increased thickness, and heterogeneous mechanical strength, we successfully imaged a mouse brain sample with VISTA (**Fig. 2d**). Similar to cell samples, protein rich structures including cell nuclei, blood vessels, and neuronal processes are observed (**Fig. 2d**, arrowheaded). The limitation of brain tissues is that we only achieved 3.4-times expansion which makes the effective the resolution of VISTA in brain samples to be 99 nm[22]. We validated the structural origin of SRS signals by correlative dye and antibody staining, in which DAPI stains for nuclei and lectin stains for blood vessels (**Fig. 2e**, **f**). Neuronal cell bodies and processes were also delineated by immunofluorescence from NeuN and MAP2[22]. With the trained convolutional neural network (CNN) algorithm utilizing the correlative fluorescence images as ground truth, the single-channel VISTA images were then segmented into specific protein-structure channel for multiplex VISTA images[22].

Finally, we aim to interrogate pathologic Aβ plaques in brains of 5xFAD mice, a well-known animal model for Alzheimer's disease[29]. After standard VISTA procedures, we acquired a 3-dimensional SRS image of amyloid plaques deposition in brain tissues (**Fig. 2g**). Puncta with high protein concentration was observed (**Fig. 2g**, orange arrowed), representing the core of the Aβ plaque. Such image also reveals peripheral Aβ plaques (**Fig. 2g**, magenta arrowheaded), which is often neglected by conventional congo-red staining that only targets the Aβ core. When combined with the trained segmentation algorithm, the label-free VISTA image could be transformed into target-specific multiplex image (**Fig. 2f**) and can be performed in joint with immunofluorescence[23] to study plaque-astrocyte and plaque-microglia microenvironment interactions[30], in a comprehensive and high-throughput manner.

4.5 Discussion

In summary, we present the protocol for VISTA, which is a label-free modality to image protein-rich cellular and subcellular structures of cells and tissues. By targeting endogenous CH_3 from proteins in hydrogel embedded cell and tissues, VISTA achieves an effective imaging resolution down to 78 nm in biological samples and resolves minor extrusion in Huntingtin aggregates and fibrils in Aβ plaques. VISTA is the first instance to report sub-100 nm resolution for label-free imaging modalities[22]. Compared to existing expansion methods[6-8,31-33], VISTA inherits the merit of label-free SRS imaging and hence is free from photobleaching, inactivation, or quenching caused by laser-illuminations. In addition, as a label-free method, VISTA circumvents the demanding, inefficient, and potentially artifact-causing antibody-labeling that is always involved in methods such as DISCO[12,34] and ExM[32,33] and thus offers high throughput sample preparation and uniform imaging throughout tissues. To address the lack of multiplexity in the label-free approach, VISTA, implemented with CNN-based image segmentation algorithm[25], provides protein-specific multi-component images without any labels in brain tissues[22]. We further applied VISTA on 5xFAD mouse brains and enabled a holistic volumetric view of aggregates core and periphery fibrils, nuclei and blood vessels[23]. We envision that VISTA would scale up well for larger samples such as primate or human brain slices and could ultimately be useful for clinical investigations.

There are three essential steps that ensure the successful implementation of VISTA method. First, maximum protein retention in hydrogel sample hybrid is crucial and required for VISTA[22]. To achieve this goal, we modified our fixation condition to contain high concentration of acrylamide[34] and replaced the protein digestion procedure with high-concentration detergent delipidation that saves significant protein loss from protein digestion. The addition of AA quenches intermolecular crosslinking of proteins and enables the isotropic expansion without protein digestions[34]. Second, proper correlations between SRS and immuno-labeling and distinctions between different protein targets need to be established. As VISTA relies on image-segmentation algorithms to add multiplexity to monochromatic SRS images, crosstalk between different protein targets in immunofluorescence will significantly compromise the quality of VISTA images. We meticulously selected protein rich structures that are obvious in SRS images and validated their corresponding immunofluorescence features. Third, before using the model to predict fluorescence patterns from new SRS data sets, the validity and reliability of trained ML model should be testified. Distinct features that are not included in the training sets will likely cause issues in prediction. If the prediction results are not satisfying, the user should try to include more data for training and avoid predicting patterns that are not included in the training sets. Pearson's correlations of the testing sets and validation sets should also be monitored to ensure the quality of the prediction[22,23]. We suggest the user to have at least 100 corresponding image sets for training.

While VISTA has immense potential for biological studies, there are certain limitations awaiting creative solutions. First and foremost, the sensitivity of VISTA needs further improvement. The detection limit of label-free simulated Raman scattering is in low millimolar range and the isotropic expansion of samples in three dimensions significantly dilutes chemical bonds and weakens VISTA signal. We hence are limited to imaging the total ensemble of endogenous proteins that lacks specificity and multiplexity. Combining VISTA with ultrasensitive SRS[35], we could possibly extend VISTA to image low abundance proteins and study aggregate structures and compositions at super-resolution level by targeting orthogonal chemical bonds[36]. Second, the current 3.4 times expansion ratio in brain tissues only gives moderate resolution improvement. Although we have already resolved minor extrusions in Aβ plaques that are previously indistinguishable, higher resolution is always desirable. In this case, innovations in protein-anchoring and hydrogel chemistry would greatly benefit VISTA. For example, different gel formulation could enable larger expansion ratios for even higher image resolution[37-38,40]. New procedures in sample processing would allow VISTA to be applied with widely available FFPE histology samples[37,39], making VISTA even better suited for large-scale clinical studies.

4.6 Figures

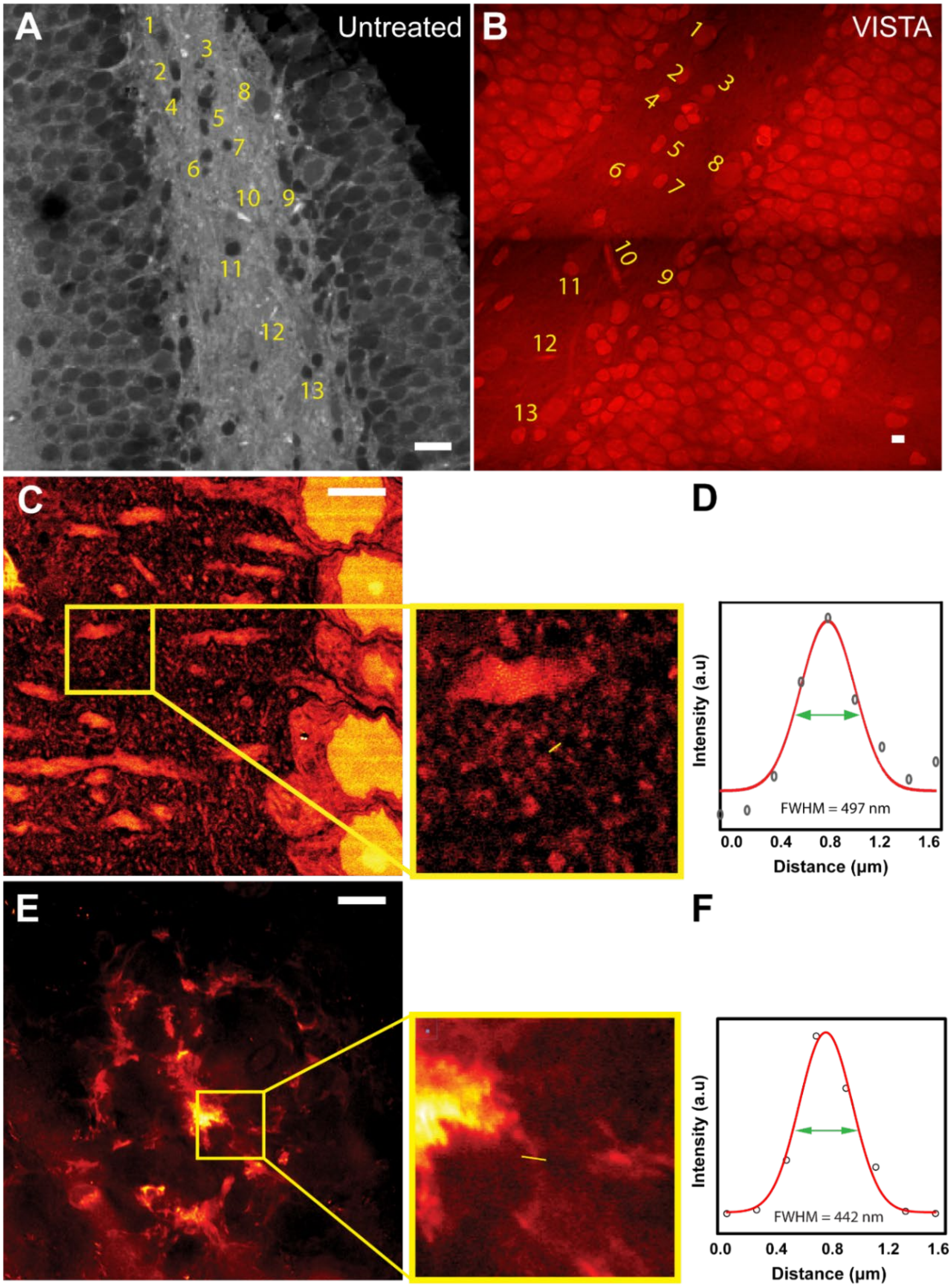

A
Untreated
1
2
3
4
5
6
7
8
9
10
11
12
13
B
VISTA
1
2
3
4
5
6
7
8
9
10
11
12
13
C
D
Intensity (a.u)
FWHM = 497 nm
0.0
0.4
0.8
1.2
1.6
Distance (µm)
E
F
Intensity (a.u)
FWHM = 442 nm
0.0
0.4
0.8
1.2
1.6
Distance (µm)

Fig. 1: Sample expansion strategy enables super-resolution label-free imaging in mouse brain tissues. a) An SRS image at CH_3 frequency of mouse hippocampus. b) A VISTA image at the same field of view of mouse hippocampus. Labeled area show the corresponding features before and after VISTA treatments. c). A VISTA image of a normal mouse cortex that shows finer features. Inset shows region of interest. d) Resolution quantification for the fine structure observed in expanded samples. FWHM of 497 nm corresponds to effective resolution of 146 nm with 3.4 times expansion. e) A VISTA image of an amyloid beta plaque in mouse brain tissues. Inset shows enlarged region of interest. f) Resolution quantification for the extrusion fiber structure of the expanded amyloid beta plaque. FWHM of 442 nm corresponds to effective resolution of 130 nm with 3.4 times expansion. Scale bars: 20 μm.

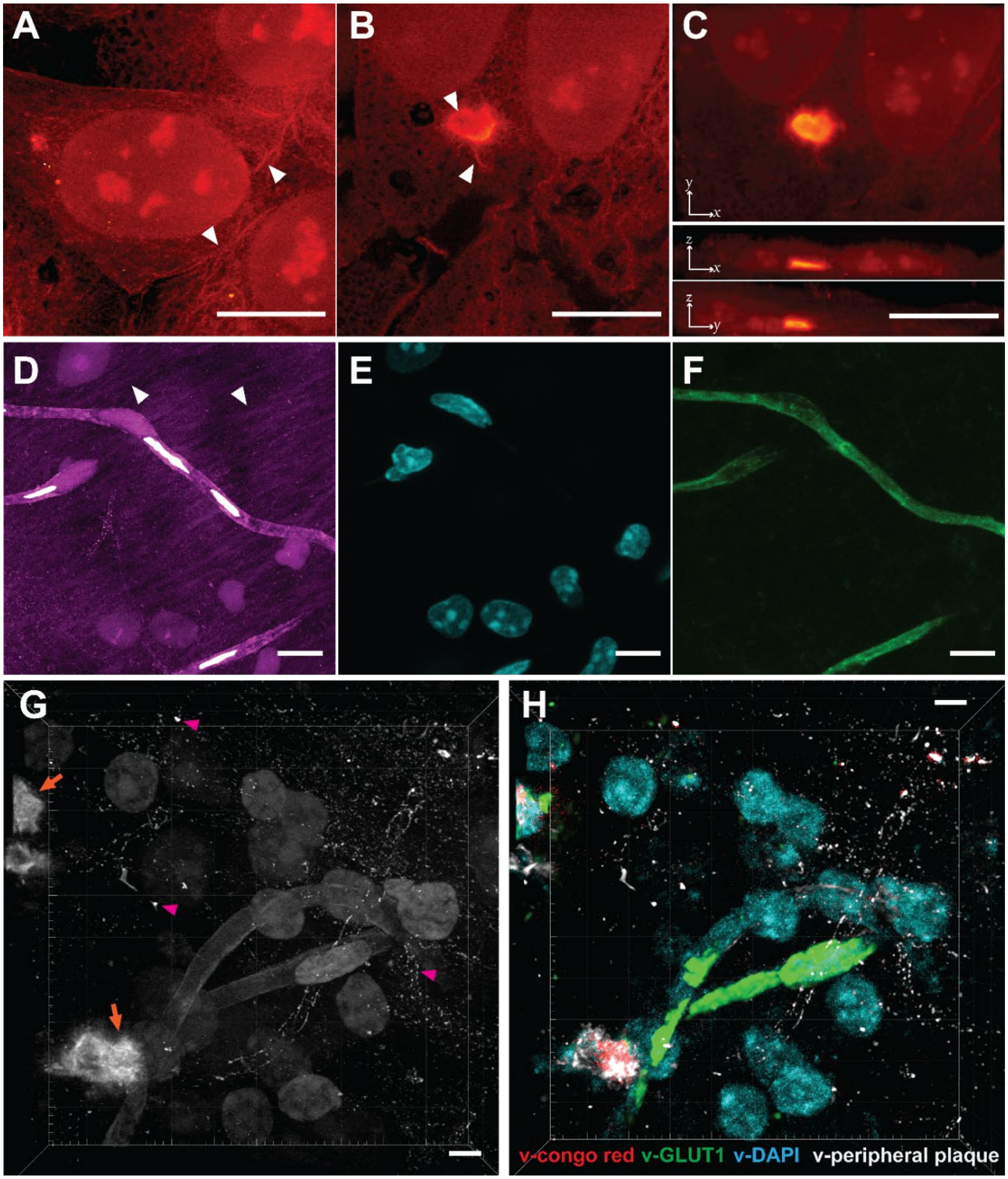


Fig. 2: Label-free super-resolution volumetric imaging in cells and tissues enabled by VISTA. a) Volumetric VISTA image of a normal HeLa cell. Arrowheaded: cytoskeleton like structure. b) Single z-slice VISTA image of a polyQ aggregate expressed in HeLa cells. Arrowheaded: hollow structure and fibril extrusions. c) Maximum intensity projections of x-

y, x-z, y-z directions show the volumetric view of the polyQ aggregate containing cell. d) Volumetric VISTA image of a coronal section of mouse brain. Arrowheaded: neuronal processes. e) Fluorescence image of nuclei (DAPI staining) at the same sample region shows 1 to 1 correlation with nuclei in VISTA image. f) Fluorescence image of blood vessels (anti-lectin) at the same sample region shows 1 to 1 correlation with vessel structures in VISTA image. g) Volumetric VISTA image of Aβ (orange arrowed) containing brain tissue. Magenta arrowheaded: peripheral Aβ plaque. h) Multiplex VISTA image from g), predicted by trained image segmentation algorithm. v-Congo red represents core of Aβ plaque. v-GLUT1 represents blood vessels. v-DAPI represents nuclei. v-peripheral plaque represents the Aβ plaque that won't be stained by Congo red dye. Scale bars: 10 μm. The length scale is in terms of distance before expansion (adjusted for different expansion ratios).

4.7 Reference

1. Ntziachristos, V. Going deeper than microscopy: the optical imaging frontier in biology. *Nat. Methods* **7** (8), 603–614 (2010)
2. Lavis, L. D.; Raines, R. T. Bright ideas for chemical biology. *ACS Chem. Biol.* **3** (3), 142–155 (2008).
3. Tsien, R. Y. Constructing and exploiting the fluorescent protein paintbox (Nobel Lecture). *Angew. Chem. Int. Ed.* ***48*** (31), 5612–5626 (2009).
4. Huang, B.; Bates, M.; Zhuang, X. Super-resolution fluorescence microscopy. *Annu. Rev. Biochem.* ***78***, 993–1016 (2009).
5. Sahl, S. J.; Hell, S. W.; Jakobs, S. Fluorescence nanoscopy in cell biology. *Nat. Rev. Mol. Cell Biol.* ***18*** (11), 685–701 (2017).
6. Wassie, A. T.; Zhao, Y.; Boyden, E. S. Expansion microscopy: principles and uses in biological research. *Nat. Methods* ***16*** (1), 33–41 (2019).
7. Chen, F.; Tillberg, P. W.; Boyden, E. S. Expansion microscopy. *Science* **347** (6221), 543–548 (2015).
8. Gambarotto, D.; Zwettler, F. U.; Le Guennec, M.; Schmidt-Cernohorska, M.; Fortun, D.; Borgers, S.; Heine, J.; Schloetel, J.-G.; Reuss, M.; Unser, M.; Boyden, E. S.; Sauer, M.; Hamel, V.; Guichard, P. Imaging cellular ultrastructures using expansion microscopy (U-ExM). *Nat. Methods* **16** (1), 71–74 (2019).
9. Demchenko, A. P. Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection. *Methods Appl. Fluoresc.* **8** (2), 022001 (2020).
10. Murray, E.; Cho, J. H.; Goodwin, D.; Ku, T.; Swaney, J.; Kim, S.-Y.; Choi, H.; Park, Y.-G.; Park, J.-Y.; Hubbert, A.; McCue, M.; Vassallo, S.; Bakh, N.; Frosch, M. P.; Wedeen, V. J.; Seung, H. S.; Chung, K. Simple, scalable proteomic imaging for high-dimensional profiling of intact systems. *Cell* **163** (6), 1500–1514 (2015).
11. Kim, S.-Y.; Cho, J. H.; Murray, E.; Bakh, N.; Choi, H.; Ohn, K.; Ruelas, L.; Hubbert, A.; McCue, M.; Vassallo, S. L.; Keller, P. J.; Chung, K. Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proc. Natl. Acad. Sci. U.S.A.* **112** (46), E6274–E6283 (2015).
12. Liebmann, T.; Renier, N.; Bettayeb, K.; Greengard, P.; Tessier-Lavigne, M.; Flajolet, M. Three-dimensional study of Alzheimer’s disease hallmarks using the IDISCO clearing method. *Cell Rep.* 16 (4), 1138–1152 (2016).
13. Min, W.; Freudiger, C. W.; Lu, S.; Xie, X. S. Coherent nonlinear optical imaging: beyond fluorescence microscopy. *Annu. Rev. Phys. Chem.* **62** (1), 507–530 (2011).
14. Saar, B. G.; Freudiger, C. W.; Reichman, J.; Stanley, C. M.; Holtom, G. R.; Xie, X. S. Video-rate molecular imaging in vivo with stimulated Raman scattering. *Science* **330** (6009), 1368–1370 (2010).
15. Cheng, J.-X.; Xie, X. S. Vibrational spectroscopic imaging of living systems: an emerging platform for biology and medicine. *Science* **350** (6264), 1054–1063 (2015).

16. Ji, M.; Orringer, D. A.; Freudiger, C. W.; Ramkissoon, S.; Liu, X.; Lau, D.; Golby, A. J.; Norton, I.; Hayashi, M.; Agar, N. Y. R.; Young, G. S.; Spino, C.; Santagata, S.; Camelo-Piragua, S.; Ligon, K. L.; Sagher, O.; Xie, X. S. Rapid, Label-free detection of brain tumors with stimulated Raman scattering Microscopy. *Sci. Transl. Med.* **5** (201), 201ra119-201ra119 (2013).
17. Wei, M.; Shi, L.; Shen, Y.; Zhao, Z.; Guzman, A.; Kaufman, L. J.; Wei, L.; Min, W. Volumetric chemical imaging by clearing-enhanced stimulated Raman scattering Microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116** (14), 6608–6617 (2019).
18. Ji, M. *et al.* Label-free imaging of amyloid plaques in Alzheimer's disease with stimulated Raman scattering microscopy. *Sci. Adv.* **4**, eaat7715 (2018).
19. Silva, W. R.; Graefe, C. T.; Frontiera, R. R. Toward label-free super-resolution microscopy. *ACS Photonics* **3** (1), 79–86 (2016).
20. Gong, L.; Zheng, W.; Ma, Y.; Huang, Z. Higher-order coherent anti-stokes Raman scattering microscopy realizes label-free super-resolution vibrational imaging. *Nat. Photonics* **14** (2), 115–122 (2020).
21. Watanabe, K.; Palonpon, A. F.; Smith, N. I.; Chiu, L.; Kasai, A.; Hashimoto, H.; Kawata, S.; Fujita, K. Structured line illumination Raman microscopy. *Nat. Comm.* **6** (1), 10095 (2015).
22. Qian, C.; Miao, K.; Lin, L.-E.; Chen, X.; Du, J.; Wei, L. Super-resolution label-free volumetric vibrational imaging. *Nat. Comm.* **12** (1), 3648 (2021).
23. Lin, L.-E.; Miao, K.; Qian, C.; Wei, L. High spatial-resolution imaging of label-free in vivo protein aggregates by VISTA. *Analyst* **146** (13), 4135–4145 (2021).
24. Ounkomol, C.; Seshamani, S.; Maleckar, M. M.; Collman, F.; Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15** (11), 917–920 (2018).
25. Falk, T.; Mai, D.; Bensch, R.; Çiçek, Ö.; Abdulkadir, A.; Marrakchi, Y.; Böhm, A.; Deubner, J.; Jäckel, Z.; Seiwald, K.; Dovzhenko, A.; Tietz, O.; Dal Bosco, C.; Walsh, S.; Saltukoglu, D.; Tay, T. L.; Prinz, M.; Palme, K.; Simons, M.; Diester, I.; Brox, T.; Ronneberger, O. U-Net: Deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16** (1), 67–70 (2019).
26. Yang, B.; Treweek, J. B.; Kulkarni, R. P.; Deverman, B. E.; Chen, C.-K.; Lubeck, E.; Shah, S.; Cai, L.; Gradinaru, V. Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158** (4), 945–958 (2014).
27. Mlodzianoski, M. J.; Cheng-Hathaway, P. J.; Bemiller, S. M.; McCray, T. J.; Liu, S.; Miller, D. A.; Lamb, B. T.; Landreth, G. E.; Huang, F. Active PSF shaping and adaptive optics enable volumetric localization microscopy through brain sections. *Nat. Methods* **15** (8), 583–586 (2018).
28. Querol-Vilaseca, M.; Colom-Cadena, M.; Pegueroles, J.; Nuñez-Llaves, R.; Luque-Cabecerans, J.; Muñoz-Llahuna, L.; Andilla, J.; Belbin, O.; Spires-Jones, T. L.; Gelpi,

E.; Clarimon, J.; Loza-Alvarez, P.; Fortea, J.; Lleó, A. Nanoscale structure of Amyloid-β plaques in Alzheimer's disease. *Sci. Rep.* **9** (1), 5181 (2019).

29. Oakley, H.; Cole, S. L.; Logan, S.; Maus, E.; Shao, P.; Craft, J.; Guillozet-Bongaarts, A.; Ohno, M.; Disterhoft, J.; Van Eldik, L.; Berry, R.; Vassar, R. Intraneuronal beta-Amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in Amyloid plaque formation. *J. Neurosci.* **26** (40), 10129–10140 (2006).
30. T. Bartels, S. D. Schepper and S. Hong. Microglia modulate neurodegeneration in Alzheimer's and Parkinson's diseases. *Science*. **370**, 66 —69 (2020).
31. Chen, F.; Wassie, A. T.; Cote, A. J.; Sinha, A.; Alon, S.; Asano, S.; Daugharthy, E. R.; Chang, J.-B.; Marblestone, A.; Church, G. M.; Raj, A.; Boyden, E. S. Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* **13** (8), 679–684 (2016).
32. Tillberg, P. W.; Chen, F.; Piatkevich, K. D.; Zhao, Y.; Yu, C.-C. (Jay); English, B. P.; Gao, L.; Martorell, A.; Suk, H.-J.; Yoshida, F.; DeGennaro, E. M.; Roossien, D. H.; Gong, G.; Seneviratne, U.; Tannenbaum, S. R.; Desimone, R.; Cai, D.; Boyden, E. S. Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34** (9), 987–992 (2016).
33. Ku, T.; Swaney, J.; Park, J.-Y.; Albanese, A.; Murray, E.; Cho, J. H.; Park, Y.-G.; Mangena, V.; Chen, J.; Chung, K. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* **34** (9), 973–981 (2016).
34. Renier, N.; Wu, Z.; Simon, D. J.; Yang, J.; Ariel, P.; Tessier-Lavigne, M. IDISCO: A simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* **159** (4), 896–910 (2014).
35. Zhuge, M.; Huang, K.-C.; Lee, H. J.; Jiang, Y.; Tan, Y.; Lin, H.; Dong, P.-T.; Zhao, G.; Matei, D.; Yang, Q.; Cheng, J.-X. Ultrasensitive vibrational imaging of retinoids by visible preresonance stimulated Raman scattering microscopy. *Adv. Sci.* **8** (9), 2003136 (2021).
36. Miao, K.; Wei, L. Live-cell imaging and quantification of polyQ aggregates by stimulated Raman scattering of selective deuterium labeling. *ACS Cent. Sci.* **6** (4), 478–486 (2020).
37. Klimas A, *et al.* Nanoscale imaging of biomolecules using molecule anchorable gel-enabled nanoscale in-situ fluorescence microscopy. *Microsc. Microanal*. **28**, 1568–1569 (2022).
38. Shi, L. et al. Super-resolution vibrational imaging using expansion stimulated Raman scattering microscopy. *Adv. Sci.* **9**, 2200315 (2022).
39. Zhao, Y.; Bucur, O.; Irshad, H.; Chen, F.; Weins, A.; Stancu, A. L.; Oh, E.-Y.; DiStasio, M.; Torous, V.; Glass, B.; Stillman, I. E.; Schnitt, S. J.; Beck, A. H.; Boyden, E. S.

Nanoscale imaging of clinical specimens using pathology-optimized expansion microscopy. *Nat. Biotechnol.* **35** (8), 757–764 (2017).

40. M'Saad, O.; Kasula, R.; Kondratiuk, I.; Kidd, P.; Falahati, H.; Gentile, J. E.; Niescier, R. F.; Watters, K.; Sterner, R. C.; Lee, S.; Liu, X.; Camilli, P. D.; Rothman, J. E.; Koleske, A. J.; Biederer, T.; Bewersdorf, J. All-optical visualization of specific molecules in the ultrastructural context of brain tissue. *bioRxiv* April 5, 2022, p 2022.04.04.486901.

Chapter 5

Hydrogen-deuterium exchange on terminal alkynes enables Raman based spectral sensing and imaging heterogenous local environments

The content of this chapter has been published in:
Bi, X.*; Miao, K.*; Wei, L. Alkyne-Tagged Raman Probes for Local Environmental Sensing by Hydrogen–Deuterium Exchange. *J. Am. Chem. Soc.* **2022**, 144 (19), 8504–8514. https://doi.org/10.1021/jacs.2c01991.

5.1 Summary

Alkyne-tagged Raman probes have shown high promise for noninvasive and sensitive visualization of small biomolecules to understand their functional roles in live cells. However, the potential for alkynes to sense cellular environments that goes beyond imaging remains to be further explored. Here, we report a general strategy for Raman imaging-based local environment sensing by hydrogen–deuterium exchange (HDX) of terminal alkynes (termed alkyne-HDX). We first demonstrate, in multiple Raman probes, that deuterations of the alkynyl hydrogens lead to remarkable shifts of alkyne Raman peaks for about 130 cm–1, providing resolvable signals suited for imaging-based analysis with high specificity. Both our analytical derivation and experimental characterizations subsequently establish that HDX kinetics are linearly proportional to both alkyne pKas and environmental pDs. After validating the quantitative nature of this strategy, we apply alkyne-HDX to sensing local chemical and cellular environments. We establish that alkyne-HDX exhibits high sensitivity to various DNA structures and demonstrates the capacity to detect DNA structural changes in situ from UV-induced damage. We further show that this strategy is also applicable to resolve subtle pD variations in live cells. Altogether, our work lays the foundation for utilizing alkyne-HDX strategy to quantitatively sense the local environments for a broad spectrum of applications in complex biological systems.

5.2 Introduction

The emergence of diverse imaging probes, including organic molecules and genetically encodable proteins (e.g., GFP) has significantly advanced the development of modern optical microscopy. For vibrational spectro-microscopy, alkynes have become one of the most popular functional Raman imaging probes, especially for non-invasive visualization of small biomolecules in live cells, which presented as a fundamental limit in fluorescence microscopy due to the large size of fluorophores.[1-8] The tiny, exogenous, and biorthogonal alkyne moiety exhibits a distinct, sharp, and strong Raman peak in the cell-silent spectral window (1800–2600 cm^{-1}, **Figure 1a**, **S1a**), boosting Raman detection sensitivity and specificity and providing superb biocompatibility with minimal perturbation. In particular, when coupled with stimulated Raman scattering (SRS) microscopy, alkyne-tagged small precursors or analogs have been applied to interrogating a wide range of biomolecules in living organisms with subcellular resolution, including nucleic acids, proteins, glycans, lipids, small metabolites, natural products, and drug candidates.[9-13]

In addition to imaging, alkyne-tagged Raman probes have found wide utilities in sensing local chemical and cellular environments. For instance, reactions between the probes and various substrates, which lead to shifts in Raman frequency of alkynes due to chemical structure change, have been designed to sense the presence of small molecules,[14] ions,[15-17] or enzymes[18-19]. One major issue here is that a new alkyne-bearing probe needs to be devised and synthesized for each specific chemical transformation to achieve high selectivity of targeted reactions. An alternative sensing strategy is to measure the Raman peak shift produced by the vibrational Stark effect[20-21], which is caused by noncovalent interactions between an alkyne probe and its adjacent environment. This phenomenon allows exploiting diverse existing alkyne probes with proven live-cell compatibility to sense protein motions[22] and conformational changes[23]. However, the reported vibrational frequency shifts are rather limited (~ 5 cm^{-1}) and hence require prolonged data acquisition and fine spectrum fitting to determine the alkyne peak positions. The small peak shift also poses challenges for combing vibrational Stark sensing with the fast imaging-based SRS microscopy, which often has a spectral resolution of around 10-15 cm^{-1} without complex instrument customization.[12] A general strategy that further explores the potential of existing alkyne probes towards quantitative environmental sensing with high spatial and temporal resolution would hence be highly desirable.

Here, we report a general imaging-based platform that utilizes terminal alkyne-tagged Raman probes for local environmental sensing by alkyne hydrogen–deuterium exchange (HDX) (**Figure 1b**). Conventionally, amide HDX has been adopted by NMR[24-25] and mass spectrometry[26-29] to examine protein conformations and micro-environments, through detecting the spectroscopic or mass signatures that correspond to the deuteration at the amide backbone of proteins after exposure to deuterated solvents. The deuteration sites correlate with the accessibility of the amino acid residues at various positions of the protein secondary or tertiary structures. Similar spectroscopic characterizations were demonstrated by IR[30-34] and resonance Raman spectroscopy[35-37], yet with limited resolvability due to peak crowdedness in the fingerprint spectral region. Inspired by these studies, we here explore the Raman spectral shift and the kinetics of deuterium exchange on alkynyl hydrogen of terminal alkynes in the background-free cell-silent region. We reasoned that HDX should occur for terminal alkynes in suitable deuterium sources (e.g., D_2O, **Figure 1b**), because the known pK_as of alkynes (~ 20-25) are much smaller than those of sp^3 CH bonds (~ 40-60) and are close to those of amides (~ 20)[38-39]. Therefore, by tracking alkyne-HDX dynamics with corresponding Raman signatures, we could investigate micro-environments of diverse alkyne-tagged probes in heterogeneous biological systems. We first theoretically establish and experimentally validate the generality of the proposed alkyne-HDX strategy. We then demonstrate unique imaging-based sensing applications of local chemical and cellular environments, including investigation of UV-induced DNA damage in cells and visualization of subtle variations in intracellular pHs. Our approach provides a novel dimension to existing sensing strategies and offers new opportunities to investigate complex biological microenvironments in their native states.

5.3 Results

Raman peak shift upon alkyne-HDX. Using established alkyne-tagged cell-imaging probes, the thymidine analog 5-ethynyl-2'-deoxyuridine (EdU) and choline analog propargylcholine (PCho) (**Figure 1a**), we first examined the achievable alkyne Raman peak shift before and after HDX in D_2O. After EdU and PCho were diluted and equilibrated in the D_2O solutions, spontaneous Raman spectra showed a distinct red-shift over 100 cm^{-1} for the alkyne stretching frequency. 130 cm^{-1} shift was observed for EdU from 2124 cm^{-1} to 1994 cm^{-1} and 136 cm^{-1} shift was observed for PCho from 2142 cm^{-1} to 2006 cm^{-1} (**Figure 1c**). This extended peak shift is a result of the combined effect of adding more mass to the alkyne "oscillator" (similar to introducing heavier-mass ^{13}C labeling[40]) and the coupling between the C≡C and adjoining C-D bonds. Such significant shifts are consistent with previous reports of deuterated acetylide compounds[41-42]. We also performed DFT calculations on EdU after alkyne deuteration, in which the peak shift was predicted to be 149 cm^{-1}, close to our experimental data. Similar peak shift values were detected for other alkyne-tagged molecules, including the cytidine analog ethynyl-2'-deoxycytidine (EdC, 128 cm^{-1} shift) and the uridine analog 5-ethynyl-uridine (EU, 134 cm^{-1} shift) (**Figure S1a**, b). In addition to the peak shift, we observed a slight decrease in peak intensities comparing fully converted alkyne-D peaks with the original alkyne-H ones. (**Figure 1c**, S1b, dashed lines). The ratios of the peak intensity after and before alkyne-HDX are 0.77, 0.46, 0.86, 0.79 for EdU, PCho, EdC and EU, respectively. As relative peak intensities decrease, the corresponding full-width-at-half-maximums (FWHMs) increase slightly, maintaining approximately identical integrated peak areas between alkyne peaks from before and after alkyne-HDX (integrated peak-area ratios: EdU, 0.94; PCho, 0.81; EdC, 0.96; EU, 0.93). These spectroscopic characterizations validate our first hypothesis that the alkyne-HDX strategy offers a clear spectral separation for robust peak distinction and analysis across diverse live-cell compatible alkyne probes.

Theoretical analysis and experimental validation of alkyne-HDX kinetics. We next sought to establish the theoretical basis for alkyne-HDX processes. The analytical framework of amide-HDX in proteins has been founded as acid-, base- or water-catalyzed.[43-45] We modeled the alkyne-HDX process as a base-catalyzed reaction in D_2O buffer solutions, as the physiological pD (pD=-lg[D^+]) is higher than 7. The rate-determining step (RDS) in HDX between alkyne-tagged probes (R-C≡C-H) and catalytic base OD^- is shown in Eq. (1) below (more details in SI). It includes three elementary steps: a) diffusional collision, b) equilibrium redistribution of the hydrogen in the intermediate state, and c) dissociation.

$$R-C\equiv C-H+OD^- \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} R-C\equiv C-H \ldots O-D^- \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} R-C\equiv C^- \ldots H-O-D \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} R-C\equiv C^- + HOD \qquad (1)$$

We can assume that: 1) the temperature remains constant; 2) the reaction is diffusion-limited; 3) the intermediates are in steady states.[43] The overall rate (k) for transferring a proton from R-C≡C-H to OD^- then becomes (details in Supporting Information):

$$k = k_1 \cdot (10^{pK_a(\,donor,\ R-C\equiv C-H)-pK_a(acceptor,\ HOD)} + 1)^{-1} \cdot [OD^-] \cdot [R-C\equiv C-H]. \qquad (2)$$

k_1 is the diffusion-limited collision constant, upper-bounded by 10^{10} $M^{-1}s^{-1}$.[43] As fluctuation of OD^- concentration is negligible during HDX, the RDS is considered as a pseudo-first-order reaction. Since the acceptor pK_a is much smaller than the donor pK_a, the corresponding exchange half-life ($t_{1/2}$) is reduced to Eq. (3):

$$\lg(t_{1/2}) = pK_a - pD + constant. \qquad (3)$$

Here pK_a designates the donor pK_a, i.e., $pK_a(R-C\equiv C-H)$. Taking the pK_a of alkynyl hydrogen as 20-25 and the pK_a of HOD as ~15,[46] we estimated the $t_{1/2}$ of alkyne-HDX in physiological pD (7.6) to be on the order of minutes, reasonable kinetics to be captured by both spontaneous Raman spectroscopy and live-cell SRS microscopy.

With the above theoretical derivation, we next acquired time series of spontaneous Raman spectra for the four alkyne imaging probes (i.e., EdU, PCho, EU and EdC), to experimentally examine the sensitivity of HDX kinetics on alkynes with minor changes of their surrounding chemical structures. By diluting each stock solution into a pD=7.6 D_2O buffer (DPBS-D_2O buffer solution), a gradual decrease of alkyne-H peak intensity (I(R-C≡C-H)) and an increase of alkyne-D peak intensity (I(R-C≡C-D)) were observed as time elapsed (**Figure 2a, c** and Figure **S2a, c**). For each peak, a Gaussian peak fitting was implemented to generate a precise peak height as the intensity. Complying with pseudo-first-order reactions, the decrease of I(R-C≡C-H) indeed follows an exponential decay. The obtained average $t_{1/2,\ \mathrm{EdU}}$ is 103.9±5.3 min (**Figure 2b**, n=3), consistent with our above theoretical estimations. Interestingly, the average exchange $t_{1/2,\ \mathrm{PCho}}$ is 8.4±0.9 min (**Figure 2d**, n=3), only about one-tenth of the $t_{1/2}$ for EdU, indicating that the pK_a of the terminal alkyne in PCho is smaller than that in EdU. For the other two alkyne-tagged nucleoside probes, EU and EdC[47], which have similar structures and close alkyne Raman peaks to those of EdU, while $t_{1/2,\ \mathrm{EU}}$ is close to $t_{1/2,\ \mathrm{EdU}}$, which is expected due to the shared uracil structure (**Figure S2a, b**), $t_{1/2,\ \mathrm{EdC}}$ is about 5 times shorter than $t_{1/2,\ \mathrm{EdU}}$, although EdC is only the amination product on the beta carbon to the alkyne of EdU (**Figure S2c, d**). These kinetics results demonstrate that our alkyne-HDX strategy is highly sensitive to resolve the subtle changes of local chemical structures, which is reflected in the varying alkyne pK_a values.

Establishing a general rule between alkyne-HDX and the probe structures. Guided by Eq. (3), it would be highly desirable to experimentally establish a quantitative relationship between the pK_as of alkynes and their exchange kinetics to inform the future structural design for alkyne-HDX probes. Even though powerful pK_a prediction methodologies have been developed, precise calculation of the pK_a for a given alkynyl hydrogen remains challenging, limited by the solvent leveling effect.[39] Hence, we adopted the Hammett substituent constant σ (**Figure 3a**), defined in Eq. (4) below,[48] as a surrogate for pK_a, to explore the relationship between the pK_a values and the HDX kinetics.

$$\sigma_X = \lg K_X - \lg K_H \qquad (4)$$

K_H is the ionization constant for benzoic acid in water at 25 °C and K_X is the corresponding constant for a para-substituted benzoic acid. A positive (negative) σ value represents an electron-withdrawing (donating) group (**Figure 3a**), and more positive σ values were reported to correlate with lower pK_as. The same correlation should also hold for alkynyl hydrogen in alkyne probes.[48]

We chose five phenyl-acetylene derivatives with various σ as our model molecules (4-fluorophenylacetylene, methy-4-ethynylbenzoate, 4-ethynylbenzaldehyde, 1-ethynyl-4-nitrobenzene, 1-ethynyl-4-methoxybenzene). Due to limited solubility of substituted phenylacetylene molecules in aqueous solutions, we measured their alkyne-HDX kinetics in two different deuterium conditions of 1:1 DMSO-D_2O solutions (pD=7.6) and methanol-OD. We observed a nice linear relationship between $lg(t_{1/2})$ and σ for the model phenylacetylenes in both deuterium solutions (**Figure 3b**). With the lowest σ value, 1-ethynyl-4-methoxybenzene with an electron-donating group (-OMe) indeed shows the slowest kinetics. Similarly, with the highest σ value, 1-ethynyl-4-nitrobenzene with the strongest electron-withdrawing group (-NO_2) has the fastest kinetics. The observed linear relationship between $lg(t_{1/2})$ and σ together with Eq. (3) confirms that σ is a quantitative surrogate for alkyne pK_as. These data hence offer a general interpretation for the dependence of alkyne-HDX kinetics on chemical structure changes and would help guide the design of alkyne probes with specific alkyne-HDX kinetics.

Experimentally confirming a linear relationship between alkyne-HDX and pD. In addition to pK_a, another factor that directly influences HDX rate is the pD in exchange solutions, as shown in Eq. (3). To experimentally establish the relationship between pD and HDX kinetics, we measured the $t_{1/2}$ on all above alkyne probes dissolved in D_2O buffers with varying pD values. Consistently, $lg(t_{1/2})$ displays a nice linear relationship with pDs for all four probes (**Figure 3c, d**). It is worth pointing out that the theoretical $lg(t_{1/2})$ - pD slope is -1 in the simplified Eq. (3). While our data for EdC and PCho meet this criterion well, the slopes for EdU and EU (e.g., negative 0.5) deviate from the prediction. Similar mismatch in slopes has been shown in previous work on amide-HDX in proteins.[49-50] Further efforts to understand other factors that lead to the deviation are still ongoing.

Fast and homogenous D_2O diffusion throughout the whole cells. After establishing the quantitative nature of the alkyne-HDX approach, we now sought to apply it for cellular sensing. One potential issue moving from *in vitro* homogenous solution measurements to heterogeneous cells is the cellular uptake and intracellular diffusion of D_2O, which could influence the alkyne-HDX kinetics. Previous studies by magnetic field gradient NMR have reported the outer membrane water permeability as ~0.001 cm/s and the intracellular diffusion coefficient as $\sim 10^{-6}$ cm^2/s.[51-52] Taking the typical cell size (~10 μm) into consideration, we calculated that cells should experience a full D_2O environment within a second of D_2O incubation. Such rapid D_2O diffusion timescale should not influence alkyne-HDX rates in cells.

To experimentally justify that D_2O diffusion would not influence alkyne-HDX, we performed time-lapse Raman spectroscopy and imaging by targeting the O-D stretching peak (~2490 cm^{-1}) as an indicator to trace the D_2O uptake and diffusion in both live and fixed cells after immersion in the DPBS-D_2O buffer. Our spontaneous Raman spectra confirmed that D_2O reaches cell nuclei (evidenced by the hump at ~ 2490 cm^{-1}) and achieves equilibrium (> 98% compared to the D_2O solution signal outside the cells) within tens of seconds of incubation in live cells (**Figure 4a**). Additionally, the signal from O-H stretching peak (~3430cm^{-1}) is nearly negligible, indicating that there is almost no residual H_2O in cells (**Figure 4a**), consistent with the fast water diffusion. To exclude the potential issue of heterogeneous intracellular distribution of D_2O, we also performed time-lapse SRS imaging at 2490 cm^{-1}. Images taken after 3 min and 65 min of incubation are almost identical (**Figure 4b, c**), indicating that the intracellular D_2O level reaches equilibrium within 3 min, similar to what is shown in spontaneous Raman (**Figure 4a**). The D_2O distribution across cells is mostly homogenous with an observed heterogeneity less than 5%, consistent with the normal distribution of H_2O recorded by O-H stretching peak.[53] We note that the dark spots shown in

Figure 4b-c are hydrophobic lipid droplets. Similar spontaneous Raman and SRS measurements for diffusions were also captured in fixed cells (**Figure S3a-c**). All above results confirm that D_2O diffusion is a fast and homogeneous process that would not interfere with the measurements of the alkyne-HDX kinetics in cells.

Sensing the changes of DNA structures by alkyne-HDX. The ability to analyze DNA structural changes is essential yet challenging. Majority of existing methodologies rely on solution characterizations *in vitro*, which may cause perturbations to the original structures during extraction, purification, and reconstitution processes[54-55]. Understanding DNA structures in their native environments should gain new insights for scenarios such as DNA damage, genome instability, and DNA-protein interactions. Toward this goal, EdU has been proven as a well-established live-cell compatible SRS reporter for DNA where alkynes on EdU presumably extends into the major grooves of double helix[56] . We hypothesize that the changes in DNA structures would modulate the local microenvironments of alkynes and this process could be reflected and monitored by the alkyne-HDX of EdU. To verify the hypothesis, we acquired the HDX kinetics in cellular DNA by collecting a time series of SRS images for alkyne-HDX at the alkyne-H (2224 cm^{-1}) and the alkyne-D (1992 cm^{-1}) frequencies on EdU-labeled cells immersed in the DPBS-D_2O buffer (Figure S4). Surprisingly, the alkyne-HDX kinetics for EdU in cells are significantly (~ 10 times) slower (**Figure S4**, $t_{1/2} = 16.5$ hours) compared to that for EdU in the DPBS-D_2O buffer solution.

To better understand this slow-down effect in cells, we evaluated the alkyne-HDX for potential EdU-labeled downstream products (Figure 4d) *in vitro* in the same DPBS-D_2O buffer (**Figure 4e**, orange bars). We first tested the HDX kinetics of EdUTP (EdU triphosphate), because EdU needs to be phosphorylated before being incorporated into genomic DNA and the phosphate groups bring extra negative charges to the vicinity of the alkynes. The $t_{1/2}$ for EdUTP is 105.5 min that is virtually the same as EdU (**Figure 4e, Figure S5c**). Hence, we ruled out the possible contribution of added phosphate groups to the slower alkyne-HDX kinetics. We then hypothesized that the formation of polynucleotides where

EdU was compressed into conformationally restricted polymer chains leads to the slowdown. Consequently, we generated EdU-labeled double-stranded DNA (dsDNA, 746 bp) and single-stranded DNA (ssDNA, 746 nt) (**Figure 4d**) by replacing dTTP with EdUTP in PCR mixture and using the EdUTP-tolerating pwo DNA polymerase.[57] The amplification of ssDNA was based on asymmetric PCR, in which concentrations of forward and reverse primers are different.[58] The successful formation of dsDNA and ssDNA was confirmed through gel electrophoresis and ssDNA-specific Exonuclease I (ExoI) digestion[58] (**Figure S5f, g**). The purified EdU-labeled ssDNA and dsDNA were then used for alkyne-HDX kinetics measurements (Figure 4e, $t_{1/2}$(EdU-dsDNA) = 36 hours and $t_{1/2}$(EdU-ssDNA) = 58 hours). We also attempted to measure the alkyne-HDX kinetics of purified EdU-labeled genomic DNA extracted from cells in solutions, but unfortunately failed due to the limited solubility and lower labeling efficiency of EdU inside cells compared to PCR reactions. These data indicate that the formation of polynucleotide chains plays a major role in the slowdown of HDX kinetics in EdU-incorporated cell.

Interestingly, among all EdU labeled polynucleotides, EdU-ssDNA showed the slowest alkyne-HDX, whereas EdU-labeled cellular DNA has the fastest kinetics (**Figure 4e**, orange bars). This result implies that our alkyne-HDX strategy is indeed sensitive to the change of DNA structures. To gain more insights on the causes of varying alkyne-HDX kinetics across different EdU-labeled DNA structures, we next investigated how the high-salt condition (DPBS-D_2O buffer with 2 M NaCl),[59] which could perturb the DNA electrostatic interactions, modulates the alkyne-HDX of EdU (**Figure 4e**, light blue bars). We confirmed that the high osmolarity wouldn't alter the D_2O diffusion rate across cells by both spontaneous Raman and SRS measurements (**Figure S3d-f**). Interestingly,while the high-salt condition accelerated the HDX for both EdU-cells and EdU-dsDNA (**Figure 4e, S5**), it did not influence the HDX kinetics for both EdUTP and EdU-ssDNA. As a control, alkyne-HDX for the EdU was instead slowed down in the high-salt buffer compared to that in the regular DPBS-D_2O buffer. This means that the accelerating effect of the high-salt condition should be mostly attributed to the DNA structural changes.

Based on above data, one possible explanation for the significant HDX slowed-down of ssDNA could be the formation of intramolecular hydrogen-bonds (H-bonds) around alkynes,[60] which effectively block the alkyne-HDX by shielding. Similar H-bonds formation is known to slow down the HDX.[36] This postulation for forming strong H-bonding is also consistent with the minimum influence on the alkyne-HDX for ssDNA in high-salt condition (Figure 4e). As a comparison, in dsDNA (and also cellular DNA) with well-organized helical structures, the formation of Watson–Crick base pairs pushes alkynes towards major grooves, likely eliminating H-bonding around EdU alkynes. They both hence present faster alkyne-HDX compared to ssDNA (Figure 4e). The high-salt condition is also known to compromise the double helical structure and lead to partial denaturation of dsDNA.[59] This change of structural integrity was then reflected in the alkyne-HDX speed-up for both EdU-labeled dsDNA and cellular DNA. Collectively our results for different DNA species under varied conditions establish the viability of employing alkyne-HDX to sense and study the local DNA environments.

Sensing the changes of cellular DNA structures through UV-induced damage *in situ* by alkyne-HDX. DNA damage inhibits DNA replication, interrupts the normal function of cells and potentially leads to cell death. Studies of structural changes upon DNA damage have been mostly limited to *in vitro* examinations.[54, 61-62] As alkyne-HDX of EdU is a viable approach to monitor DNA double helices, next we aimed to apply our strategy to investigate UV-induced DNA damage in native cellular environments. DNA damage induced by UV irradiation involves the formation of cyclobutene pyrimidine dimers (CPDs), in which the thymine dimer (T-T dimer) is the most common one (Figure 5a).[63-66] As a thymidine analogue, EdU would be a perfect reporter for interrogating the T-T dimer formation. To establish the proper dimerization condition, we first implemented a reported procedure to induce thymine dimerization in solutions by UV irradiation.[62] The reported change in spontaneous Raman spectra was observed (Figure S6a), confirming the dimer formation.[62] Similar UV condition was also suited for inducing T-T dimer formation in live cells, which was confirmed by immunostaining (anti-T dimer, Figure S6b, c). With the established condition, we next examined the formation of T-T dimer in EdU-labeled cellular DNA by alkyne-HDX (**Figure 5b**). Interestingly, we observed an about 30% faster alkyne-HDX kinetics for UV-cells (**Figure 5c**, $t_{1/2}$(UV-cell) = 11.6 h vs $t_{1/2}$(cell, with no UV) = 16.5 h). We speculated that the increase of HDX kinetics could originate from two non-mutually exclusive factors: first, the explicit chemical structure change upon EdU-EdU dimer formation, and second, the change in local DNA structures and thus alkyne microenvironments upon formations of T-T (or EdU-EdU or T-EdU) dimers (**Figure 5a**).

To test our hypothesis, we purified the EdU dimer using reverse-phase column chromatography and further characterized the spectral properties and the alkyne-HDX kinetics of EdU dimer. While no obvious Raman peak shift was observed, the SRS alkyne intensity of EdU dimer became only one-quarter of that from EdU. This means the intensity of alkyne in each EdU becomes 1/8 of the original intensity, probably due to the loss of conjugation to alkyne (**Figure 5d**). The alkyne-HDX kinetics of EdU dimer solution increased by almost two folds, indicating a decreasing alkyne-pK_a upon dimer formation (**Figure 5c**, EdU vs EdU dimer). With the observed 30% faster HDX kinetics in cells, we would expect a decent percentage of dimer formation that would result in a significant drop in SRS signals. This contradicts our observed SRS images for UV-irradiated cells whose intensities remain approximately unchanged (**Figure 5b**). Furthermore, the limited EdU-labeling efficiency in live cells should also restrict the probability of the EdU dimer formation, which requires two EdU residing in proximity. Collectively, our data largely ruled out that the direct EdU dimer formation is the primary reason for the increased alkyne-HDX kinetics. We proposed that the accelerating HDX kinetics instead senses the disruption of double-helical structures caused by UV-induced T-T dimerization (**Figure 5a**). This is also consistent with previous reports that the formation of T-T dimers would unwind the DNA double helix about 10° and lead to a bend of about 30°.[61] In this scenario, the EdU essentially serves as a reporter for T-T dimer formation: the EdU could experience a compromised double-helical structure due to the formation of CPDs and shows faster alkyne-HDX kinetics

(**Figure 5c**). The high salt treatment and UV-irradiation result in similar EdU-HDX in cellular DNA, suggesting both dimer formation and the high-salt condition might have similar levels of disruptions in DNA double-helical structures.

Sensing minor pD variations in live cells by alkyne-HDX. In addition to sensing chemical environments, our alkyne-HDX is also highly sensitive to the environmental pDs, as shown above with the logarithmic dependence of $t_{1/2}$ on pD. We therefore implemented alkyne-HDX to potentially resolve subtle pD changes in the cellular environment. Such detectability would find wide utility, particularly in differentiating the local and heterogeneous acidity in acidic diseases including cancer, inflammation and others.[67] To achieve such a goal, we first calibrated the intracellular pDs with a fluorescent ratiometric pH sensor BCECF (2',7'-is-(2-Carboxyethyl)-5-(and-6)-Carboxyfluorescein) in both D_2O buffer solutions (**Figure S7a**) and live cells immersed in these buffers with varying pD values (**Figure 6a**, left, control). The relative constant ratios (F_{488}/F_{445}) for cells indicate a cellular buffering capacity to the external pD changes (**Figure 6a**, control, left). We further adopted the treatment of nigericin, an ionophore that is utilized to modulate intracellular pH by changing the efflux of K^+/H^+. It could hence help overcome the cellular buffer capacity and modulates the cytosolic pH by extracellular buffer pH (**Figure 6a**, right, nigericin; **Figure S7b-e**).[68-69]

We next captured alkyne-HDX live-cell ratiometric images (I(R-C≡C-D)/I(R-C≡C-H), i.e., SRS_D/SRS_H) for PCho-labeled cells in D_2O buffers without (control) and with nigericin (**Figure. 6b-d**) at the same exchange time point. The same buffering capacity is shown with SRS ratios (**Figure 6b**, left, control). The increase in SRS_D/SRS_H ratio indicates faster alkyne-HDX. The trend of SRS_D/SRS_H is also consistent with the nigericin-balanced intracellular pD values (**Figure 6b**, left, nigericin). The fold change on our SRS ratios on pD changes is much higher than that from fluorescence ratios in BCECF (**Figure 6a-b**), demonstrating our ability to readout small cellular pD differences in live cells. Ideally, we would acquire exponential decay kinetics for PCho alkyne-HDX in cells and derive a $t_{1/2}$. Instead, we discovered interesting exchange kinetics of a biphasic trend with a fast (close to what is shown for PCho solution) and a much slower-exchanging component. One possible explanation for the observed 2-stage HDX is that PCho experiences two different environments as the headgroup of phospholipids: phospholipids exposed to an aqueous environment exhibit a faster exchange rate and those buried in a hydrophobic environment exhibit a slower rate with almost no exchange in the detect time window (~2 hours).

"Two-color" mapping of DNA and lipids in live cells via distinct kinetics of the subcellular alkyne-HDX maps. As EdU and PCho have vastly distinct alkyne-HDX kinetics in cells, we expected to differentiate EdU and PCho in ratiometric (SRS_D/SRS_H) SRS images. We first labeled DNA and lipids in HeLa cells with EdU and PCho and then incubated cells with custom-made DMEM-D_2O. We chose to take SRS images at the overlapping frequency channels 2137 cm^{-1} (SRS_H) and 1998 cm^{-1} (SRS_D) (**Figure 1c**), so that both targets could be mapped at the same channels, and we could avoid acquiring four images, which takes longer time and may introduce complications from cell movement. As EdU-DNA now goes through HDX about 100 times slower than PCho in cells, pixels from EdU-enriched DNA in the nuclear regions should show much smaller ratios than those from PCho-labeled lipid membranes. In ratiometric maps at 7 min, 11 min, and 33 min of D_2O incubation, nuclei indeed always have a smaller ratio compared to cytoplasm (**Figure S8**). Our live-cell ratiometric imaging hence provides an additional dimension to perform "two-color" based on the differentiation of alkyne-exchange kinetics, but not from frequency difference, which might be potentially useful to extract multi-component information in cells for alkyne-tagged Raman probes with close vibrational frequency, but distinct alkyne-HDX kinetics.

5.4 Discussion

We reported a novel Raman-based imaging platform for sensing local chemical and environmental changes in cells by analyzing HDX kinetics of alkyne-tagged probes. We theoretically investigated the alkyne-HDX mechanism and quantitatively established the relationship for the alkyne-HDX kinetics (i.e., $lg(t_{1/2})$) with probe structures, and local chemical and cellular environments (i.e., pK_a and pD). The drastic peak shift accompanied with alkyne-HDX finds its perfect marriage with SRS microscopy, enabling sub-cellular exchange analysis with high spatial and temporal resolution. The exchange also provides an additional dimension for multiplex imaging for alkyne-tagged probes with large spectral cross-talks, but distinct alkyne-HDX kinetics. These properties render alkyne-HDX a uniquely suited imaging-based strategy for *in situ* sub-cellular sensing applications, including sensing small pD changes and DNA structure changes, which all remained challenging to be interrogated previously. Despite limited methods for detecting thymine dimer lesions, our alkyne-HDX provides a fast and noninvasive approach to understand thymine dimerization *in situ*.

It is worth discussing the use of pD values in our work. pD was reported to be directly measured by pH meter with an offset $pD_{real}=pH_{meter\ reading} + 0.4$.[70] We also extensively evaluated the relationship between our own pH meter reading and the actual pD values by both making the standard DPBS buffer in H_2O and D_2O, and by comparing the pD and pH values using fluorescent pH indicator, BCECF (**Figure S7a**). Our own results reliably established that $pD_{real}=pH_{meter\ reading}-0.25$ (more details in Methods). The small difference of the offsets from our investigation and prior report might arise from different pH meter electrode probes or room temperatures. To avoid discrepancy and confusion, we reported all our pD values above directly from the pH meter reading results.

As an outlook, there are a number of fronts that could be further explored for alkyne-HDX. Imaging-wise, when coupled with spectrally resolvable alkyne-probes, parallel multiplex alkyne-HDX sensing could be implemented to simultaneously investigate the local cellular environment in the vicinity of multiple alkyne-tagged biological species and structures. Chemically, alkyne-HDX should find wider utility with further chemical derivatization of the alkyne-probes., We expect alkyne-HDX could serve as a general strategy for sensing enzymatic reactions in cells with properly designed alkyne probes. Biologically, alkyne-HDX should provide unique insights on compartmentalized pDs in various cellular states with further studies. We envision alkyne-HDX to be further extended to investigate other local heterogeneity, such as phase separation. Our preliminary data (not shown) detected a weak but reproducible EdU signal in the non–membrane-bound sub-compartment in nucleoli, likely originated from ribosomal DNA, with almost no observable exchange kinetics. We suspect this non-exchange behavior is due to that rDNA is located in either fibrillar centers (FCs) or dense fibrillar components (DFCs) of the nucleoli,[71-72] the crowding of which blocks alkyne-HDX of EdU in rDNA in nucleolar regions. We are furthering our investigations along this line. Considering all the unique properties of our strategy, we believe our alkyne-HDX could extend the utility of alkynes from imaging to sensing and would present a valuable addition to the optical-imaging-based sensing toolbox in understanding complex biology.

5.5 Figures

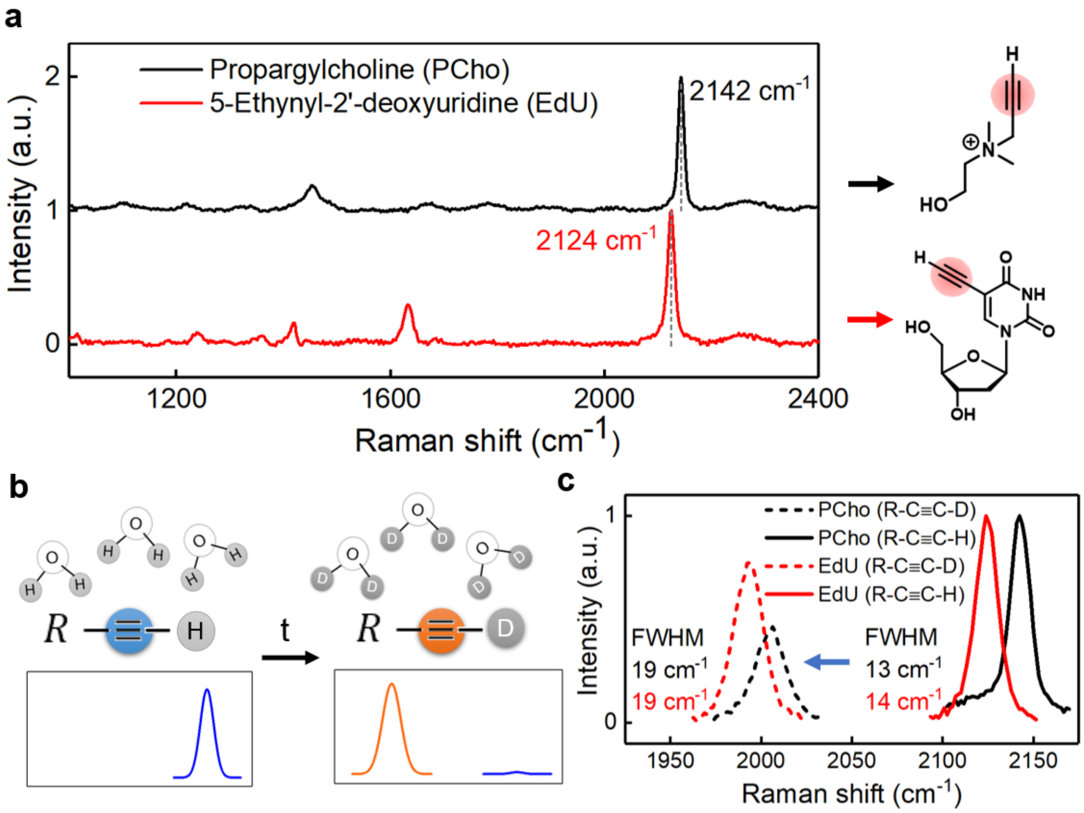


Figure 1. Characterization of hydrogen-deuterium exchange in terminal alkyne groups (alkyne-HDX) from live-cell compatible Raman probes by spontaneous Raman spectroscopy. a) Chemical structures and corresponding spontaneous Raman spectra of EdU (10 mM) and PCho (40 mM) solutions. Terminal alkynes are highlighted in pink. b) Scheme of our designed alkyne-HDX-based Raman sensing. c) Spontaneous Raman peaks of alkynes in EdU (10 mM, red) and PCho (40 mM, black) solutions before (solid lines, normalized) and after (dashed lines) alkyne-HDX.

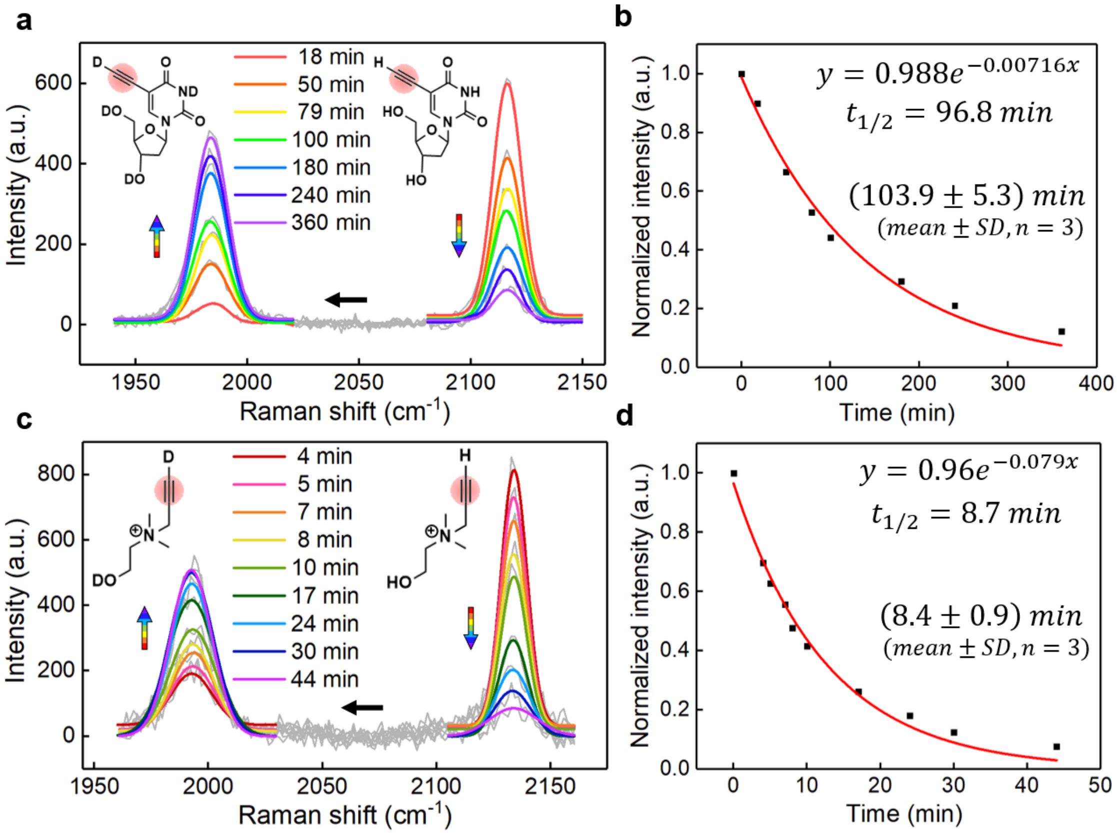


Figure 2. Distinct alkyne-HDX rates for EdU (top) and PCho (bottom) solutions. a) The kinetics trace of the spontaneous Raman spectra for the decrease of alkyne-H peaks and the increase of alkyne-D peaks from EdU (10 mM) during alkyne-HDX in pD=7.6 D_2O buffer solution. b) Exponential curve fitting of normalized alkyne peak intensities for EdU (10 mM) in (a). The average of $t_{1/2}$ over three independent measurements is also shown. c) The kinetics trace of the spontaneous Raman spectra for the decrease of alkyne-H peaks and the increase of alkyne-D peaks from PCho (40 mM) during the alkyne-HDX process in pD=7.6 D_2O buffer solution. d) Exponential curve fitting of normalized alkyne peak intensities for PCho (~40 mM) in (c). The average of $t_{1/2}$ over three independent measurements is also shown.

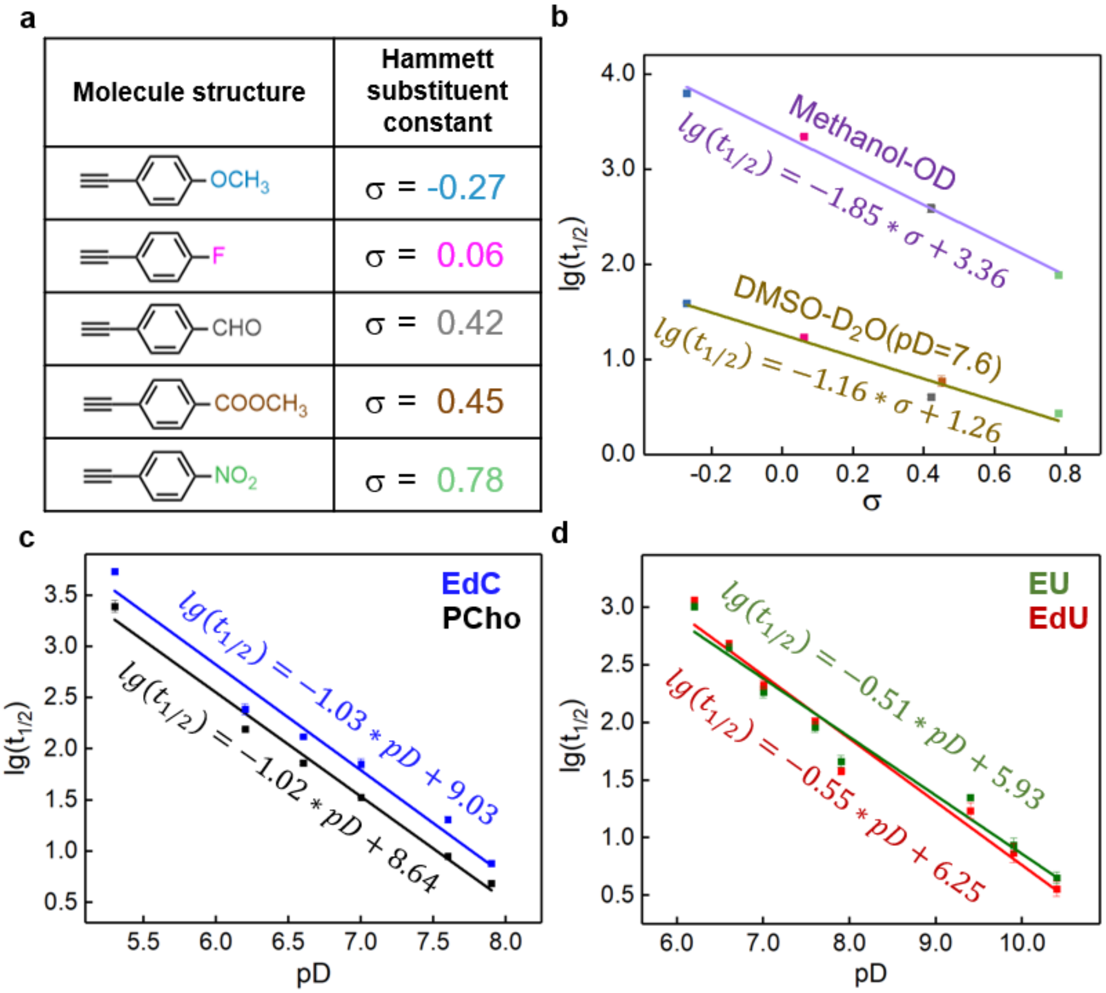


Figure 3. The effect of pK_a and pD on the alkyne-HDX kinetics. a) Structures of model molecules and the corresponding Hammett substituent constant σ. b) Linear relationship between lg($t_{1/2}$) and σ for model molecules in the methanol-OD solution and 1:1 D_2O-DMSO buffer solution with pD=7.6. Corresponding data points for respective molecules shown in (a) are color-coded. c-d) Linear relationship between lg($t_{1/2}$) and pD for EdC (10 mM) and PCho (40 mM) (c) and EdU (10 mM) and EU (10 mM) (d) in D_2O buffer solutions with different pD values. Error bar: SD.

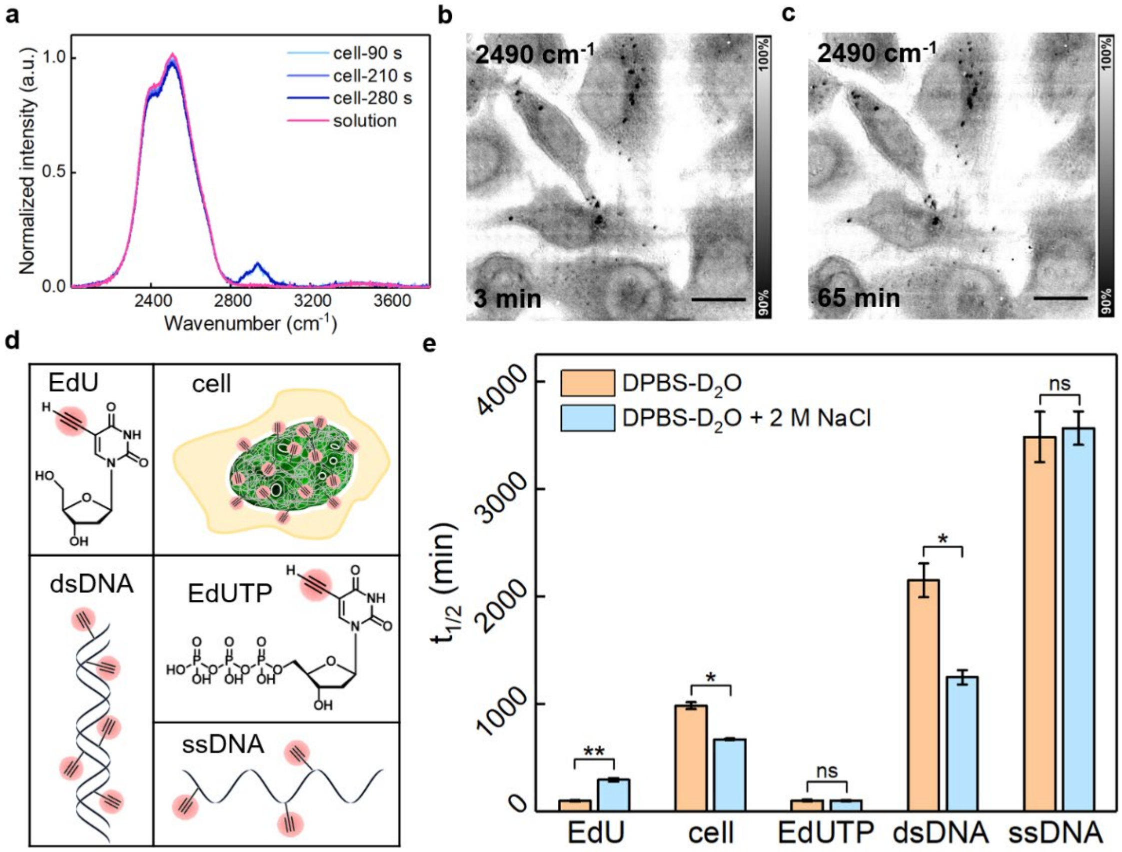


Figure 4. Fast D$_2$O diffusion across cells and the alkyne-HDX-based DNA structure sensing. a) Time-trace spontaneous Raman spectra in the nuclear region of a living cell at the indicated time points after incubation in the DPBS-D$_2$O solution (90 s, 210 s, 280 s). The solution spectrum (magenta) is taken in the surrounding region without cells in the same sample. b-c) SRS image of O-D vibrational peak (2490 cm^{-1}) for live cells in DPBS-D$_2$O buffer solution after 3 min (b) and 65 min (c) incubation. Scale bar: 20 μm. d) Cartoon structures of EdU, EdU-incorporated cells (cell), 5-Ethynyl-dUTP (EdUTP), EdU-labeled double-stranded DNA (dsDNA) and EdU-labeled single-stranded DNA (ssDNA). e) The corresponding t$_{1/2}$ values of alkyne-HDX for EdU, EdU-incorporated cells, EdUTP, EdU-dsDNA, EdU-ssDNA in DPBS-D$_2$O buffer solution (light blue) and in the high salt buffer solution (DPBS-D$_2$O + 2 M NaCl, orange). Error bar: SD. $*p < 0.05$, $**p < 0.01$, ns $p > 0.1$ from unpaired t-tests.

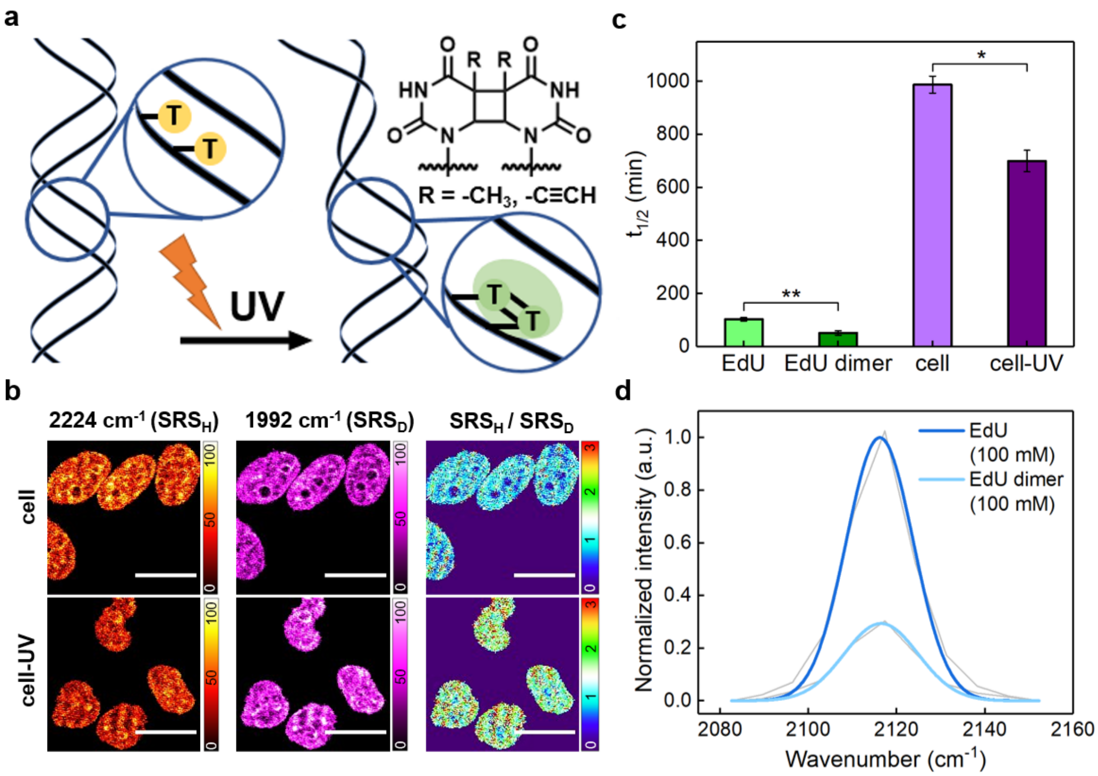


Figure 5. Sensing the UV-induced damage on DNA structures by alkyne-HDX from EdU-labeled cells. a) The dimerization scheme in regular or EdU-labeled DNA induced by UV (254 nm) irradiation. b) Ratiometric imaging of SRS_D / SRS_H (right) generated by dividing the SRS imaging at the alkyne-H channel (2224 cm^{-1}) (left) by that at the alkyne-D channel (1992 cm^{-1}) (middle) for cells without (top, average ratio of 1.11±0.47) and with (bottom, average ratio of 1.61±0.53) UV irradiation immersed in DPBS-D_2O after 23.5h incubation. Experiments were repeated with three independent measurements. Scale bar: 20 μm. c) SRS-imaging measured $t_{1/2}$ of alkyne-HDX from EdU and EdU dimer solutions, EdU-incorporated cells, and EdU-incorporated cells after UV irradiation. Error bar: SD. d) Raman spectra for characterization of alkyne Raman intensity in EdU (100 mM in DMSO) and EdU

dimer (100 mM in DMSO) based on SRS measurements. Error bar: SD. $^{*}p < 0.05$, $^{**}p < 0.01$, ns $p > 0.1$ from unpaired t-tests.

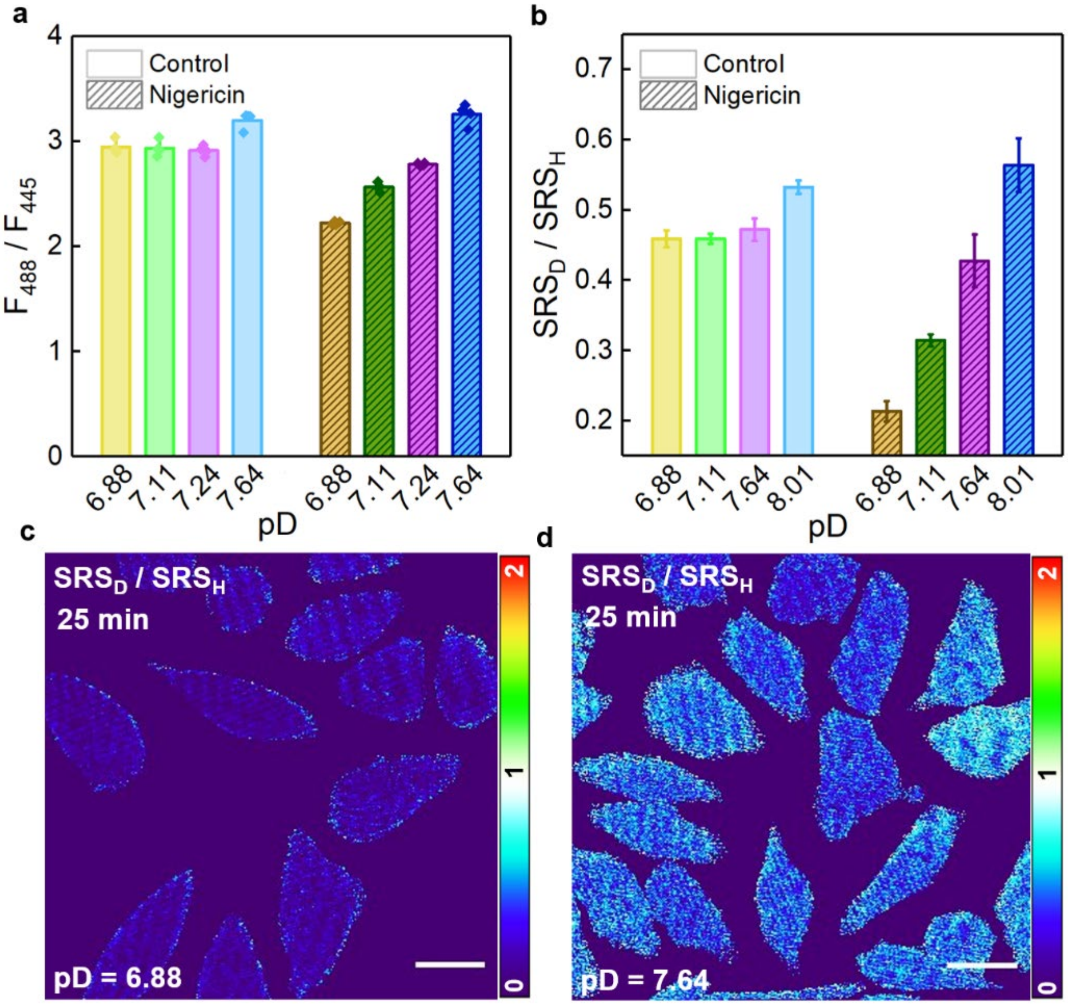


Figure 6. Sensing intracellular pH changes by alkyne-HDX. a) Benchmarking of intracellular pD values by the standard fluorescence ratiometric (F_{488} / F_{445}) pH indicator BCECF from live cells immersed in DPBS-D_2O buffers with minor changes around the physiological pD range with (right) and without (left) extra nigericin for intracellular pD calibration from the extracellular environment. b) The corresponding SRS-imaging ratios (SRS_D / SRS_H) of PCho in live cells immersed in the DPBS-D_2O buffers with similar pD range with (right) and without (left) extra nigericin. Error bar: SD. c-d) Ratiometric imaging

of SRS_D / SRS_H for PCho at 25 min exchange in live cells in pD=6.88 (c) and pD=7.64 (d) DPBS-D_2O buffer solutions. The average ratio for (c) is 0.19±0.13 and for (d) is 0.47±0.21. Scale bar: 20 µm.

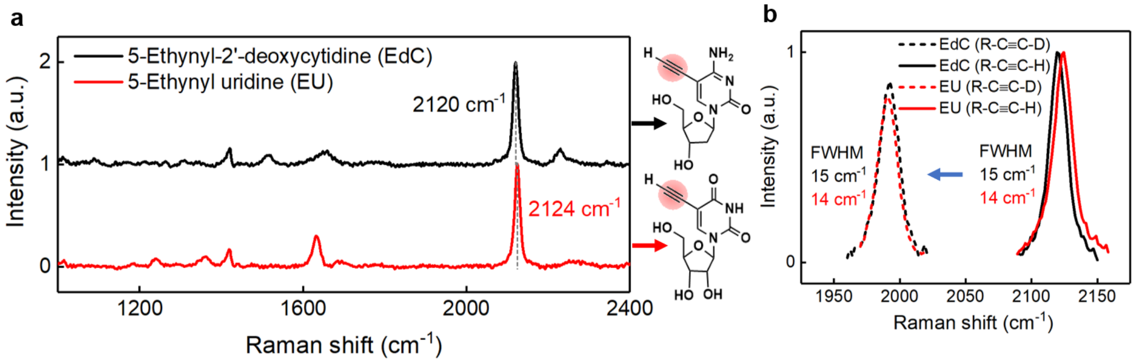


Figure S1 Characterizations of hydrogen-deuterium exchange on terminal alkyne groups (alkyne-HDX) by spontaneous Raman spectroscopy. a) Chemical structures and the corresponding spontaneous-Raman spectra of EdC (10 mM) and EU (10 mM) solutions. Pink-shade highlight the terminal alkynes. b) Spontaneous-Raman peaks of alkyne in EU (10 mM, red) and EdC (10 mM, black) solutions before (solid lines) and after (dashed lines) HDX.

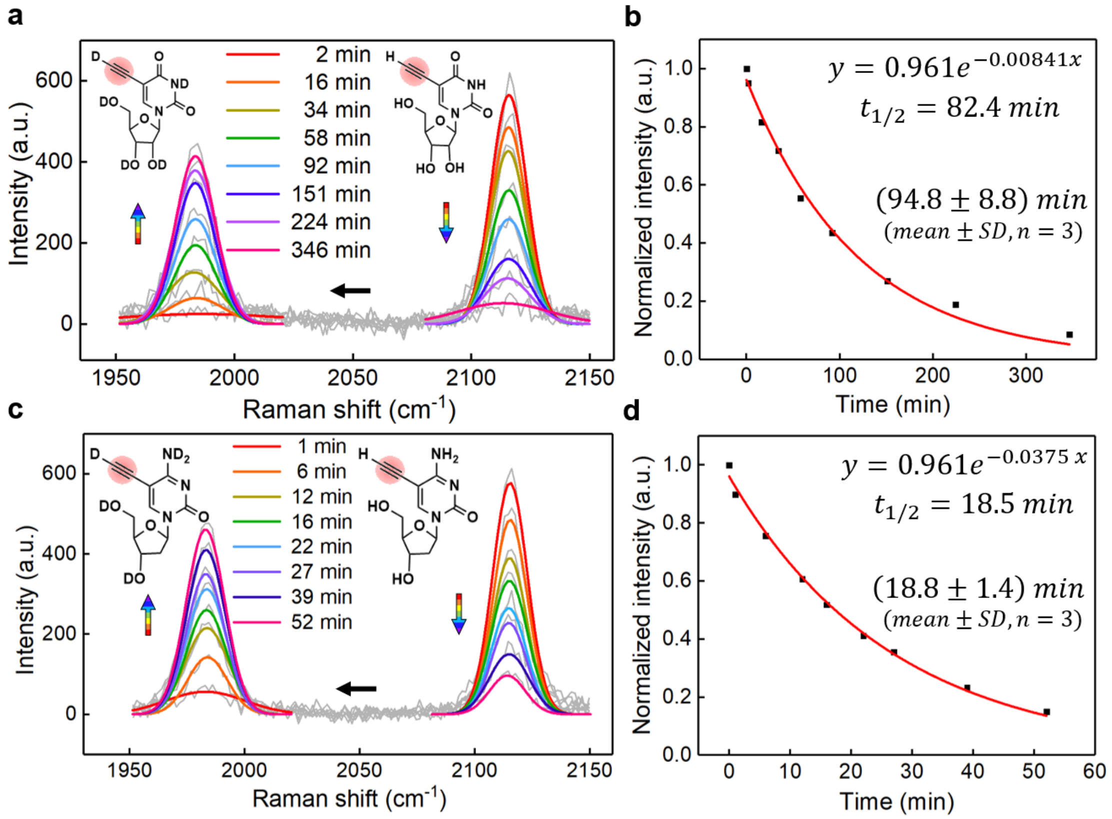


Figure S2 The alkyne-HDX rates for EdC and EU solutions. a) The Kinetics trace of the spontaneous Raman spectra for the decrease of alkyne-H peaks and the increase of alkyne-D peaks from EU (10 mM) during alkyne-HDX in the pD=7.6 D_2O buffer solution. b) Exponential curve fitting of normalized alkyne peak intensities for EU (10 mM) from (a). The average of $t_{1/2}$ over three independent measurements is also shown. c) The kinetics trace of the spontaneous Raman spectra for the decrease of alkyne-H peaks and the increase of alkyne-D peaks from EdC (20 mM) during alkyne-HDX in the pD=7.6 D_2O buffer solution. d) Exponential curve fitting of normalized alkyne peak intensities for EdC (20 mM) in (c). The average of $t_{1/2}$ over three independent measurements is also shown.

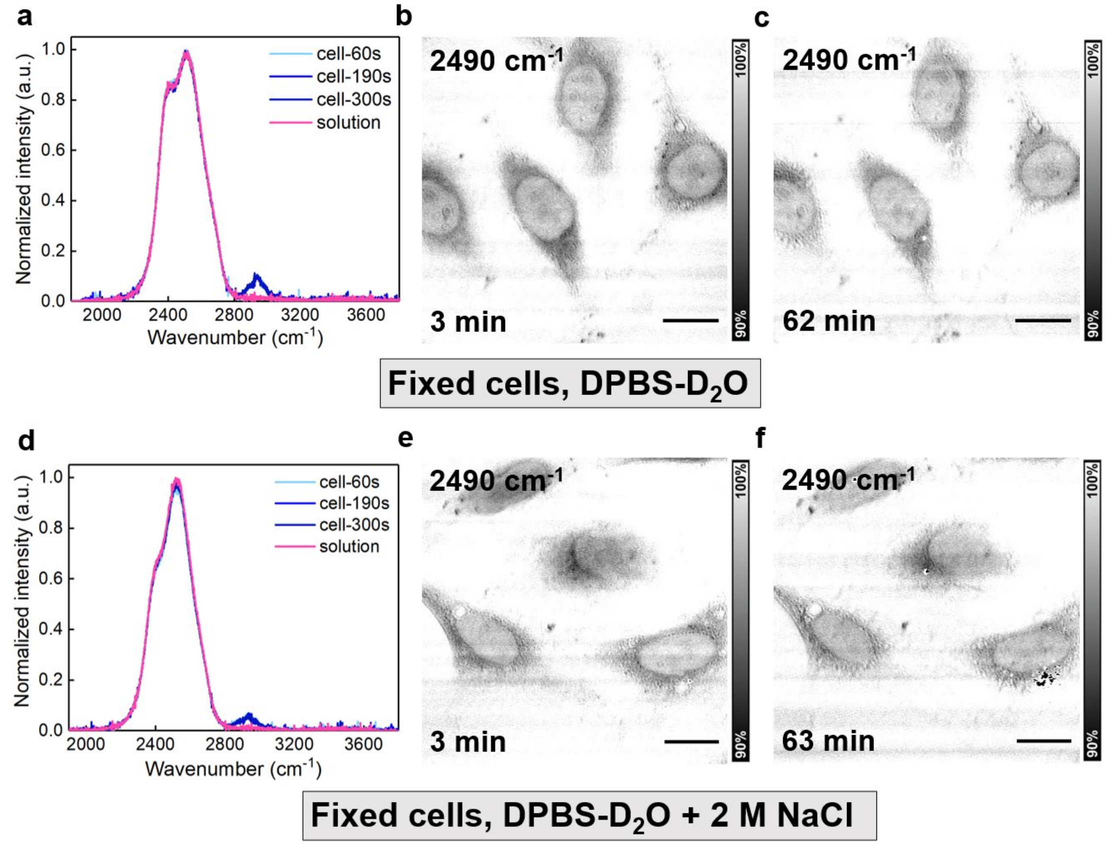


Figure S3 Fast D_2O diffusion across cells in solutions with different osmolarities. a) Time-trace spontaneous Raman spectra in the nuclear region of a fixed cell in the DPBS-D_2O solution. The solution spectrum is taken in the surrounding region without cells from the same sample. b-c) SRS image of O-D vibrational peak (2490 cm^{-1}) for the same set of fixed cells in DPBS-D_2O buffer solution after 3 min (b) and 62 min (c) incubation. d) Spontaneous Raman spectra in the nuclear region of a fixed cell in the high osmolarity solution (DPBS-D_2O + 2 M NaCl). The solution spectrum is taken in the surrounding region without cells from the same sample. e-f) SRS image of O-D vibrational peak (2490 cm^{-1}) for the same set of fixed cells in DPBS-D_2O + 2 M NaCl buffer solution after 3 min (e) and 63 min (f) incubation. Scale bar: 20 μm.

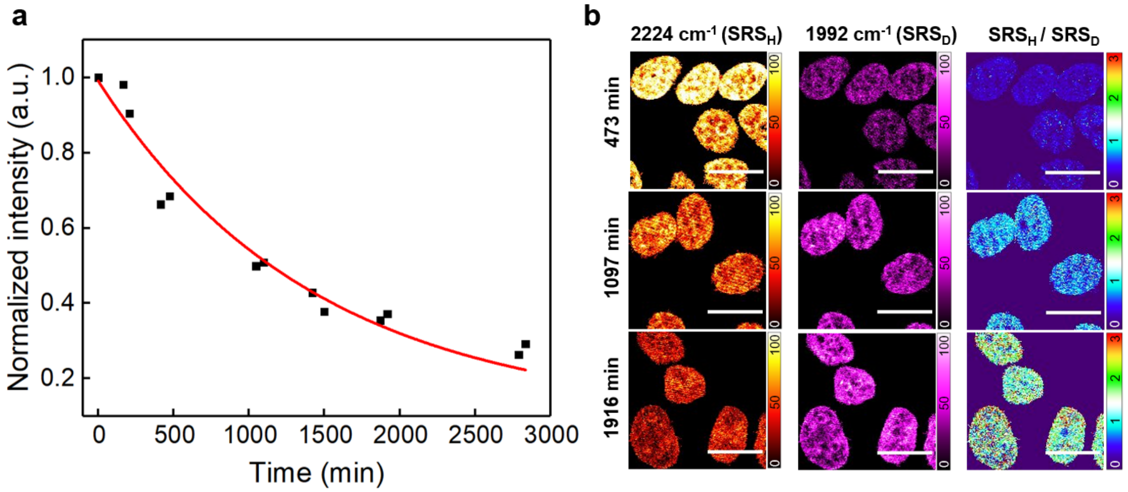


Figure S4 Representative data sets for SRS imaging-based alkyne-HDX kinetics on EdU-labeled cells. a) Exponential curve fitting of normalized alkyne peak intensities for EdU-incorporated cells in a pD=7.6 D_2O buffer solution. b) Representative ratiometric imaging of SRS_D / SRS_H (right) generated by dividing the SRS images at the alkyne-H channel (2224 cm^{-1}) (left) by those at the alkyne-D channel (1992 cm^{-1}) (middle) for cells immersed in DPBS-D_2O from a series of alkyne-HDX time points (473 min, 1097 min, 1916 min).

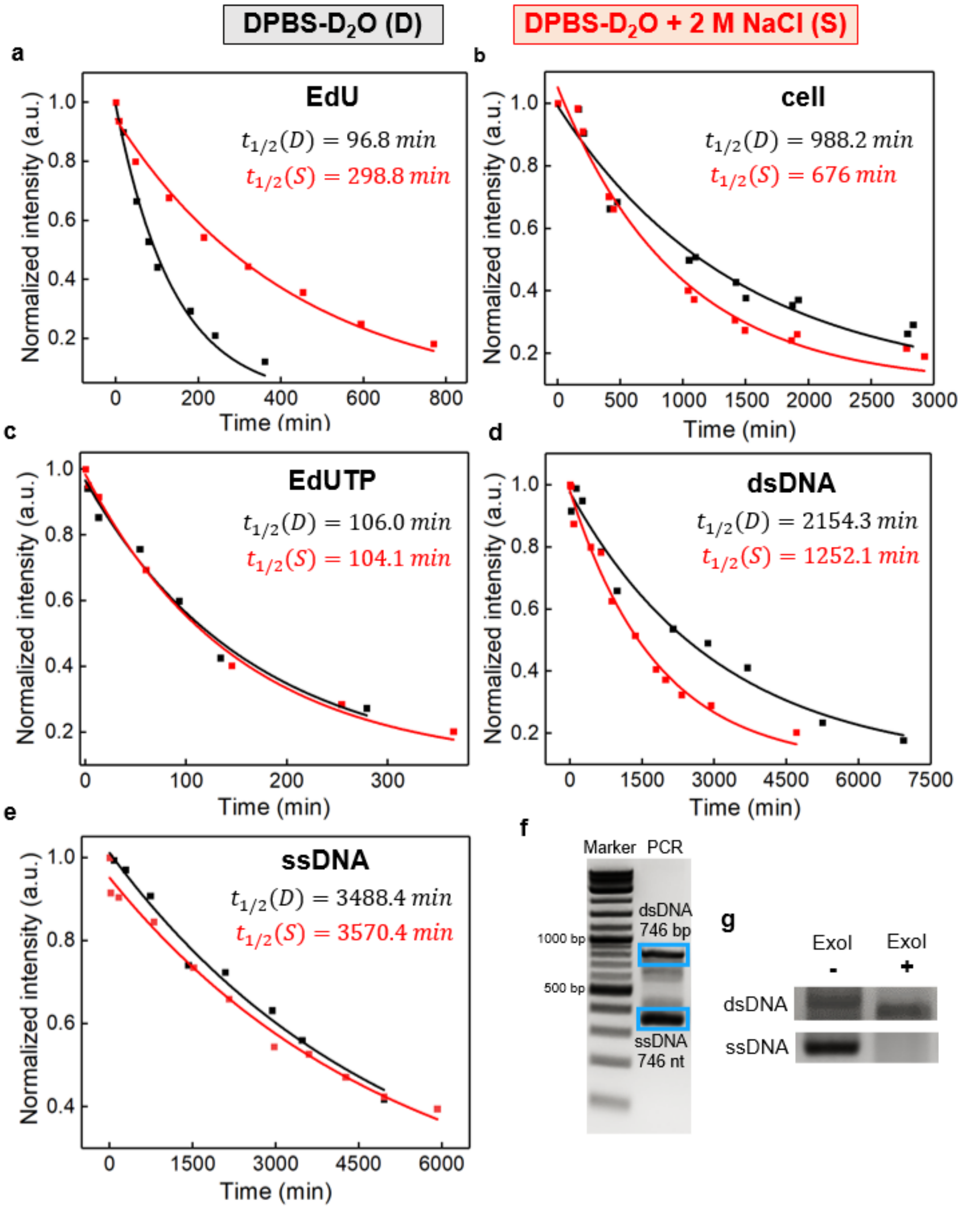


Figure S5 The alkyne-HDX kinetics for EdU and EdU-labeled structures at the regular (DPBS-D$_2$O, D, black) and high (DPBS-D$_2$O+2M NaCl, S, red) salt concentration. a-e) Exponential curve fitting of normalized alkyne peak intensities for EdU (10 mM) (a); EdU-incorporated cells (b); EdUTP (10 mM) (c); EdU-labeled dsDNA (d); and EdU-labeled ssDNA (e) in a pD=7.6 D$_2$O buffer solution (black) and pD=7.6 D$_2$O buffer solution with 2M NaCl (red), respectively. f) DNA gel electrophoresis for confirming the EdU-incorporated products of dsDNA and ssDNA. Left: 1 kb DNA ladder; right: EdU-labeled

PCR products, containing 746 bp dsDNA and 746 nt ssDNA. g) DNA gel electrophoresis confirming the presence of EdU-dsDNA and EdU-ssDNA with Exonuclease I (ExoI) digestion. Left: control PCR products; right: PCR products after Exonuclease I (ExoI) digestion.

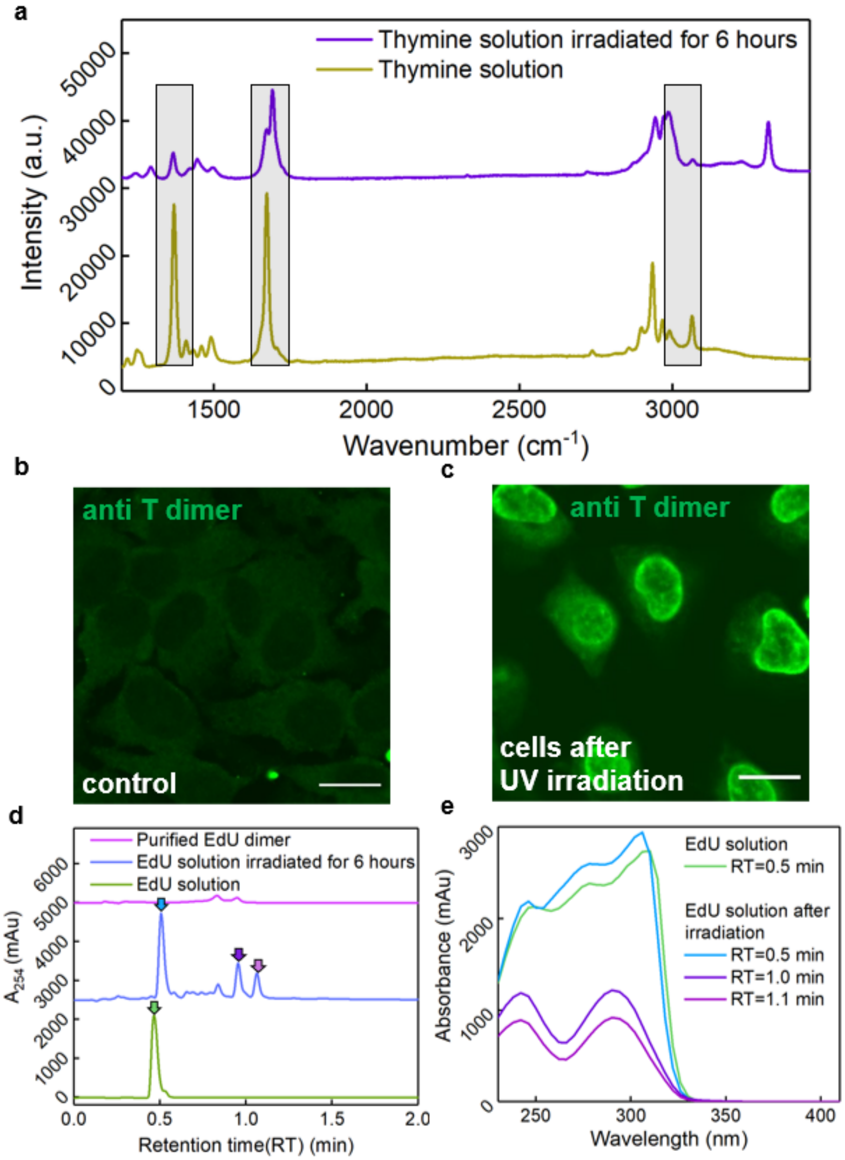


Figure S6 Confirmation for the UV-induced dimer formation. a) Spontaneous Raman spectra of thymine solution (yellow) and thymine solution irradiated for 6 hours (purple).

The gray-boxed regions show the featured Raman spectra changes for thymine dimer formation, consistent with what was reported in Ref. 62. b-c) Immunofluorescence staining images with anti-T dimer-Alexa Fluor 647 for control cells without UV irradiation (control, b) and UV-irradiated cells (c). d) HPLC trace of EdU solution (light green), EdU solution irradiated for 6 hours (light purple) and purified EdU dimer (magenta). Arrowed peaks indicate the corresponding color-coded UV absorption traces in (e). e) UV absorption traces of peaks indicated by color-coded arrows in the HPLC trace (d) of the EdU solution before and after UV irradiation.

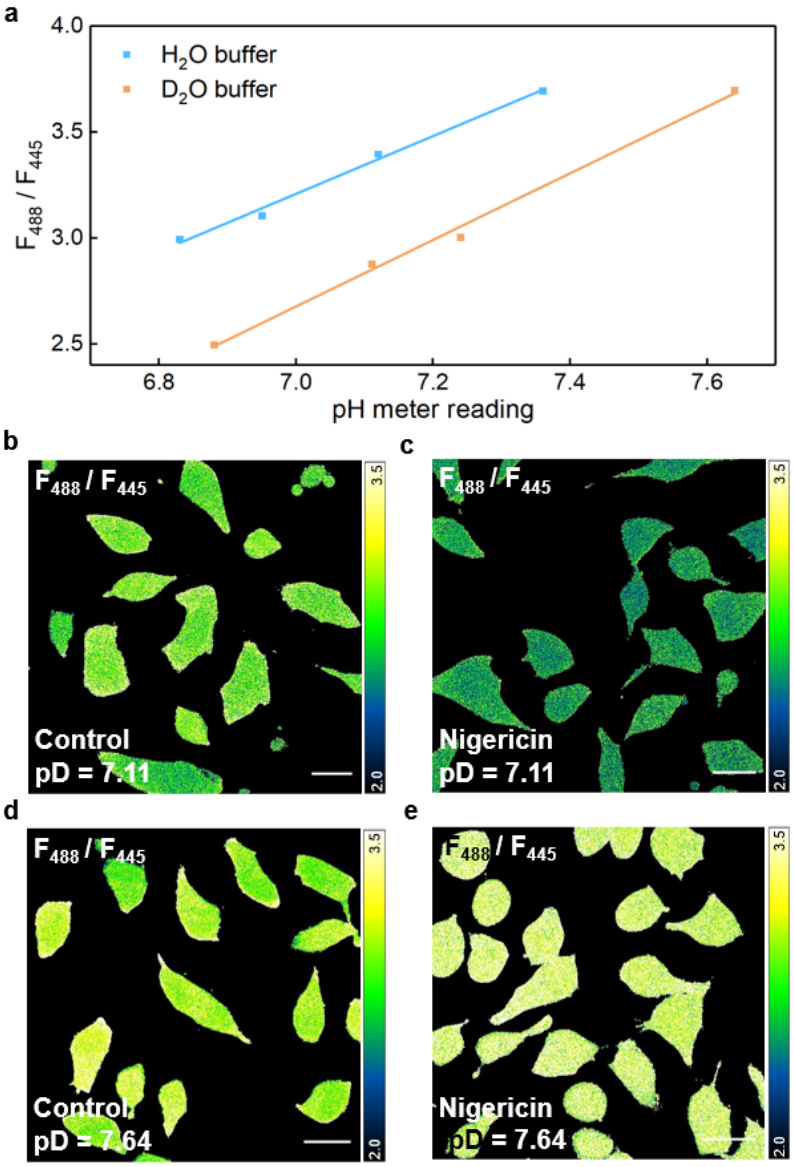


Figure S7 Calibration of solution and intracellular pD by a ratiometric fluorescent pH sensor BCECF. a) Linear relationship between ratios (F_{488} / F_{445}) taken at the 488 nm excitation and 445 nm excitation of BCECF in both H_2O (light blue) and D_2O (orange) buffers with varying reading results from the pH meter over the physiological-relevant range. b-c) Representative ratiometric images (F_{488} / F_{445}) for BCECF in live cells in pD=7.11 DPBS-D_2O buffer without (b, control, average ratio 2.93±0.38) and with nigericin (c, average ratio 2.55±0.31). d-e) Representative ratiometric images (F_{488} / F_{445}) for BCECF in live cells in pD=7.64 DPBS-D_2O buffer, without (d, control, average ratio 3.01±0.31) and with nigericin (e, average ratio 3.34±0.44). Scale bar: 20 μm.

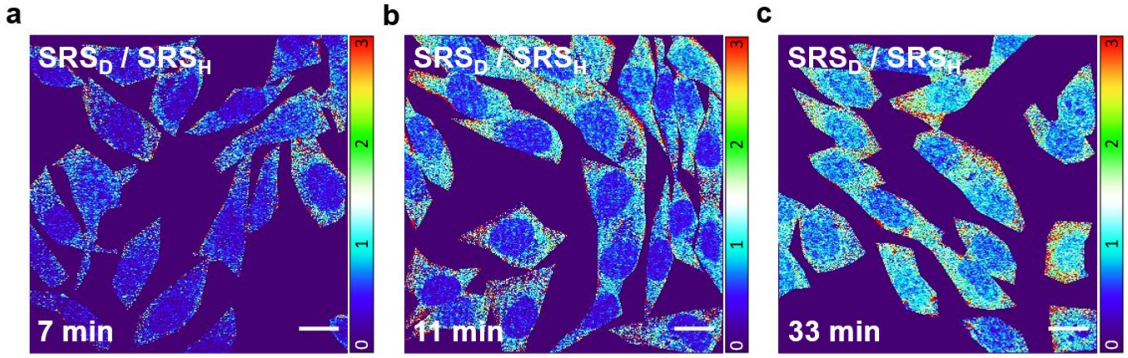


Figure S8 Ratiometric imaging for "two-color" identification of EdU-labeled DNA and PCho-labeled membranes by distinct alkyne-HDX kinetics between the two labeled species. a-c) Ratiometric imaging (SRS_D (1998 cm^{-1})/SRS_H (2137 cm^{-1})) for both EdU and PCho in live cells at three exchange time points. Scale bar: 20 μm.

5.6 Method and materials

Chemicals

All chemicals were purchased from Sigmal-Aldrich unless otherwise specified. 5-Ethynyl-2′-deoxyuridine (EdU, CAS# 61135-33-9) was purchased from TCI America. 5'-Ethynyl-2'-deoxycytidine (EdC, CAS# 69075-47-4) was purchased from Cayman Chemicals. 5-Ethynyl Uridine (EU, CAS# 69075-42-9) was purchased from Sigma-Aldrich. 5-Ethynyl-2'-deoxyuridine 5'-triphosphate (5-EdUTP) was purchased from Jena Bioscience. BCECF, AM (2',7'-Bis-(2-Carboxyethyl)-5-(and-6)-Carboxyfluorescein, Acetoxymethyl Ester) and BCECF, free acid were purchased from ThermoFisher scientific. Pwo DNA polymerase was purchased from Roche. Exonuclease I (ExoI) was purchased from Sigma-Aldrich. Nigericin sodium salt (CAS# 28643-80-3) was purchased from Sigma-Aldrich. Propargylcholine bromide (PCho, CAS# 111755-76-1) was synthesized according to Ref. 3. Briefly, propargyl bromide (80 wt. % solution in toluene, Sigma-Aldrich) was added dropwise to a stirring solution of 2-dimethylaminoethanol (Sigma-Aldrich) in anhydrous THF in ice bath under argon gas protection. The reaction mixture was slowly warmed up to room temperature and stirred overnight. The resulting white solids were filtered and washed extensively with cold anhydrous THF to obtain pure PCho.

Buffer preparation

For solution experiments, we used different buffering systems according to the final pD region to prepare D_2O buffers. (Buffer range: citric acid-Na_2HPO_4, 2.6-7.6; NaH_2PO_4-Na_2HPO_4, 6.2-8.2; CAPSO-CAPSO sodium salt, 8.9-10.3). We prepared D_2O buffers by directly diluting corresponding buffer powders to the same amount of D_2O as that for the H_2O system. To keep ion strength consistent for all buffers to avoid the influence on the alkyne-HDX kinetics, we added extra NaCl to those buffers with less salt concentrations. The detailed buffer recipes are shown below: pD= 5.3 (46.4 mM citric acid, 107.2 mM Na_2HPO_4); pD= 6.2 (33.9 mM citric acid, 132.2 mM Na_2HPO_4); pD= 6.6 (27.25 mM citric acid, 145.5 mM Na_2HPO_4); pD= 7.0 (17.65 mM citric acid, 164.7 mM Na_2HPO_4); pD= 7.6 (DPBS); pD= 7.9 (3 mM NaH_2PO_4, 30 mM Na_2HPO_4, 60 mM NaCl); pD= 9.4 (28mM CAPSO, 7 mM CAPSO sodium salt, 143 mM NaCl); pD= 10.4 (19 mM CAPSO, 16 mM CAPSO sodium salt, 134 mM NaCl); pD= 10.4 (10 mM CAPSO, 25 mM CAPSO sodium salt, 125 mM NaCl). The final pD values were determined by the pH-meter. For salt concentration experiment, 2 M NaCl was added into the DPBS-D_2O buffer solution.

For BCECF experiments, high K^+ conditions were used to meet the requirement of nigericin. The detailed high K^+ buffer recipes are shown below: High K^+ D_2O buffer, pD = 7.64 (120 mM KCl, 5 mM NaCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 in D_2O); pD=6.88, 7.11, 7.24 buffers were made by adjusting pD of pD=7.64 buffer with drops of 0.1 M HCl in D_2O. pD= 8.01 buffer was made by adjusting pD of pD=7.64 with drops of 0.1 M $NaHCO_3$ in D_2O. High K^+ H_2O buffer, pH = 7.35 (120 mM KCl, 5 mM NaCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 in H_2O); pH = 6.83, 6.95, 7.12 buffers were made by adjusting pH of pH=7.64 buffer using 0.1 M HCl in H_2O.

pH/pD determination

pD/pH reading for all D_2O buffers and H_2O buffers were acquired on a pH meter (Mettler Toledo FiveEasy Plus FP20, with LE 410 sensor) at room temperature (~ 22 °C). The pH meter was calibrated with standard solutions (pH=4.01, 7.00, 10.01, Mettler Toledo) before measuring the custom-made buffers. As we discussed in the main manuscript, to avoid confusion, we show all the pD values in our manuscript from direct pH meter reading (i.e., the $pH_{\text{meter reading}}$) for all D_2O buffers. As shown below, the calibration factor is small and would not influence any of our conclusions.

The offset between $pH_{\text{meter reading}}$ and pD values was obtained by two control experiments. First, we compared the $pH_{\text{meter reading}}$ of DPBS-H_2O solution (7.35, considered as real pD if H_2O is replaced with D_2O) to that of DPBS-D_2O solution (7.6, considered as $pH_{\text{meter reading}}$). Second, we used the pH-sensitive ratiometric fluorophore BCECF to provide additional calibration between $pH_{\text{meter reading}}$ and pD values reported through BCECF ratios. In brief, we dissolved BCECF acid in water and made a 2 mM stock solution. We then diluted the stock BCECF solution in the high K^+ H_2O/D_2O buffers with adjusted $pH_{\text{meter reading}}$ (directly reading from the pH meter) to be in the range 6.8-8.1. The final BCECF concentration is 4 μM. We then acquired fluorescence images using 445 nm and 488 nm excitation lasers (ZEISS LSM 980). The ratios of fluorescence intensity at 488 nm over that at 445 nm were plotted against respective readings from the pH meter for H_2O buffers and D_2O buffers (Figure S7a). The offset between linearly fitted ratio vs pH/pD curves is 0.25. The reading from the pH meter for D_2O buffers is shown as $pH_{\text{meter reading}}$ (and are reported as the pD values in our manuscript). The determined relationship is consistent with the above two different experiments: $pD=pH_{\text{meter reading}} - 0.25$.

DFT calculation

DFT calculations were performed using the Gaussian09 software. Structures were optimized and then characterized using frequency calculations at the B3LYP/6-311(G)++(d,p) level of theory.

Spontaneous Raman Spectroscopy

Spontaneous Raman spectra were acquired using an upright confocal Raman spectrometer (Horiba Raman microscope; Xplora plus). A 532 nm YAG laser was used to illuminate the sample with a power of 12 mW through a 100×, N.A. 0.9 objective (MPLAN N; Olympus). Data acquisition was performed with 10 s integration by the LabSpec6 software. For whole spectra recording, background was subtracted by measuring signal from the same solution without probe molecules. The spectra shown in Figure 1a and S1a are normalized to the alkyne peak.

EdU, EdC and EU were dissolved into DMSO to make 100 mM stock solutions. PCho was dissolved into H_2O to make 2M stock solution. For measurement, EdU, EdC and EU are 1:10 diluted into corresponding H_2O buffers or D_2O buffers, while PCho is 1:50 diluted into corresponding H_2O buffers or D_2O buffers.

Model molecules (4-fluorophenylacetylene (CAS# 766-98-3), methyl-4-ethynylbenzoate (CAS# 3034-86-4), 4-ethynylbenzaldehyde (CAS# 63697-96-1), 1-ethynyl-4-nitrobenzene (CAS# 937-31-5), 4-Ethynylanisole (CAS# 768-60-5)) were dissolved into DMSO to make 100 mM stock solutions. For measurement in DMSO-D_2O system, model molecules were diluted into the 1:1 DMSO-D_2O (DPBS- D_2O, pD=7.6) solution to ensure good dissolvability for all model molecules with corresponding dilution factors. For measurement in methanol-OD system, model molecules were 1:10 diluted into methanol-OD (CAS# 1455-13-6).

All data are confirmed by at least three sets of independent experiments.

Stimulated Raman Scattering (SRS) Microscopy

A picoEmerald laser system (Applied Physics and Electronics) was used as the light source for SRS microscopy. Briefly, it produces 2 ps pump (tunable from 770 nm – 990 nm, bandwidth 0.5 nm, spectral bandwidth ~ 7 cm^{-1}) and Stokes (1031.2 nm, spectral bandwidth 10 cm^{-1}) pulses with 80 MHz repetition rate. Stokes beam is modulated at 20 MHz by an internal electro-optic modulator. The spatially and temporally overlapped Pump and Stokes beams are introduced into an inverted laser-scanning microscope (FV3000, Olympus), and then focused onto the sample by a 25X water objective (XLPLN25XWMP, 1.05 N.A., Olympus). Transmitted Pump and Stokes beams are collected by a high N.A. condenser lens (oil immersion, 1.4 N.A., Olympus) and pass through a bandpass filter (893/209 BrightLine, 25mm, Semrock) to filter out Stokes beam. A large area (10×10 mm) Si photodiode (S3590-09, Hamamatsu) is used to measure the pump beam intensity. A 64 V reverse-biased DC voltage is applied on the photodiode to increase the saturation threshold and reduce response time. The output current is terminated by a 50Ω terminator and pre-filtered by a 19.2-23.6-MHz band-pass filter (BBP-21.4+, MiniCircuits) to reduce laser and scanning noise. The signal is then demodulated by a lock-in amplifier (SR844, Stanford Research Systems) at the modulation frequency. The in-phase X output is fed back to the Olympus IO interface box (FV30-ANALOG) of the microscope. 30 μs time constant is set for the lock-in amplifier. Correspondingly, 80 μs pixel dwell time is used, which gives a speed of 21.3 s/frame for a 512-by-512-pixel image, with two frame-averaging. Laser powers are monitored throughout image acquisition by an internal power meter and power fluctuations are controlled within 1%. The power for the pump and Stocks beam is about 25 mW and 220 mW, respectively. 16-bit greyscale images were acquired by Olympus Fluoview 3000 software. To minimize the line-pattern issue likely due to an interfering Radio frequency (RF) picked up by our lock-in amplifier detection (demodulation at 20 MHz), we have optimized our alignment and replaced a few bandpass filters between our photodiode and lock-in amplifier.

For EdU and PCho measurement, the wavelengths of pump lasers for SRS_H are 845.9 and 844.6 nm, respectively. For ratiometric imaging of EdU and PCho after the exchange, the wavelengths of pump laser for SRS_D are 855.5 and 854.5 nm, respectively. Off-resonance images were taken under 851.3 nm pump wavelength. For D_2O diffusion measurement, the pump wavelength is 820.5 nm. For EdU/EdU dimer spectra recordings, the wavelengths of pump lasers tuned from 843.9 to 847.9 nm with a 0.5 nm interval. For two-color ratiometric imaging of EdU and PCho during the exchange, the wavelengths of pump lasers for SRS_H and SRS_D are 845 and 855 nm, respectively.

Cell culture, sample preparation and alkyne-HDX in cells

For all SRS imaging experiments, cultured HeLa-CCL2 (ATCC) cells were seeded onto coverslips (12mm, #1.5, Fisher) with a density of 1×10^5/mL in 4 well plate with 0.3 mL DMEM culture medium (DMEM+10%FBS+1% penicillin-streptomycin) for 20 h at 37 °C and 5% CO_2. Prior to imaging, coverslips were collected and attached to a microscope slide (1mm thick, VWR) with an imaging spacer (0.12mm thick, Sigma-Aldrich).

For the EdU experiment, DMEM culture medium was then changed to DMEM medium (FBS-free, Gibco) for 20-22 h for cell cycle synchronization. After synchronization, the medium was replaced back to DMEM culture medium and EdU (10 mM stock in DPBS) was simultaneously added to a concentration of 100 μM for 20-24 h. Then 4% PFA was added for 20 min for fixation. After that, DPBS was used to wash away PFA and fixed cells could be stored in DPBS at 4 °C for several days.

For UV-irradiation on live cells, cells were put inside the biosafety cabinet (BSC) with UV (254 nm) on for one hour. Morphologies were quickly checked with no severe abnormality under a transmission light microscope. The cells were fixed by 4% PFA immediately after the UV irradiation.

For all the fixed cells alkyne-HDX experiments, corresponding D_2O buffers were used to wash the cells three times and then the coverslip was taken out to make an imaging chamber filled with designating D_2O buffers for SRS imaging.

For live-cell BCECF experiments, cells were first loaded with 2 μM BCECF-AM in HBSS for 20 min in the CO_2 incubator and were subsequently treated with 10 μM nigericin in pH=7.35 high K^+ H_2O buffer for 5 min. Cells were switched into high K^+D_2O buffers containing 10 μM nigericin with designated pD and put onto the microscope slide. The image was collected on an inverted confocal microscope (ZEISS LSM 980) using 40x water immersion objective (NA 1.2) with either 445 or 488 nm excitation laser. Both images were acquired using the same PMT which was set to collect photons from 530-570 nm. Control experiments were done under similar conditions yet with the addition of DMSO instead of nigericin in all the buffer solutions.

For live-cells alkyne-HDX experiment for pD sensing, HeLa cells were first incubated with 10 μM nigericin in high K^+ H_2O buffer for 5 min at 37 ˚C. The high K^+ H_2O buffer was then removed, and cells were washed with high K^+ D_2O buffer with 10 μM nigericin of different pD values. We started timing as the cells were washed with high K^+ D_2O buffers. The coverslip with cells was placed onto a microscope slide with spacer filled with 10 μM nigericin high K^+ D_2O buffers. Control experiments were done under similar conditions but with the addition of DMSO instead of nigericin in all the buffer solutions.

For the PCho and EdU experiment, DMEM culture medium was changed to DMEM medium (FBS-free, Gibco) for synchronization. After synchronization, medium was replaced back to DMEM culture medium by simultaneously adding both propargylcholine (100 mM stock in DPBS) and EdU (10 mM stock in DPBS) to the culture medium with a final concentration of 1 mM and 100 μM, respectively, for 20-24 h. For the live cells alkyne-HDX experiments, home-made DMEM-D_2O buffer through dissolving DMEM powder into D_2O, was used to wash the cells three times. Then the coverslip was taken out to make an imaging chamber filled with the DMEM-D_2O buffer for SRS imaging.

Synthesis and purification of dsDNA and ssDNA

Pwo DNA polymerase, an enzyme showing good performance to accept modified nucleotides, was used to incorporate EdUTP into DNA through PCR process. A typical PCR reaction contained ~100 ng plasmid template, 0.05 mM each of the forward and reverse primers, 2.5 U polymerase and 10x polymerase buffer with magnesium, 100μM of dNTPs (dATP, dCTP and dGTP, NEB) instead of dTTP and

100μM modified EdUTP. The reactions were done in an overall volume of 50 μL with the addition of UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen). For ssDNA synthesis, asymmetric PCR was used. The concentration of the forward primer and the reverse primer is 0.001 mM and 0.05 mM, respectively, while other conditions are the same with that indicated above. PCR experiments were performed on a Bio-Rad C1000 thermal cycler.

Primers and the template employed for PCR experiments:

Forward primer: 5´- GGAAATCGGTACTGGCTTTCCATTCGAC

reverse primer: 3´- GTGAGTTAAAGTTGTACTCGAGTTTGTGTCCG

Template (sequence of 746 bp, contains 382 T):

GGAAATCGGTACTGGCTTTCCATTCGACCCCCATGATGGTTCCGTTCAACTAGC
AGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAG
ACAACCATTACCTGTCGACACAATCTGTCCTTTCGAAAGATCCCAACGAAAAG
CGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGC
ATGGATGAGCTCTACAAAGGCGGTGGGTCGGGCGGGGGCTCCCCCGGGGGTG
GCGGTTCATGATCAGGTGGAGGGTCAGGGGGCGGATCAATGAGCAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAA
TGGGCACAAATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGA
AAACTCACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCA
ACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGGAT
CACATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAACGCACTATATCTTTCAAAGATGACGGGACCTACAAGACGCGTGCTGAA
GTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAAGGTATTGA
TTTTAAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTTTAACT
CAC

PCR products were purified with a Monarch PCR & DNA Cleanup Kit (NEB) and confirmed by gel electrophoresis and Sanger sequencing. The products of PCR reaction were separated by 1% agarose (UltraPure™ Low Melting Point Agarose, Invitrogen) gel electrophoresis. dsDNA and ssDNA were purified with a Monarch Gel Extraction Kit (NEB). Exonuclease I (Thermo Scientific, 20 U/μL) was used to digest ssDNA in 37℃, which was confirmed by gel electrophoresis as well. Before enzyme digestion, 5 min 95℃ heat shock was used for denaturation. The images were recorded with Bio-Rad Gel and Blot Imaging Systems. The combined PCR products (~50 tubes per sample) were loaded into Microcon-10kDa Centrifugal Filter (Millipore Sigma) and washed with UltraPure™ DNase/RNase-Free Distilled Water three times to remove the remaining salts and EDTA in the elusion buffer through buffer exchange. Then the PCR products in pure water were concentrated into ~1 uL through vacufuge (Vacufuge Plus Concentrator, Eppendorf). The concentrated PCR products were diluted into DPBS-D_2O / DPBS-D_2O + 2 M NaCl buffer solutions for alkyne-HDX kinetics measurements.

Formation and purification of EdU dimers

For solution samples, EdU (or thymine) was dissolved into water to make saturated solutions. After frozen into ice, the EdU (or thymine) solution was put into a homemade dry ice chamber and was irradiated with a UV lamp (254nm, UVLS-24 EL Series UV Lamp, 4 Watt) for 6 hours. Analytical HPLC coupled with mass spectrometry (LC-MS) for the UV-irradiated EdU product was performed on Agilent 1290 infinity LC system using ZORBAX RRHD Eclipse Plus C18, 95Å, 2.1 x 50 mm, 1.8 μm column with Agilent 6140 Series Quadrupole LCMS/LC-MS/MSD/Mass Spectrometer System. The mobile phase is water (0.1% AcOH) and acetonitrile with a running method of gradient 40% - 95% acetonitrile (1 ml/min, 10 min for total running time). The data shown in the supplementary figures (Figure S6e) are the absorption (254 nm) intensity traces.

The aqueous reaction mixture (100 mg) post UV-radiation was first concentrated under reduced pressure at 40 ˚C and was then load on a reverse phase Biotage cartridge (12g SNAP Ultra C18). The flash column chromatography was automated by the Biotage Isolera System, with acetonitrile and water as the mobile phases. The flushing gradient was set to 5% to 40% acetonitrile over 10 column volume (CV) and the EdU dimer should be the second UV active portion eluted around the second to third CV. The collect portion was confirmed by normal phase TLC. EdU dimer would have an R_f value of 0.15 and EdU has an R_f of 0.6 when running with pure Ethyl Acetate. The purified EdU dimer was confirmed through high-resolution mass spectrometry (calculated exact mass: 505.1571, detected mass: 505.1575) and dissolved into H_2O to prepare a stock solution for further alkyne-HDX measurements.

Immunofluorescence staining

The UV-irradiated cells and control cells were first fixed and treated with 0.2% Triton-X-100 in 1x DPBS (no calcium, no magnesium, Gibco) at room temperature for 30 min. Then the Triton-X-100 was removed, and cells were washed with 1x DPBST (DPBS + 0.1% Tween 20) three times. Cells were incubated with 1% BSA, 22.52 mg/mL glycine in PBST at room temperature for 60 min for blocking. After washed with PBST, cells were incubated with 1:300 diluted anti-thymine dimer antibody (Mouse monoclonal, ~2 mg/mL, MilliporeSigma) in 0.1% BSA, PBST overnight at 4°C. The primary antibody solution was removed, and cells were washed three times with DPBST. Then cells were incubated with 3% BSA in PBST at room temperature for 60 min for blocking. After washed with PBST, cells were incubated with 1:400 diluted Goat anti-Mouse IgG (H+L), (Invitrogen, Superclonal™ Recombinant Secondary Antibody, Alexa Fluor 647, 1 mg/mL) in 0.1% BSA, PBST for 2 hours at room temperature. The secondary antibody solution was removed, and cells were washed three times with DPBST. Fluorescent imaging was conducted immediately after sample preparation through the same Olympus FV3000 confocal microscope with CW laser excitation (640 nm, Coherent OBIS LX laser) and standard bandpass filter sets.

Data processing

All spectra were processed using LabSpec6 software. Spectral baselines were subtracted. Peak centers and intensity were read out by Gaussian peak fitting. All images were processed using ImageJ software. Corresponding off-resonance images were subtracted.

For precise $t_{1/2}$ fitting, we obtained the time-zero intensity (i.e., I_0(R-C≡C-H)) by the normalization of I_1(R-C≡C-H) and I_1(R-C≡C-D), the intensity measured from the first exchange data point at the alkyne-H and the alkyne-D channels, respectively. I_0(R-C≡C-H) is defined as I_1(R-C≡C-H)+I_1(R-C≡C-D)/r, in which r is the intensity correction ratio, defined as dividing I(R-C≡C-D), the alkyne-probe solution after equilibrium in D_2O (i.e., finished with exchange), by I(R-C≡C-H), the corresponding alkyne-probe solution in H_2O with the same concentration without any exchange. This strategy was used for both the spontaneous Raman measurement of solution samples and the SRS imaging recording of cell samples. I(R-C≡C-H) and I(R-C≡C-D) from spontaneous Raman measurements were read out by Gaussian peak fitting. I(R-C≡C-H) and I(R-C≡C-D) from SRS measurements (also refer as SRS_H and SRS_D) were read out from all the nucleus regions in one field of view based on SRS images.

Derivation of alkyne-HDX kinetics

The rate-determining step (RDS) in HDX between alkyne-tagged probes (R-C≡C-H) and catalytic base OD^- is shown in Eq. (1) below. It includes three elementary steps: a) diffusional collision, b) equilibrium redistribution of the hydrogen in the intermediate state, and c) dissociation.

$$R-C\equiv C-H+OD^- \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} R-C\equiv C-H \ldots O-D^- \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} R-C\equiv C^- \ldots H-O-D \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} R-C\equiv C^- + HOD \qquad (1)$$

We can assume that: 1) the temperature remains constant; 2) the reaction is diffusion-limited; 3) the intermediates are in steady states. So the concentrations of the two intermediates remain the same in the reaction, shown in Eq. (2).

$$\frac{d[(R-C\equiv C-H\cdots OD)^-]}{dt} = \frac{d[(R-C\equiv C\cdots HOD)^-]}{dt} = 0 \qquad (2)$$

i.e.,

$$\begin{aligned} &k_1[R-C\equiv C-H][OD^-] - k_{-1}[(R-C\equiv C-H\cdots OD)^-] + k_{-2}[(R-C \\ &\qquad \equiv C\cdots HOD)^-] - k_2[(R-C\equiv C-H\cdots OD)^-] = 0 \\ &k_2[(R-C\equiv C-H\cdots OD)^-] - k_{-2}[(R-C \\ &\qquad \equiv C\cdots HOD)^-] - k_3[(R-C\equiv C\cdots HOD)^-] = 0. \end{aligned}$$

So we can get the concentration of the intermediate shown in Eq. (3).

$$[(R-C\equiv C\cdots HOD)^-]=\frac{k_1k_2}{k_{-1}k_{-2}+k_{-1}k_3+k_2k_3}[R-C\equiv C-H][OD^-]. \qquad (3)$$

The overall rate (k) for transferring a proton from R-C≡C-H to OD^- then becomes:

$$k=\frac{d[R-C\equiv C^-]}{dt}=k_3[(R-C\equiv C\cdots HOD)^-]=\frac{k_1k_2k_3}{k_{-1}k_{-2}+k_{-1}k_3+k_2k_3}[R-C\equiv C-H][OD^-]. \quad (4)$$

Since the reaction is diffusion-limited, there are some restrictions on rate constants:

$$k_{-1}\ll k_2 \text{ or } k_3\ll k_{-2}, \qquad k_{-1}\approx k_3.$$

So the overall rate can be simplified as Eq. (5).

$$k=\frac{k_1}{\frac{k_{-2}}{k_2}+1}[R-C\equiv C-H][OD^-]. \qquad (5)$$

In addition, $\frac{k_{-2}}{k_2}=10^{pK_a(R-C\equiv C-H)-pK_a(HOD)}\gg 1$, so Eq. (5) can be further simplified as Eq. (6).

$$k=k_1\cdot(10^{pK_a(\,donor,\ R-C\equiv C-H)-pK_a(acceptor,\ HOD)}+1)^{-1}\cdot[OD^-]\cdot[R-C\equiv C-H] \quad (6)$$

k_1 is the diffusion-limited collision constant, upper-bounded by 10^{10} $M^{-1}s^{-1}$. As fluctuation of OD^- concentration is negligible during HDX, the RDS is considered as a pseudo-first-order reaction. Since the acceptor pK_a is much smaller than the donor pK_a, the corresponding exchange half-life ($t_{1/2}$) is reduced to Eq. (7),

$$\lg(t_{1/2}) = pK_a-pD + l\,g(ln2)-10-pK_a\,(HOD)+pK_a(D_2O). \quad (7)$$

Here pK_a designates the donor pK_a, i.e., $pK_a(R-C\equiv C-H)$. Taking the pK_a of alkynyl hydrogen as 20-25 and the pK_a of HOD as ~15, we estimated the $t_{1/2}$ of alkyne-HDX in physiological pD (7.6) to be on the order of minutes. For example, if we used 20 for pK_a of alkynyl hydrogen, the calculated $t_{1/2}$ could be 174 s.

5.7 Reference

1. Jamieson, L. E.; Greaves, J.; McLellan, J. A.; Munro, K. R.; Tomkinson, N. C. O.; Chamberlain, L. H.; Faulds, K.; Graham, D., Tracking intracellular uptake and localisation of alkyne tagged fatty acids using Raman spectroscopy. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **197**, 30-36 (2018).
2. Shen, Y.; Hu, F.; Min, W., Raman imaging of small biomolecules. *Annu. Rev. Biophys.* **48**, 347-369 (2019).
3. Ando, J.; Asanuma, M.; Dodo, K.; Yamakoshi, H.; Kawata, S.; Fujita, K.; Sodeoka, M., Alkyne-tag SERS screening and identification of small-molecule-binding sites in protein. *J. Am. Chem. Soc.* **138** (42), 13901-13910 (2016).
4. Chen, Y.; Ren, J. Q.; Zhang, X. G.; Wu, D. Y.; Shen, A. G.; Hu, J. M., Alkyne-Modulated Surface-Enhanced Raman Scattering-Palette for Optical Interference-Free and Multiplex Cellular Imaging. *Anal. Chem.* **88** (12), 6115-6119 (2016).
5. Yamakoshi, H.; Dodo, K.; Palonpon, A.; Ando, J.; Fujita, K.; Kawata, S.; Sodeoka, M., Alkyne-tag Raman imaging for visualization of mobile small molecules in live cells. *J. Am. Chem. Soc.* **134** (51), 20681-20689 (2012).
6. Prince, R. C.; Frontiera, R. R.; Potma, E. O., Stimulated Raman scattering: from bulk to nano. *Chem. Rev.* **117** (7), 5070-5094 (2017).
7. Wei, L.; Hu, F.; Shen, Y.; Chen, Z.; Yu, Y.; Lin, C.-C.; Wang, M. C.; Min, W., Live-cell imaging of alkyne-tagged small biomolecules by stimulated Raman scattering. *Nat. Methods* **11** (4), 410-412 (2014).
8. Hong, S.; Chen, T.; Zhu, Y.; Li, A.; Huang, Y.; Chen, X., Live-cell stimulated Raman scattering imaging of alkyne-tagged biomolecules. *Angew. Chem. Int. Ed.* **53** (23), 5827-5831 (2014).
9. Hu, F.; Chen, Z.; Zhang, L.; Shen, Y.; Wei, L.; Min, W., Vibrational imaging of glucose uptake activity in live cells and tissues by stimulated Raman scattering. *Angew. Chem. Int. Ed.* **54** (34), 9821-9825 (2015).
10. Lee, H. J.; Zhang, W.; Zhang, D.; Yang, Y.; Liu, B.; Barker, E. L.; Buhman, K. K.; Slipchenko, L. V.; Dai, M.; Cheng, J. X., Assessing cholesterol storage in live cells and C. elegans by stimulated Raman scattering imaging of phenyl-Diyne cholesterol. *Sci. Rep.* **5** (7930), 1-10 (2015).
11. Tipping, W. J.; Lee, M.; Serrels, A.; Brunton, V. G.; Hulme, A. N., Imaging drug uptake by bioorthogonal stimulated Raman scattering microscopy. *Chem. Sci.* **8** (8), 5606-5615 (2017).
12. Hu, F.; Shi, L.; Min, W., Biological imaging of chemical bonds by stimulated Raman scattering microscopy. *Nat. Methods* **16** (9), 830-842 (2019).
13. Seidel, J.; Miao, Y.; Porterfield, W.; Cai, W.; Zhu, X.; Kim, S. J.; Hu, F.; Bhattarai-Kline, S.; Min, W.; Zhang, W., Structure-activity-distribution relationship study of anti-cancer antimycin-type depsipeptides. *Chem. Comm.* **55** (63), 9379-9382 (2019).
14. Zeng, C.; Hu, F.; Long, R.; Min, W., A ratiometric Raman probe for live-cell imaging of hydrogen sulfide in mitochondria by stimulated Raman scattering. *Analyst* **143** (20), 4844-4848 (2018).
15. Tipping, W. J.; Wilson, L. T.; Blaseio, S. K.; Tomkinson, N. C. O.; Faulds, K.; Graham,

D., Ratiometric sensing of fluoride ions using Raman spectroscopy. *Chem. Comm.* **56** (92), 14463-14466 (2020).

16. Wilson, L. T.; Tipping, W. J.; Jamieson, L. E.; Wetherill, C.; Henley, Z.; Faulds, K.; Graham, D.; Mackay, S. P.; Tomkinson, N. C. O., A new class of ratiometric small molecule intracellular pH sensors for Raman microscopy. *Analyst* **145** (15), 5289-5298 (2020).
17. Talley, C. E.; Jusinski, L.; Hollars, C. W.; Lane, S. M.; Huser, T., Intracellular pH sensors based on surface-enhanced Raman scattering. *Anal. Chem.* **76** (23), 7064-7068 (2004).
18. Qin, X.; Lyu, M.; Si, Y.; Yang, J.; Wu, Z.; Li, J., Alkyne-based surface-enhanced Raman scattering nanoprobe for ratiometric imaging analysis of caspase 3 in live cells and tissues. *Anal. Chim. Acta* **1043**, 115-122 (2018).
19. Si, Y.; Bai, Y.; Qin, X.; Li, J.; Zhong, W.; Xiao, Z.; Li, J.; Yin, Y., Alkyne-DNA-functionalized alloyed Au/Ag nanospheres for ratiometric surface-enhanced Raman scattering imaging assay of endonuclease activity in live cells. *Anal. Chem.* **90** (6), 3898-3905 (2018).
20. Fried, S. D.; Boxer, S. G., Measuring electric fields and noncovalent interactions using the vibrational stark effect. *Acc. Chem. Res.* **48** (4), 998-1006 (2015).
21. Marr, J. M.; Schultz, Z. D., Imaging electric fields in SERS and TERS using the vibrational stark effect. *J. Phys. Chem. Lett.* **4** (19), 3268-3272 (2013).
22. Epstein, S. C.; Huff, A. R.; Winesett, E. S.; Londergan, C. H.; Charkoudian, L. K., Tracking carrier protein motions with Raman spectroscopy. *Nat. Comm.* **10** (2227), 1-6 (2019).
23. Flynn, J. D.; Gimmen, M. Y.; Dean, D. N.; Lacy, S. M.; Lee, J. C., Terminal alkynes as Raman probes of alpha-synuclein in solution and in cells. *Chembiochem* **21** (11), 1582-1586 (2020).
24. Craven, C. J.; Derix, N. M.; Hendriks, J.; Boelens, R.; Hellingwerf, K. J.; Kaptein, R., Probing the nature of the blue-shifted intermediate of photoactive yellow protein in solation by NMR: Hydrogen-deuterium exchange data and pH studies. *Biochemistry* **39** (47), 14392-14399 (2000).
25. Yagi-Utsumi, M.; Chandak, M. S.; Yanaka, S.; Hiranyakorn, M.; Nakamura, T.; Kato, K.; Kuwajima, K., Residual structure of unfolded ubiquitin as revealed by hydrogen/deuterium-exchange 2D NMR. *Biophys. J.* **119** (10), 2029-2038 (2020).
26. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L., Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **100** (12), 7057-7062 (2003).
27. Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G., Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 1-25 (2003).
28. Konermann, L.; Pan, J.; Liu, Y. H., Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **40** (3), 1224-1234 (2011).
29. Jansson, E. T.; Lai, Y. H.; Santiago, J. G.; Zare, R. N., Rapid hydrogen-deuterium exchange in liquid droplets. *J. Am. Chem. Soc.* **139** (20), 6851-6854 (2017).
30. Vigano, C.; Smeyers, M.; Raussens, V.; Scheirlinckx, F.; Ruysschaert, J. M.;

Goormaghtigh, E., Hydrogen-deuterium exchange in membrane proteins monitored by IR spectroscopy: a new tool to resolve protein structure and dynamics. *Biopolymers* **74** (1-2), 19-26 (2004).

31. DeFlores, L. P.; Tokmakoff, A., Water penetration into protein secondary structure revealed by hydrogen-deuterium exchange two-dimensional infrared spectroscopy. *J. Am. Chem. Soc.* **128** (51), 16520-16521 (2006).
32. Brielle, E. S.; Arkin, I. T., Site-Specific Hydrogen Exchange in a Membrane environment analyzed by infrared spectroscopy. *J. Phys. Chem. Lett.* **9** (14), 4059-4065 (2018).
33. Smith, J. R.; Cicerone, M. T.; Meuse, C. W., Measuring hydrogen-deuterium exchange in protein monolayers. *Surf. Interface Anal.* **41** (11), 878-885 (2009).
34. Dunkelberger, E. B.; Woys, A. M.; Zanni, M. T., 2D IR cross peaks reveal hydrogen-deuterium exchange with single residue specificity. *J. Phys. Chem. B.* **117** (49), 15297-15305 (2013).
35. Doukas, A. G.; Pande, A.; Suzuki, T.; Callender, R. H.; Honig, B.; Ottolenghi, M., On the mechanism of hydrogen-deuterium exchange in bacteriorhodopsin. *Biophys. J.* **33** (2), 275-279 (1981).
36. Deng, H.; Huang, L. W.; Callender, R.; Ebrey, T., Evidence for a bound water molecule next to the retinal Schiff base in bacteriorhodopsin and rhodopsin: a resonance Raman study of the Schiff base hydrogen/deuterium exchange. *Biophys. J.* **66** (4), 1129-1136 (1994).
37. Shashilov, V.; Xu, M.; Makarava, N.; Savtchenko, R.; Baskakov, I. V.; Lednev, I. K., Dissecting structure of prion amyloid fibrils by hydrogen-deuterium exchange ultraviolet Raman spectroscopy. *J. Phys. Chem. B.* **116** (27), 7926-7930 (2012).
38. Streitwi.A; Reuben, D. M. E., Acidity of hydrocarbons. XXXV. Equilibrium acidities of phenylacetylene and tert-butylacetylene in cyclohexylamine. *J. Am. Chem. Soc.* **93** (7), 1794-1795 (1971).
39. Yang, Q.; Li, Y.; Yang, J. D.; Liu, Y.; Zhang, L.; Luo, S.; Cheng, J. P., Holistic prediction of the pKa in diverse solvents based on a machine-learning approach. *Angew. Chem. Int. Ed.* **59** (43), 19282-19291 (2020).
40. Chen, Z.; Paley, D. W.; Wei, L.; Weisman, A. L.; Friesner, R. A.; Nuckolls, C.; Min, W., Multicolor live-cell chemical imaging by isotopically edited alkyne vibrational palette. *J. Am. Chem. Soc.* **136** (22), 8027-8033 (2014).
41. Bell, R. A.; Chisholm, M. H., Reactions of alkynyl- and alkenylplatinum(II) compounds. 2. Study of the reaction between anhydrous hydrogen chloride and diethynylbis (dimethylphenylphosphine) platinum(II). Evidence for platinum-stabilized vinyl cations. *Inorg. Chem.* **16** (3), 687-697 (1976).
42. Egoshi, S.; Dodo, K.; Ohgane, K.; Sodeoka, M., Deuteration of terminal alkynes realizes simultaneous live cell Raman imaging of similar alkyne-tagged biomolecules. *Org. Biomol. Chem.* **19** (38), 8232-8236 (2021).
43. Eigen, M., Proton transfer, acid-base catalysis, and enzymatic hydrolysis. *Angew. Chem. Int. Ed.* **3** (1), 1-19 (1964).
44. Englander, S. W.; Downer, N. W.; Teitelbaum, H., Hydrogen-exchange. *Annu. Rev. Biochem.* **41**, 903-924 (1972).

45. Englander, S. W.; Mayne, L.; Bai, Y.; Sosnick, T. R., Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci.* **6** (5), 1101-1109 (1997).
46. Krezel, A.; Bal, W., A formula for correlating pKa values determined in D_2O and H_2O. *J. Inorg. Biochem.* **98** (1), 161-166 (2004).
47. Qu, D.; Wang, G.; Wang, Z.; Zhou, L.; Chi, W.; Cong, S.; Ren, X.; Liang, P.; Zhang, B., 5-Ethynyl-2'-deoxycytidine as a new agent for DNA labeling: detection of proliferating cells. *Anal. Biochem.* **417** (1), 112-121 (2011).
48. Hansch, C.; Leo, A.; Taft, R. W., A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **91** (2), 165-195 (1991).
49. Dempsey, C. E., pH dependence of hydrogen exchange from backbone peptide amides in apamin. *Biochemistry* **25** (13), 3904-3911 (1986).
50. Houliston, R. S.; Liu, C. S.; Singh, L. M. R.; Meiering, E. M., pH and urea dependence of amide hydrogen-deuterium exchange rates in the beta-trefoil protein hisactophilin. *Biochemistry* **41** (4), 1182-1194 (2002).
51. JE, T., Intracellular diffusion of water. *Arch. Biochem. Biophys.* **224** (2), 416-428 (1983).
52. Mills, R., Self-diffusion in normal and heavy water in the range 1 to 450. *J. Phys. Chem. C* **77** (5), 685-688 (1973).
53. Shi, L.; Hu, F.; Min, W., Optical mapping of biological water in single live cells by stimulated Raman excited fluorescence microscopy. *Nat. Comm.* **10** (4764), 1-8 (2019).
54. Zwang, T. J.; Tse, E. C. M.; Barton, J. K., Sensing DNA through DNA charge transport. *ACS Chem. Biol.* **13** (7), 1799-1809 (2018).
55. Maeshima, K.; Imai, R.; Tamura, S.; Nozaki, T., Chromatin as dynamic 10-nm fibers. *Chromosoma* **123** (3), 225-237 (2014).
56. Fantoni, N. Z.; El-Sagheer, A. H.; Brown, T., A Hitchhiker's guide to click-chemistry with nucleic acids. *Chem. Rev.* **121** (12), 7122-7154 (2021).
57. Burley, G. A.; Gierlich, J.; Mofid, M. R.; Nir, H.; Tal, S.; Eichen, Y.; Carell, T., Directed DNA metallization. *J. Am. Chem. Soc.* **128** (5), 1398-1399 (2006).
58. Veneziano, R.; Shepherd, T. R.; Ratanalert, S.; Bellou, L.; Tao, C.; Bathe, M., In vitro synthesis of gene-length single-stranded DNA. *Sci. Rep.* **8** (1), 6548 (2018).
59. Maity, A.; Singh, A.; Singh, N., Differential stability of DNA based on salt concentration. *Eur. Biophys. J.* **46** (1), 33-40 (2017).
60. Baillargeon, P.; Lussier, T.; Dory, Y. L., Hydrogen bonds between acidic protons from alkynes (C–H···O) and amides (N–H···O) and carbonyl oxygen atoms as acceptor partners. *J. Appl. Crystallogr.* **2014**, 1-5 (2014).
61. Park, H.; Zhang, K. J.; Ren, Y. J.; Nadji, S.; Sinha, N.; Taylor, J. S.; Kang, C. H., Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. U.S.A.* **99** (25), 15965-15970 (2002).
62. Nagpal, A.; Dhankhar, D.; Cesario, T. C.; Li, R.; Chen, J.; Rentzepis, P. M., Thymine dissociation and dimer formation: A Raman and synchronous fluorescence spectroscopic study. *Proc. Natl. Acad. Sci. U. S. A.* **118** (6), 1-7 (2021).
63. Rastogi, R. P.; Richa; Kumar, A.; Tyagi, M. B.; Sinha, R. P., Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J. Nucleic Acids* **2010**, 1-32 (2010).

64. Sinha, R. P.; Hader, D. P., UV-induced DNA damage and repair: a review. *Photochem. Photobiol. Sci.* **1** (4), 225-236 (2002).
65. Douki, T.; Court, M.; Sauvaigo, S.; Odin, F.; Cadet, J., Formation of the main UV-induced thymine dimeric lesions within isolated and cellular DNA as measured by high performance liquid chromatography-tandem mass spectrometry. *J. Biol. Chem.* **275** (16), 11678-11685 (2000).
66. DeRosa, M. C.; Sancar, A.; Barton, J. K., Electrically monitoring DNA repair by photolyase. *Proc. Natl. Acad. Sci. U.S.A.* **102** (31), 10788-10792 (2005).
67. Huber, V.; Camisaschi, C.; Berzi, A.; Ferro, S.; Lugini, L.; Triulzi, T.; Tuccitto, A.; Tagliabue, E.; Castelli, C.; Rivoltini, L., Cancer acidity: An ultimate frontier of tumor immune escape and a novel target of immunomodulation. *Semin. Cancer Biol.* **43**, 74-89 (2017).
68. Ozkan, P.; Mutharasan, R., A rapid method for measuring intracellular pH using BCECF-AM. *Biochim. Biophys. Acta* **1572** (1), 143-148 (2002).
69. Bassnett S, R. L., Beebe DC, Intracellular pH measurement using single excitation-dual emission fluorescence ratios. *Am. J. Physiol.* **258** (1), C171-C178 (1990).
70. Paabo, M.; Bates, R. G., Standards for a practical scale of pD in heavy water. *Anal. Chem* **41** (2), 283-285 (1969).
71. Cerqueira, A. V.; Lemos, B., Ribosomal DNA and the nucleolus as keystones of nuclear architecture, organization, and function. *Trends Genet.* **35** (10), 710-723 (2019).
72. Lafontaine, D. L. J.; Riback, J. A.; Bascetin, R.; Brangwynne, C. P., The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell. Biol.* **22** (3), 165-182 (2021).