Live-Cell Imaging and Quantification of PolyQ Aggregates by Stimulated Raman Scattering of Selective Deuterium Labeling

Kun Miao and Lu Wei*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, 91125

*Corresponding author. Email: lwei@caltech.edu

Key words: protein aggregations; neurodegenerative disease; stimulated Raman scattering microscopy; stable isotope labeling; live cells; vibrational spectroscopy

Abstract Huntington's disease, a major neurodegenerative disorder, involves expression and deposition of highly aggregation-prone proteins with a long polyglutamine (polyQ) expansion. However, whether these intracellular aggregations are pathological or protective still remains elusive. Here, we propose stimulated Raman scattering (SRS) imaging of deuterium-labeled glutamine to investigate native polyQ aggregations in live cells with subcellular resolution. Technically, enrichment of deuterated glutamine in polyQ sequence enables sensitive and specific SRS imaging of carbon-deuterium bonds (C-D) from the aggregates without GFP labeling, as typically required for fluorescence microscopy. Biologically, we first reveal that these aggregates become 1.8 times denser without GFP, confirming its perturbation. Second, with SRS quantification, we discover a two-stage aggregation formation model, in which small aggregates are dominated by cytosolic proteins in the first stage, but then prefer to sequester polyQ proteins with size growth. This suggests that small aggregates are cytotoxic while large aggregates play a protective role. Third, we illustrate that polyQ aggregations are not a highly dense solid structure by calculating their absolute protein concentrations to be around millimolar range. Furthermore, hyperspectral SRS (hSRS) mapping of C-D vibrations reveals spectroscopic features corresponding to aggregation formation. This is largely absent in label-free SRS or SRS imaging of other non-glutamine amino acid probes. Applying hSRS to interrogating polyQ aggregations by inducing heat shock responses indicates a conformational intermediate with a changed hydrogen-bonding network. Altogether, our quantitative SRS imaging and spectroscopic analysis strategy offers new structural-functional insights for the molecular role of polyQ aggregates.

Introduction

One hallmark of neurodegenerative disorders is the presence of protein aggregations in the patient brains[1-4]. Among which, polyglutamine (polyQ) diseases, such as Huntington's disease (HD), is one main kind that usually starts with motor symptoms like chorea and followed by memory deficit and depressions[3,4]. The onset of HD has been linked to abnormal CAG trinucleotide repeats that encodes polyQ expansion sequence in mutant Hungtingtin (mHtt) proteins. While Q repeats are typically fewer than 37 in healthy human, they range from 40 to 250 in mHtt proteins and are consistently found in the protein depositions of HD brain slices by immunohistology[1,3]. However, the molecular role of polyQ aggregations for their cytotoxicity remains controversial in the HD disease[2,4-6]. It has recently been suggested that soluble oligomers are cytotoxic by dynamically interacting with functional proteins and triggering apoptosis while aggregations 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 essential functional (e.g. chaperones and transcription factors) and structural (e.g. actin) proteins as well as 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 aggregations. Conventional biochemical assays and recent quantitative proteomics offer relative protein compositions from the aggregates in reference to the soluble pools. However, they 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 typically limited to short expansion length (e.g. K2Q24K2W) due to the difficulty to isolate the peptides with longer Q[14-19]. More importantly, all these *in vitro* studies cannot recapitulate the native aggregation status in live and complex intracellular environment and lack fine spatial information. For live-cell studies, fluorescence imaging offers unprecedented spatial and temporal resolution by genetically-encoded fluorescent proteins[20] or self-labeling tags (e.g. HaloTag)[21] to the C-terminus of a well-established model mHtt exon1 (ex1) sequence with various polyQ lengths (Fig. 1a). The aggregation-prone ex1 fragment, composed of 17 N-terminal amphipathic amino acids (aa), a polyQ tract followed by proline-rich residues at 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 known tendency to oligomerize[23]. This could perturb the ex1 aggregation kinetics and conformations, hence contributing to the controversy of reported toxicity. Moreover, quantification and spectroscopic information of aggregates available from *in vitro* studies is largely missing in fluorescence imaging. It is therefore highly desirable to have a new modality that combines the advantages from both *in vitro* investigations and fluorescence imaging while overcoming their limitations.

Results and Discussion

Proof-of-Principle SRS Imaging of mHtt Aggregations with Glutamine-d_5 (Gln-d_5) Labeling in Live HeLa Cells. Herein, we report a method for live-cell imaging and quantification of polyQ aggregates with minimum perturbation, subcellular resolution and spectral information by stimulated Raman scattering (SRS) microscopy of deuterium-labeled glutamine (Gln) (Fig. 1b, c). Labeling of aggregates is achieved through replacing regular Gln in

the medium with the corresponding Gln-d_5 (Fig. 1c) so that cells would metabolically incorporate and enrich Gln-d_5 into the long polyQ tail of expressed mHtt ex1 proteins (Fig. 1a). We first determined our SRS detection limit on Gln-d_5 solution to be 3 mM (Fig. 1d, when signal (S)/noise (N)=1, targeting C-D vibrational peak at 2147 cm^{-1}). This indicates that our detection limit for mHtt protein is down to 29 μM with a total of 103 Gln enrichment in the widely adopted mHtt-97Q ex1 protein (details in SI). It offers higher detection sensitivity compared to SRS imaging of alkyne tags, which is limited to about 200 μM [24]. In cells, however, the detection limit is set by aggregation-signal to cellular-background ratio (Agg/Cell) because newly synthesized proteome also incorporates Gln serving as background, but with a much lower abundance (4.2 %)[25]. The estimated detectability for mHtt-97Q ex1 is 86 μM when Agg/Cell ratio equals to 1 (details in SI). At this concentration, the SRS image quality set by S/N is high. We note that the intracellular free amino acid Gln pool is about 8 mM[26], which slightly lowers Agg/Cell limited detectability. To minimize this additional background source, we replaced medium with buffer before imaging.

We next validated our SRS imaging by transfecting cells with mHtt-97Q-GFP plasmid (Fig. 1a and S1) and culturing them in Gln-d_5 labeling medium. Fig. 2a confirms that SRS imaging of the aggregates (Fig. 2a, 2167 cm^{-1}) in live HeLa cells co-localizes well with that by confocal fluorescence from GFP. Clear off-resonance image demonstrates high SRS imaging quality (Fig. 2a, 2035 cm^{-1}). The C-D peak-shape change by shifting the highest peak from 2147 cm^{-1} in solution to 2167 cm^{-1} after Gln-d_5 being incorporated into cellular proteins should be due to varied microenvironment (Fig. S2). In contrast, label-free SRS images at CH_3 (2940 cm^{-1}) and amide I (1664 cm^{-1}), typically adopted for imaging total proteins[27], show much decreased detection specificity, particularly for smaller-size aggregations (Fig 2b and 2d, indistinguishable from the nucleoli, arrow indicated). Moreover, the average Agg/Cell ratio of high-contrast C-D images from Gln-d_5 decreases significantly if replaced with non-enriched Leucine-d_{10} (Leu-d_{10}) labeling (Fig. 2c and 2d), confirming that the high imaging specificity comes from the enrichment of Gln-d_5 in polyQ tract.

SRS Imaging and Quantification of Native mHtt-97Q Aggregates without GFP Labeling. After demonstrating the feasibility and the specificity of Gln-d_5 labeling for SRS imaging of mHtt-97Q-GFP aggregates, we aim to image native mHtt-97Q proteins without GFP (Fig. 3a). The mHtt-97Q ex1 has only 152 aa, while GFP is 238 aa long, about twice in size. We observed bright aggregates by SRS imaging at the same C-D frequency (Fig. 3b, CD on and off). Interestingly, these aggregates (Fig. 3b, 2167 cm^{-1}) are significantly brighter compared to those with GFP (Fig. 2a, 2167 cm^{-1}). We reason that since the C-D intensity should remain about the same when chopping off GFP sequence, which only contains 8 Gln, the intensity increase is from the formation of denser aggregates. Similar phenomenon was recently reported by Cryo-ET[11]. To further confirm this claim, amide I image (Fig. 3b, 1664 cm^{-1}) shows similar intensity to that with GFP (Fig. S3, 1664 cm^{-1}), because the increase of mHtt-97Q density makes up for the loss of amide I intensity from GFP. Furthermore, aggregates of mHtt-97Q in CH_3 channel become barely distinguishable from cellular background (Fig. 3b, 2940 cm^{-1}) because previous CH_3 signals for aggregates largely come from the GFP sequence (Fig. S1).

Thanks to linear concentration dependence and the multi-channel SRS ratiometric analysis, we quantitatively examined the compositional changes between mHtt-97Q and mHtt-97Q-GFP

aggregates (Fig. 3c and 3d). First, Agg/Cell ratio from SRS images of Gln-d_5 labeled mHtt-97Q aggregates is 1.8 times higher in C-D channel while marginally decreased in CH_3 than that of mHtt-97Q-GFP (Fig. 3c). These changes mostly originate from 1.8-time denser mHtt aggregations, since cytosolic background should not be influenced by the change of plasmid sequence. Second, we analyzed CH/C-D ratios, representing the ratios between non-Htt proteins (mostly containing non-Q aa) to mHtt proteins (enriched with C-D by Gln-d5 labeling) within the same aggregates (Fig. 3d). Surprisingly, the average is still 5.75 even for mHtt-97Q aggregates (Fig. 3d). Based on the sequence (Fig. S1), it should be only around 0.5 (49 non-Q aa / 103 Q) if aggregations were formed primarily by mHtt-97Q proteins, considering our labeling efficiency by Gln-d_5 is close to 100% (i.e. *de novo* glutamine synthesis is about zero, Fig. S4 and S5) and C-D Raman cross section is the same as that of CH (confirmed by SRS measurement on pure DMSO-d_6 and DMSO). Similarly, the CH/CD ratios should have decreased 5 times (from 289 non-Q aa / 112 Q to 49 non-Q aa / 103 Q). Instead, we only observed 1.5 times decrease. Altogether, this implies that the aggregations contain a rather high percentage of non-mHtt proteins (we calculated it to be around 54% in molar percentage, details in SI). Sequestration of cellular functional and structural proteins has been suggested to be one underlying mechanism of cytotoxicity by aggregation formations in HD disease[9-13]. Our method hence offers a quantitative measurement that corroborate with this mechanism.

Quantifications for the Dependence of Protein Compositions on Aggregation Sizes Reveal A Two-Stage Aggregation Sequestration. To further understand the molecular cytotoxicity of polyQ aggregations, we then interrogated the dependence of proteins compositions (i.e. CH/CD ratios) on aggregation sizes (Fig. 4a). Intriguingly, we observed a negative correlation, demonstrating that the percentage of sequestered mHtt proteins increases as the aggregations grow. By categorizing these aggregations into small (<10 μm^2) and large (>10 μm^2) groups, student t-tests confirmed the significance for the measured CH/CD ratio differences (Fig. 4b). As a comparison, such distinctive correlation is absent with Leu-d_{10} labeling (Fig. 4c-d), underscoring the necessity of specific labeling. Our observations suggest a two-stage sequestration model during mHtt aggregation formation and growth. At the initial stage, the formed small aggregations preferentially sequester non-mHtt cytosolic proteins likely functional chaperones, ribonucleoproteins and structural proteins[9-13], hence are more cytotoxic by depleting them. As the aggregates grow, they transit into a second-stage that would absorb more mHtt proteins. This indicates that large inclusion bodies may indeed play a protective role to clear toxic disease proteins as a cellular rescue mechanism[4,7,8]. Our two-stage model therefore offers a new insight that polyQ aggregations can be both cytotoxic (for small aggregates) and protective (for large aggregates), and the difference depends on sizes.

Inspired by our quantitative analysis, we set out to calculate the absolute concentrations and compositional percentages for mHtt and non-mHtt proteins from small to large aggregates in live cells. This would be highly challenging for existing methods, such as fluorescence imaging, because dense structures lead to self-absorption and fluorescence quenching of fluorophores, making it unreliable for quantifying aggregates. In addition, fluorescent proteins can only label known protein targets. Based on our concentration curve (Fig. 1d), we first determined that the mHtt concentration in a representative small aggregate (Fig. 4a, green dot, 6.7 μm^2) is between 1.1 mM (when all sequestered cytosolic non-mHtt proteins are newly synthesized and hence incorporated with Gln-d_5 contributing to the C-D signal) and 2.1 mM (assuming all sequestered

cytosolic non-mHtt proteins are pre-existing proteins) (Fig. 4e, all details in SI and Table S1). Similarly, a large aggregate (Fig. 4a, red dot, 19.1 μm^2) has mHtt concentrations between 4.2 and 5.1 mM (Fig. 4f). The concentrations continue to increase with bigger ones (Fig. 4a, magenta dot, Table S1, 5.6 - 6.4 mM). Surprisingly, the concentrations of non-mHtt proteins remain almost the same from small to large aggregations (from 4.6 (Fig. 4a, green dot) to 4.7 (Fig. 4a, red dot)) and then gradually decrease with further size increase (Fig. 4a, magenta dot, 4.2 mM). On average, non-mHtt proteins account for 74.5% molar percentage in small aggregates to 50% in large aggregates and to 40.8% in even larger ones (Table S1). Based on these data, polyQ aggregates have total protein concentrations of 5-10 mM and are less dense than we typically expected as solid inclusion bodies. To the best of our knowledge, this is the first that specific protein concentrations and percentages could be given for aggregates of different sizes in live cells.

Understand the Structures of mHtt Aggregates by Hyperspectral SRS (hSRS) Imaging. One key advantage for *in vitro* spectroscopic analysis over live-cell fluorescence imaging is the structural and conformational information of the aggregates[14-19]. For instance, multiple reports suggested beta-sheet rich structures in mHtt aggregations. 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 was well documented to probe protein secondary structures in 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 ~12 cm^{-1}), corresponding to cross-beta sheet structures[28]. Nonetheless, we observed no such spectroscopic difference from Amide I between mHtt aggregates and cellular background (Fig. 5a, 1664 cm^{-1}, blue vs red), suggesting that beta-sheet rich structures might be absent 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), similarly to that suggested before[29], highlighting the importance of performing analysis on live samples.

We next asked whether hSRS analysis on C-D spectral region could offer insightful structural information since Gln-d_5 labeling is more specific on mHtt-97Q proteins than Amide I, which is general to all aa. Interestingly, we observed a clear dip around 2146 cm^{-1} for aggregations compared to that from cellular background (Fig. 5b), implying a structural difference between aggregated and cytosolic proteins. To understand such spectral changes, we compared C-D aggregation spectra (Fig. 5c, green) to those of Gln-d_5 solids (purple) and solutions (red). The C-D spectra after Gln-d_5 being incorporated into cellular proteins show a combinational feature of solution and solid environment. Our previously observed shift of primary peak from 2147 cm^{-1} of Gln-d_5 solutions to 2167 cm^{-1} for mHtt aggregations (Fig. S2) could also be attributed to the formation toward a solid-state structure, consistent with the major peak position of Gln-d_5 solids (Fig. 5c). Our fitting results indicate that the 2146 cm^{-1} dip is resulted from the narrowing and slight red shift of the shoulder peak at 2130 cm^{-1} (Fig. S6), likely due to the formation of a more ordered structure upon aggregation. It also implies that the shoulder peak may serve as a sensitive indicator for the microenvironment of mHtt aggregates. In addition to the amide I and C-D, we also evaluated CH vibrations. CH_3 peaks (2940 cm^{-1}) have no discernable spectral information, since they mostly come sequestered non-mHtt proteins (Fig. 5d). The spectra of aggregates are almost identical to that of nucleoli, enriched with proteins and rRNA and the lower shoulder at around 2850 cm^{-1} compared to that from cell

background is due to CH_2 vibration from cytosolic lipids. It again emphasizes the importance of specific C-D labeling for studying mHtt aggregates.

Investigating the Structural Changes of mHtt Aggregates by Inducing Heat Shock Protein (HSP) Responses. Now that we have established our technique is effective in imaging and analyzing the structures of native aggregates, we then applied it to understanding molecular interactions between polyQ aggregates and HSPs, key chaperones for protein folding and inhibiting protein aggregations. This may offer fundamental insights toward therapeutic development for HD disease[6]. Geldanamycin (GA) is a naturally occurring antitumor drug that has been shown to be active in inhibiting or solubilizing the formation of polyQ aggregations by triggering a heat shock response of HSP40 and 70 in mammalian cells[30]. We first confirmed its function that the number of large aggregates decreases with the treatment of increased GA concentrations (Fig. S7). We next resorted to C-D hSRS on the native mHtt-97Q aggregates without GFP. 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 larger-size aggregates (Figs. S8a and S8b). 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 S8c), 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 upon aggregation formation[31]. We speculated whether such shoulder peak decrease is caused by a reverse phase conversion from solid mHtt aggregates toward surrounding liquid pool with up-regulation of HSP40/70. 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. S8), ruling out our phase-transition hypothesis.

Since C-D vibration is also sensitive to microenvironment, similar to that from CH[32], we speculated such a decrease at 2132 cm^{-1} shoulder peak corresponds to a change of secondary structures when interacting with Hsp40/70 proteins, as suggested by previous *in vitro* EM experiments[33]. To test this, we emerged cells containing mHtt-97Q-GFP aggregates without GA treatment in solvents with descending hydrogen-bonding capacity (Fig. 6c, from DPBS buffer (green), to MeOH (yellow) and DMSO (blue)). The 2132 cm^{-1} shoulder increases accordingly (Fig. 6c, vertical line). Quantitative fittings show a more detailed ratio increase between the shoulder (2130 cm^{-1}) and the major (2169 cm^{-1}) peaks (Fig. S10). The lowest ratio from GA treated subset of aggregations (Fig. S10f) hence demonstrates that its local environment has a more hydrated hydrogen bonding compared to regular cellular proteomes (Fig. S6d). Such hyper-hydration might be caused by a partially folded state, a folding intermediate, by interacting with HSP40/70 in live cells[34].

Conclusion

In summary, we combined SRS microscopy with Gln-d_5 labeling to achieve imaging, quantification and sensitive spectroscopic analysis of polyQ aggregations in HD disease. We demonstrated our SRS utility by imaging mHtt aggregations down to small sizes in live HeLa cells, quantifying the propensity and composition of mHtt aggregates, and performing hSRS

analysis. We revealed that the formation of aggregations become 1.8-time denser than that without GFP tagging, proving its possible perturbation. Our compositional quantifications of mHtt and non-mHtt proteins inside small to large aggregations propose a two-stage aggregation formation model with a surprising sequestration preference of non-mHtt proteins at the initial phase. This may indicate that small aggregates are toxic while large aggregates are protective. Additionally, we induced heat shock responses to inhibit aggregation formation and revealed a vibrational spectral signature, suggesting a possible change of hydrogen-bonding network for mHtt aggregations upon interacting with Hsp40/70.

Technically, our method manifests the following advantages. First, SRS imaging of C-D vibration from Gln-d_5 in the desired cell-silent region (1800-2600 cm^{-1}) allows non-invasive subcellular mapping and spectral analysis of mHtt aggregates in live cells without GFP. Labeling on the stable side chains of Gln does not interfere with internal hydrogen-bond network, which is between NH_2 and amide and is essential for the secondary structure of mHtt aggregates. Second, based on the mHtt ex1 sequence (Figs. 1a and S1), Gln is an ideal labeling target because its enrichment in the extended polyQ region provides both high mHtt-labeling specificity and imaging sensitivity. In mHtt-97Q ex1 fragment, Gln accounts for 68% of the ex1 sequence (Fig. S1), while the overall natural occurrence of Gln is only about 4.2% in human proteomes[25]. We experimentally confirmed close to 100% labeling efficiency by Gln-d_5. Compared to label-free SRS[27], or SRS of ^{13}C-Phenyalanine[35] and deuterated all essential amino acids[36] for imaging proteins, our selective Gln-d_5 labeling is significantly more specific. Unlike alkyne-tagged unnatural amino acids, with which only one tag is introduced to one protein[37], multiple Q labeling (e.g. 103Q for mHtt-97Q ex1) by Gln-d_5 offers much higher sensitivity and less manipulation and invasiveness. Third, compared to Coherent anti-Stokes Raman scattering (CARS)[38], another nonlinear microscopy technique, SRS offers high-fidelity Raman spectra (Fig. 1b) and linear concentration dependence without non-resonance background. Owing to the high-fidelity linear concentration dependence, we were able to quantify aggregation compositions *in situ* from live cells. Fourth, hSRS on the C-D spectral window is suited for investigating the microenvironment of mHtt aggregations.

Our method is generally applicable to other polyQ diseases including spinocerebellar ataxia and spinobulbar muscular atrophy[39] and other poly-aa diseases involves low complexity sequences, such as with poly-GA, poly-GR, poly-GP, poly-PR, and poly-PA[40,41] aggregates recently reported in ALS/FTD patient brain by selective deuteration of corresponding aa. In addition, since Gln can transport across blood-brain barrier and deuterium, as the stable isotope, labeling is highly nontoxic, applications to animal models or even to humans could be possible. Moreover, future correlative imaging of SRS on Gln-d_5 in live cells with recently demonstrated Cryo-ET investigation of down to angstrom-level resolution or together with quantitative proteomics can offer more comprehensive structural-function relationship for native mHtt-97Q aggregations.

Acknowledgments

We thank W. Min, C. Qian, D. Lee and J. Du for helpful discussion. L. Wei acknowledges support of start-up funds from California Institute of Technology.

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. Schneider, R.; Schumacher, M. C.; Mueller, H.; Nand, D.; Klaukien, V.; Heise, H.; Riedel, D.; Wolf, G.; Behrmann, E.; Raunser, S.; et al. Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol*. **412**, 121–136, (2011).
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. 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. Meth.* **11**(4), 410–412, (2014).
25. Kozlowski, L. P. Proteome-pI: proteome isoelectric point database. *Nucleic. Acids. Res.* **45**(Database issue), D1112–D1116, (2017).
26. Piez, K. A.; Eagle, H. The free amino acid pool of cultured human cells. *J. Bol. Chem*. **231**, 533-545, (1958).
27. 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).
28. 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).
29. 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)
30. 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).
31. 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).
32. 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).

33. 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).
34. Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature*. **475**, 324–332, (2011).
35. 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).
36. 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)
37. 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)
38. 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)
39. Orr, H. T.; Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30**, 575–621, (2007).
40. 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)
41. 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).

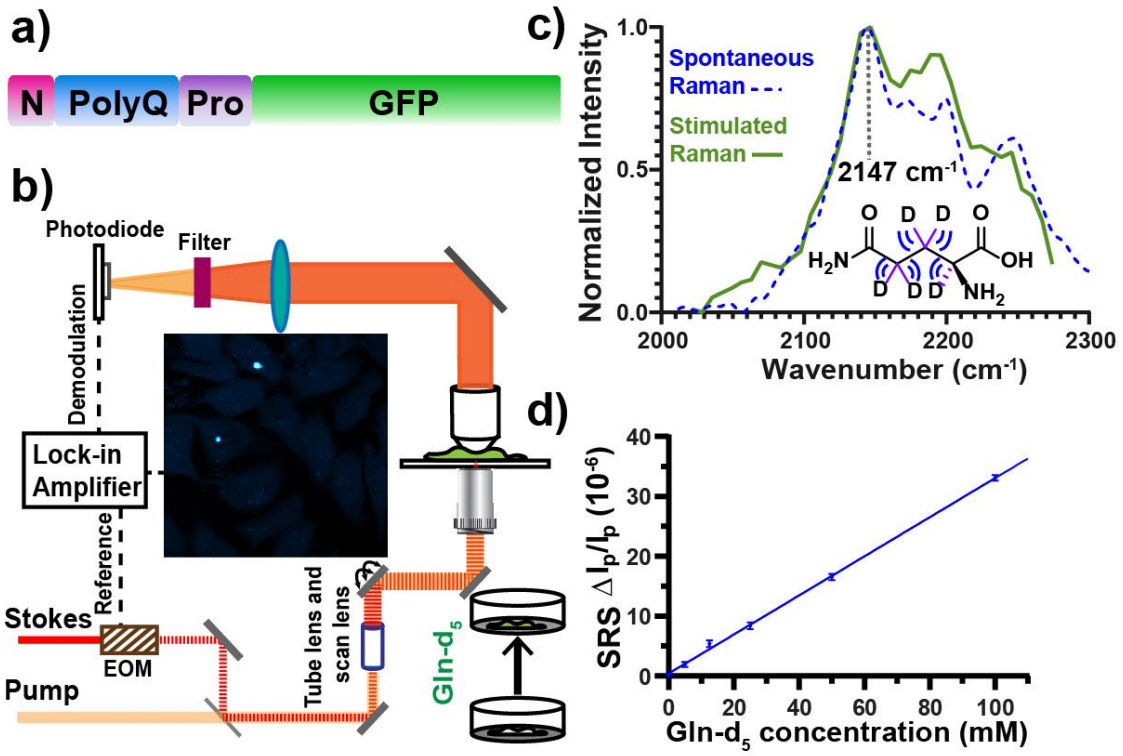

Fig. 1. Experimental scheme for stimulated Raman scattering (SRS) microscopy with deuterated glutamine (Gln or Q) labeling. a) Plasmid construct of a model mutant Huntingtin (mHtt) Exon1 (ex1) protein linked with GFP at C terminus. b) Experimental setup of SRS Microscopy. c) Spontaneous Raman (blue dashed) 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.

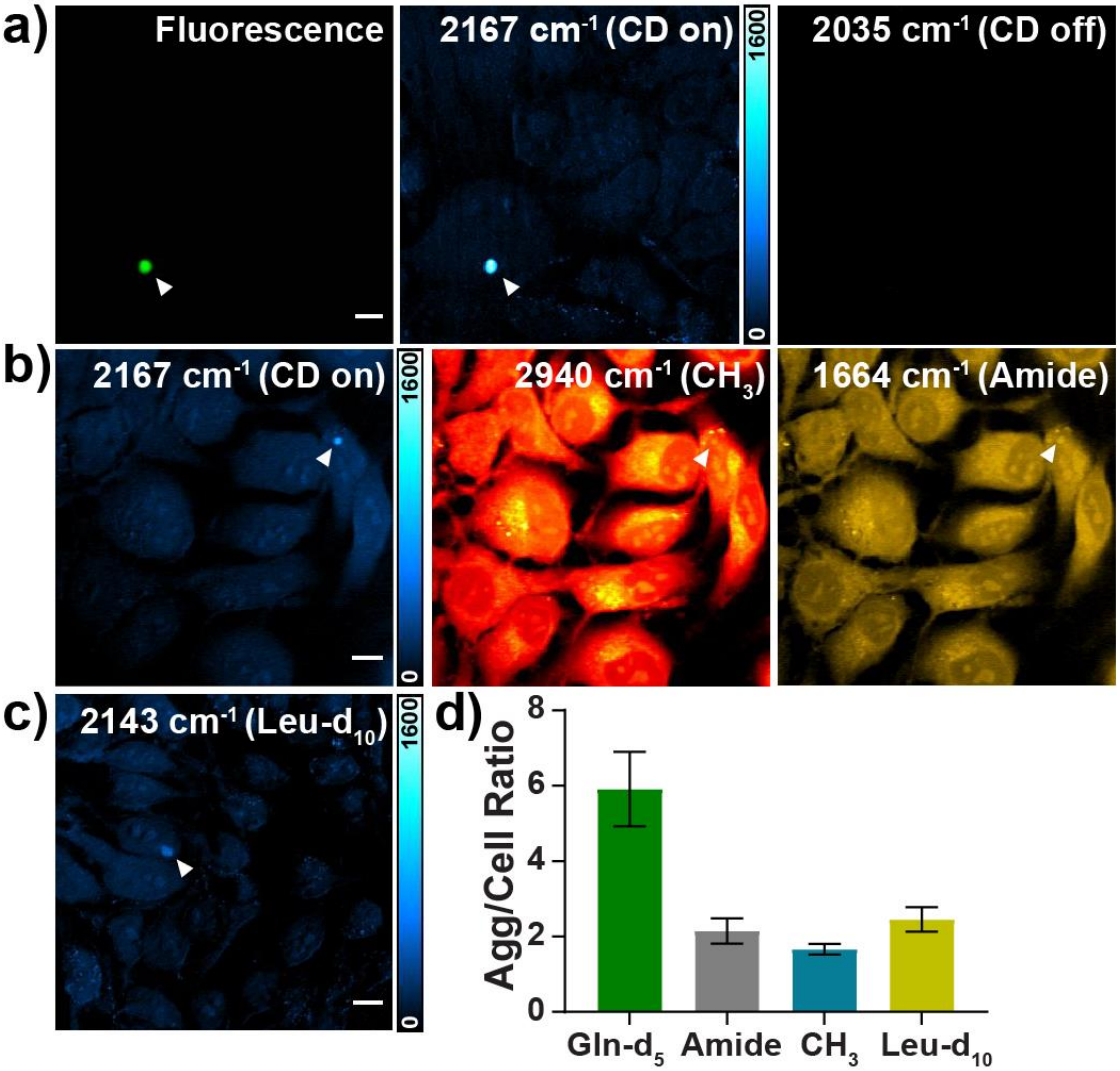

Fig. 2. Live-cell SRS imaging of mHtt-97Q-GFP aggregates with Gln-d_5 labeling. a) SRS imaging of mHtt aggregations (arrowheaded, at 2167 cm^{-1}, C-D on), validated by fluorescence microscopy through GFP (green). Off-resonance image at 2035 cm^{-1} shows no signal. b) Comparison of SRS images for the same set of cells containing a medium-size mHtt-97Q-GFP aggregate (arrowheaded) at Gln-d_5 (2167 cm^{-1}), CH_3 (2940 cm^{-1}) and Amide I (1664 cm^{-1}) channels. c) SRS imaging of an mHtt-97Q-GFP aggregate at 2134 cm^{-1} by Leucine-d_{10} (Leu-d_{10}) labeling. d) Average mHtt Agg/cell ratios from SRS images of Gln-d_5 labeling (5.75±1.03, n=13); amide (2.15±0.34, n=4); CH_3 (1.66±0.14, n=10) and Leu-d_{10} labeling (2.45±0.33, n=10.) Error bar: SD.

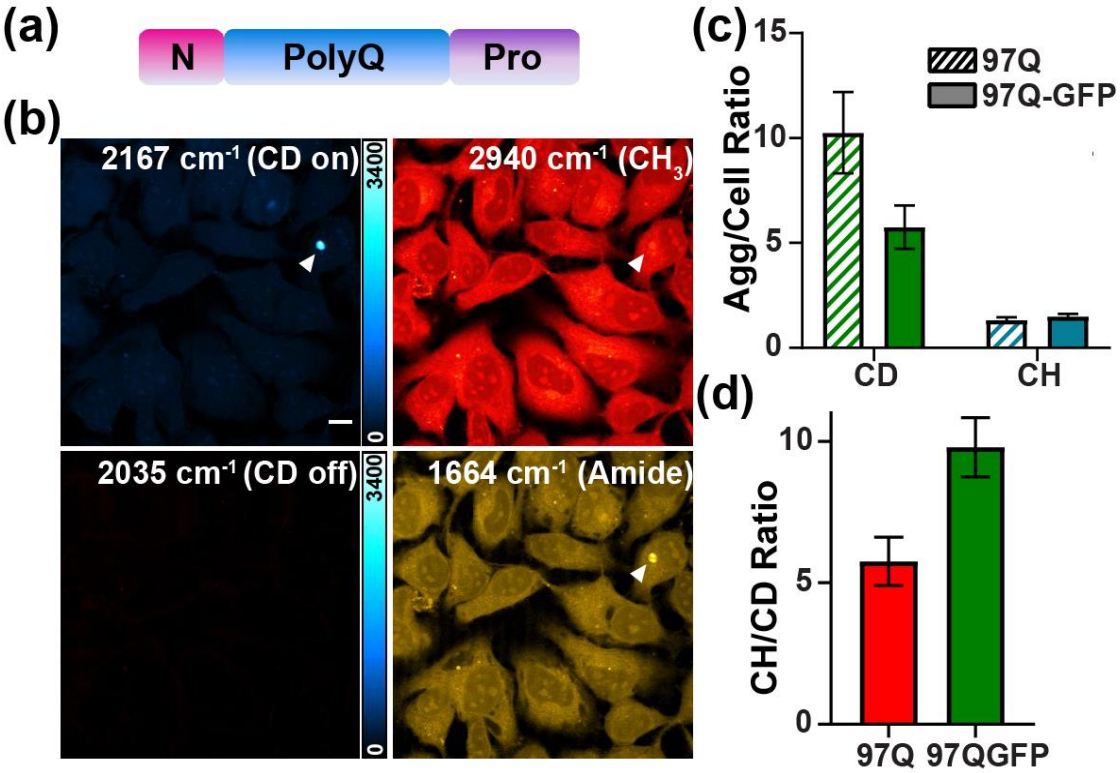

Fig. 3. SRS imaging of mHtt-97Q protein aggregates without GFP. a) Plasmid construct of mHtt-97Q by chopping off GFP sequence from Fig. 1a. b) SRS images of aggregations with Gln-d_5 labeling at 2167 cm^{-1} (C-D on-resonance), 2035 cm^{-1} (C-D off-resonance), 2940 cm^{-1} (CH_3) and 1664 cm^{-1} (Amide I). Scale bar: 10 µm. c) Agg/Cell ratios of mHtt-97Q in C-D (10.14±1.99, n=11) and CH_3 (1.50±0.12, n=10) channels compared to that for mHtt-97Q-GFP (C-D: 5.75±1.03, n=13; CH_3: 1.66±0.14, n=10). d) SRS intensity ratios between CH_3 and C-D channels (CH/CD ratio) on the same aggregates for mHtt-97Q (5.75±0.86, n=20) and mHtt-97Q-GFP (9.79±1.05, n=7). Error bar: SD.

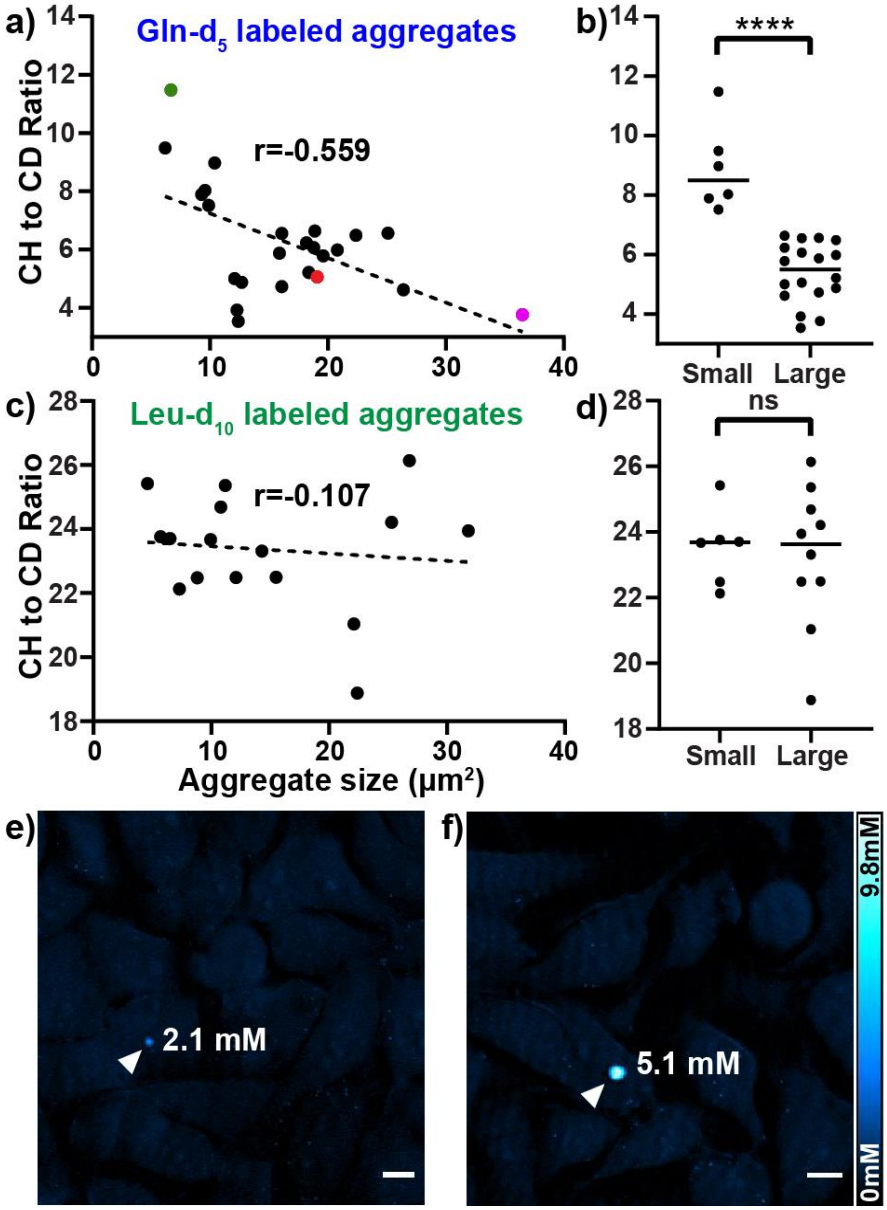

Fig. 4. SRS quantifications of mHtt-97Q aggregates with different sizes. a) CH/CD ratios in Gln-d_5 labeled aggregates decrease with the increase of aggregation sizes (The Pearson's coefficient r=-0.559). b) Student t-test is significant when grouping the aggregates in (a) into small (< 10 μm^2) and large (> 10 μm^2) categories. ****p<0.0001. c-d) Minimum correlation between CH/CD ratios and aggregation sizes in Leu-d_{10} labeled aggregates, indicated by both the Pearson's coefficient (r=-0.107) (c) and the Student t-test between small and large aggregation groups (d). e-f) Representative small (e, 6.7 μm^2) and large (f, 19.1 μm^2) Gln-d_5 labeled mHtt-97Q aggregates indicated with calculated mHtt-97Q protein concentrations. Scale bar: 10 μm.

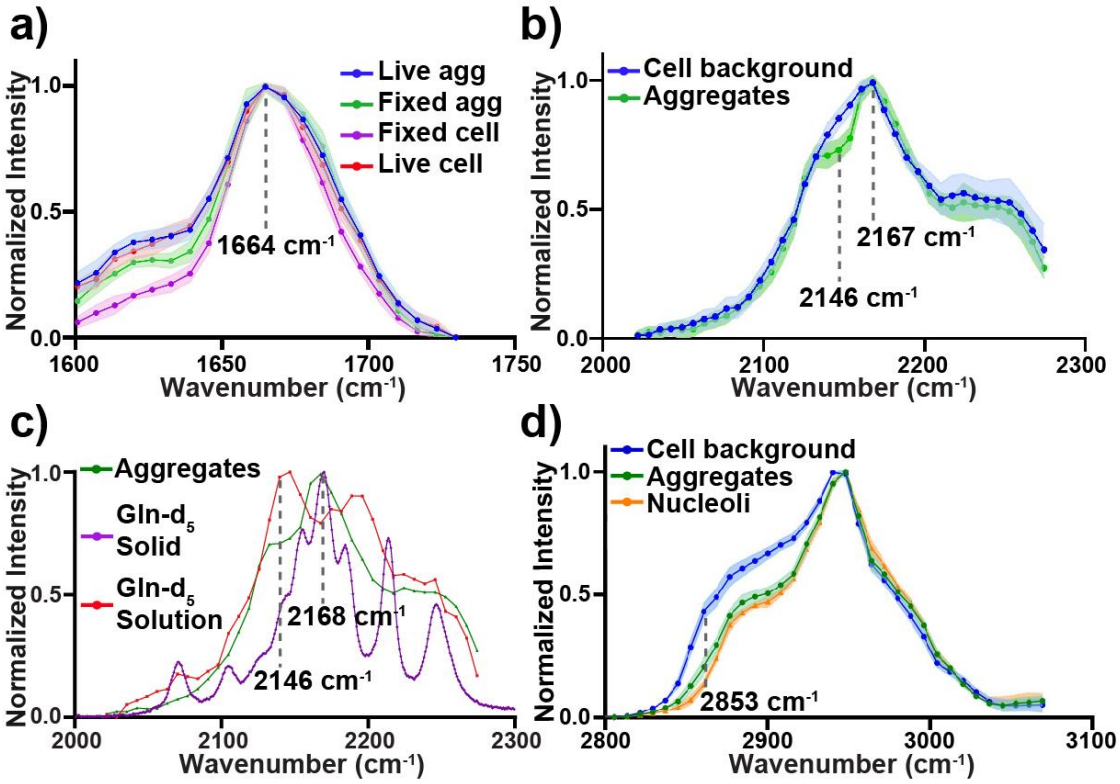

Fig. 5. Hyperspectral SRS analysis for mHtt-97Q aggregates. a) Amide I SRS spectra of live aggregates (blue), live cell-background (red), fixed aggregates (green), and fixed cell-background (magenta). b) Live-cell SRS C-D spectra of Gln-d_5 labeled aggregates (green) and cell background (blue). c) SRS C-D spectra of Gln-d_5 labeled aggregates (green) overlaid with that for Gln-d_5 solids (purple) and solutions (red). d) Live-cell SRS CH_3 spectra of aggregates (green), cell background (blue), and nucleoli (orange). Error bar: SD.

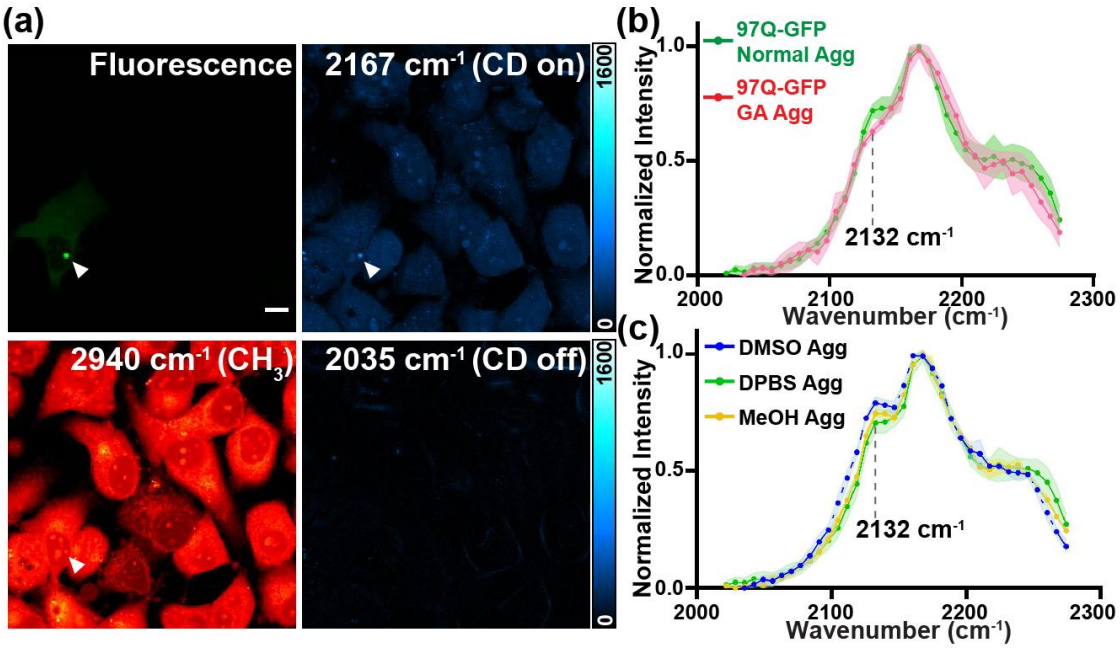

Fig. 6. Hyperspectral SRS study of interactions between mHtt aggregations and Heat Shock Proteins (HSPs). a) Fluorescence (green) and correlative SRS images for mHtt-97Q-GFP aggregates (white arrowheaded) at 2167 cm^{-1} (C-D on-resonance), 2035 cm^{-1} (off-resonance) and 2940 cm^{-1} (CH_3) in live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