Published in final edited form as:

Anal Chem. 2019 March 19; 91(6): 3810–3817. doi:10.1021/acs.analchem.8b03454.

Genetically Encoded Fluorescent Proteins Enable High-Throughput Assignment of Cell-cohorts Directly from MALDI-MS Images

Nicholas D. Schmitt^{†,‡}, Catherine M. Rawlins^{†,‡}, Elizabeth C. Randall[§], Xianzhe Wang[†], Antonius Koller[†], Jared R. Auclair^{†,||}, Jane-Marie Kowalski[∇], Paul J. Kowalski[∇], Ed Luther[○], Alexander R. Ivanov[†], Nathalie Y.R. Agar^{§,¶}, Jeffrey N. Agar^{*,†,○}

[†]Department of Chemistry and Chemical Biology, and Barnett Institute of Chemical and Biological Analysis, Northeastern University, Boston, MA, 02115, USA

[‡]These authors contributed equally to this work.

[§]Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02115, USA

^{||}Biopharmaceutical Analysis Training Laboratory (BATL), Northeastern University Innovation Campus, Burlington, MA, 01803, USA

[∇]Bruker Daltonics, 40 Manning Road, Billerica, MA, 01821, USA

[○]Department of Pharmaceutical Sciences, Northeastern University, Boston, MA, 02115, USA

[¶]Department of Neurosurgery, Brigham and Women's Hospital, Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, 02115, USA

Abstract

Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry imaging (MSI) provides a unique *in situ* chemical profile that can include drugs, nucleic acids, metabolites, lipids, and proteins. MSI of individual cells (of a known cell-type) affords a unique insight into normal and disease-related processes and is a prerequisite for combining the results of MSI and other single-cell modalities (*e.g.* mass cytometry and next-generation sequencing). Technological barriers have prevented the high-throughput assignment of MSI spectra from solid tissue preparations to their cell-type. These barriers include obtaining a suitable cell-identifying image (*e.g.* immunohistochemistry) and obtaining sufficiently accurate registration of the cell-identifying and MALDI-MS images. This study introduces a technique that overcame these barriers by assigning cell-type directly from mass spectra. We hypothesized that in MSI from mice with a defined

***Corresponding Author:** Address: 140 TF, Room 446, 360 Huntington Ave, Boston, MA 02115, United States. j.agar@northeastern.edu. Tel: +1 617-373-5909.

AUTHOR CONTRIBUTIONS

C.M.R. and J.N.A. designed experiments with input from J-M.K., P.J.K., E.C.R., and N.Y.R.A. Experiments were performed as follows: MALDI-MSI and purified YFP MS (C.M.R.); LSC analysis (E.L.); LC-MS/MS proteomics (X.W., A.K., A.R.I.); MSI and MS data analysis (N.D.S.). J.N.A., N.D.S., and C.M.R. interpreted the results and wrote the manuscript in consultation with all co-authors.

Supporting Information

Figures S-1 – S-9 and Table S-1 as described in the text are included in the Supplementary Information available on the ACS Publications website.

fluorescent protein expression pattern, the fluorescent protein's molecular ion could be used to identify cell cohorts. A method was developed for the purification of enhanced yellow fluorescent protein (EYFP) from mice. To determine EYFP's molecular mass for MSI studies, we performed intact mass analysis and characterized the protein's primary structure and post-translational modifications through various techniques. MALDI-MSI methods were developed to enhance the detection of EYFP *in situ*, and by extracting EYFP's molecular ion from MALDI-MS images, automated, whole-image assignment of cell-cohorts was achieved. This method was validated using a well-characterized mouse line that expresses EYFP in motor and sensory neurons and should be applicable to hundreds of commercially available mice (and other animal) strains comprising a multitude of cell-specific fluorescent labels.

Recent technological breakthroughs enable increasingly detailed analysis of individual cells.^{1,2} In particular, the analysis of cell types that are amenable to cell-sorting (*e.g.* tissue dissociated, immune, and circulating tumor cells) is being revolutionized by RNA sequencing (RNA Seq),³ liquid chromatography-tandem mass spectrometry (LC-MS/MS),⁴ and mass cytometry-based proteomics.⁵ Unfortunately, cells within solid tissues—especially cells with fragile projections (*e.g.* neurons)—are not amenable to high-throughput cell sorting. In addition, many normal and pathological processes depend upon cell-cell interactions, necessitating *in situ* analysis.⁶ For example, the death of motor neurons leads to paralysis, the predominant phenotype of Amyotrophic Lateral Sclerosis (ALS). Interactions of motor neurons with microglia, Schwann cells, and astrocytes modulate the age of onset and severity (rate of progression) of ALS.⁷ In ALS and other diseases of solid tissues, single-cell protein analysis has traditionally been limited to the targeted analyses of a few proteins using immunohistochemistry. This study addresses the unmet need for the unbiased, *in situ* proteomics characterization of specific cell-types, including motor neurons.

Single-cell mass spectrometry methods, including secondary-ion mass spectrometry (SIMS), imaging mass cytometry (IMC), and matrix assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) have been recently reviewed.^{8,9} Due to their unique strengths and weaknesses, SIMS, IMC, and MALDI-MSI are considered complementary techniques.^{10,11} IMC is currently the only technique that can analyze the proteome of single cells to sufficient depth—as many as 36 proteins simultaneously—to potentially identify cell-types *in situ*.¹² IMC does not suffer from the low-mass bias of other MSI techniques, affording superior sensitivity for large proteins. However, IMC is a targeted technique with many of the drawbacks of immunohistochemistry (IHC), including the requirements for antigen presentation and validated antibodies.

SIMS is the most prolific method for single-cell MSI^{13,14} and can produce multiplexed elemental (multiple stable-isotopes) and molecular images (> 20 molecules) at subcellular resolution. A unique characteristic of SIMS is its focused ion beam, which can permit imaging of biomolecules in 3D, notably at a depth resolution as high as 5 nm.¹⁵ Nano (magnetic sector) SIMS can probe the location and half-lives of proteins,¹⁶ lipids,¹⁷ and neurotransmitters at lateral spatial resolutions of 50–100 nm.¹⁸ TOF-SIMS is employed for many single-cell analyses, including imaging of vitamin,¹⁹ metabolite,²⁰ and lipid^{20,21} molecular ions at lateral spatial resolutions as high as 2 μm .²⁰ However, SIMS techniques

generally cannot detect the intact molecular ions of endogenous peptides and proteins, or other masses greater than 1,000 Da. Such molecular ions (detected by MALDI-TOF MS) have proven diagnostic utility, serving as the basis for the clinical identification of bacterial species (with regulatory approval from the US FDA, EMA, *etc.*).²²

MALDI-MSI can detect hundreds of chemically diverse molecules per image including drugs, metabolites,²³ nucleic acids,²⁴ lipids²⁵, peptides,²⁶ and proteins.²⁷ A few of the many applications of MALDI-MSI include: measuring the infiltration of pesticides into plant tissues,²⁸ identifying the substances of abuse within a single human hair,²⁹ characterizing bacterial biofilms,³⁰ augmenting the traditional tools used by pathologists to classify diseased tissue,^{31,32} and analyzing drug distribution^{33,34} in pre-clinical and clinical studies.³⁵ MALDI-MSI techniques, including advances in spatial resolution, have been reviewed.^{36,37} Laborious sample preparation techniques, such as laser microdissection,³⁸ tissue stretching,³⁹ and microinjection of matrix,⁴⁰ enable single cell MALDI-TOF MS. Innovations including lasers with increased repetition rates, improved laser optics, laser oversampling,⁴¹ and automated matrix deposition enable high-throughput imaging at ~5–10 μ m spatial resolution using commercial MALDI-MSI instrumentation, and *c.a.* 2.5 μ m with custom-built instrumentation^{27,42}.

Assigning a mass spectrum to the cell it originated from remains a major challenge. In one technique, cells are dissociated from tissue, and unsupervised classification is used to cluster spectra and infer several cell-types.⁴³ The goal of this study, however, is assigning mass spectra derived from *intact* tissue sections to their cell type. Previous studies approached this goal by the registration of MSI with a secondary cell-identifying image, but encountered several limiting factors. Obtaining a cell-identifying image can be difficult: histology stains such as hematoxylin and eosin (H&E) are not cell-selective; immunohistochemistry can be cell-selective, but requires validated antibodies; and only a limited number of cell types can currently be classified using spectroscopic methods such as FTIR imaging⁴⁴. In the event a cell identifying image is obtained, it cannot be registered to the MS image with cellular-scale accuracy.

As we demonstrate below, typical prospective MALDI-MSI methods result in registration errors of 200 μ m or more. Improvement of the size and shape of the fiducial markers, the number of fiducials (*c.a.* 20), and registration (probabilistic non-rigid transformation),⁴³ have been implemented in fit-for-purpose software.⁴⁵ These innovations improved expert's average registration accuracy from 164 μ m to 38 μ m (range of registration errors *c.a.* 10–140 μ m), permitting targeted studies of dispersed, individual cells.⁴³ McDonnell and colleagues developed an automated approach that can register MSI to histology for a variety of tissue types and instrument platforms, which reduces retrospective registration errors to 40–80 μ m.⁴⁶ Previous techniques, however, are not sufficient for identifying the densely populated cells in thin tissue sections.

The *in situ* MALDI-MS analysis of known mammalian cells is currently limited by the requirements of cell-identifying images and image registration. Here, we present a technique with potential to forgo both of these requirements. A fluorescent protein with cell-cohort specific expression is used as a “mass marker” to identify cell types directly from MALDI-

MS images. Prior to the present study the mass of our intended mass marker, enhanced YFP (EYFP), was uncharacterized. We purified EYFP from mouse brains; characterized its mass and primary structure using MS; optimized the *in situ* MALDI detection of EYFP; and used EYFP detection as a proxy for neuron-type in MSI. The strains of mice used here to label specific cells types, and hundreds of others like them, are commercially available through Jackson Laboratories.

EXPERIMENTAL SECTION

Chemicals.

Sinapic acid (SA) matrix, α -cyano-4-hydroxycinnamic acid (CHCA) matrix, HPLC grade acetonitrile (ACN), HPLC grade water + 0.1% formic acid (FA), and HPLC grade water + 0.1% trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC grade ethanol was purchased from Fisher Scientific (Hampton, NH, USA).

Vertebrate Animal Subjects.

Transgenic “thy-1 YFP-16 mice” (B6. Cg-Tg(Thy1-YFP)16Jrs/J, Stock No. 003709) (The Jackson Laboratory, Bar Harbor, ME, USA) were housed in groups at the Northeastern University Animal Care Facility. The use of these animals was approved by the Animal Care and Use Committee at Northeastern University (IACUC protocol #16-0303R) and was in accordance with the federal, local, and institutional guidelines.

Ammonium Sulfate Fractionation.

An LS50B Luminescence Spectrometer (Perkin Elmer, Waltham, MA, USA) was used to monitor EYFP at 485 nm excitation and 527 nm emission, and a dilution series of recombinant YFP standard (BioVision Incorporated, Milpitas, CA, USA) was used to determine the approximate concentration of the EYFP. For purification experiments, one EYFP-expressing mouse brain was homogenized with a Dyna-Mix homogenizer (Fisher Scientific, Hampton, NH, USA) in 10 mM Tris buffer (pH 8.0) in ten 30 second intervals on ice. This homogenate was centrifuged at 13,000 g for 10 minutes at 4 °C and the supernatant was found to contain EYFP fluorescence. Following optimization experiments at a variety of ammonium sulfate (AS) concentrations, fractionation using 20% w/v AS (centrifugation for 10 minutes, 4 °C at 13,000 g) was determined to result in optimal EYFP enrichment.

Fast Protein Liquid Chromatography (FPLC) Separation of EYFP.

The 20% AS supernatant was solvent exchanged into water, then purified on a MonoQ 10/100 GL anion exchange column (GE Healthcare, Piscataway, NJ, USA) using an ÄTKA FPLC system with INV-907 Valve System. The gradient went from 0–100% buffer B with 10 mM Tris, pH 8.0 (A) and 10 mM Tris with 1 M NaCl, pH 8.0 (B) buffers collecting 5 mL fractions. 41 fractions were collected, concentrated to 1 mL, and cleaned via ultrafiltration as described in detail on page S3 in the Supplementary Information.

LC-ESI-MS Intact Protein Analysis.

Using standard procedures (see page S3), fractions with YFP fluorescence were analyzed on a H-Class Acquity UPLC system coupled to a Xevo G2-S Q-ToF mass spectrometer (Waters Corp, Milford, MA) using an Acquity UPLC Protein BEH C4 (300 Å pore size, 1.7 µm particle size, 2.1 mm ID × 100 mm) column (Waters Corp, Milford, MA).

Top-Down EYFP Characterization.

Top-down protein characterization was performed on a Bruker 9.4 T solariX XR Mass Spectrometer (Bruker, Billerica, MA) using ESI-FSD. Intact EYFP was infused at 3 µL/min in 50:50 acetonitrile:water with 0.1% formic acid and fragmented in the region between skimmer 1 and funnel 2 with a declustering potential of 132 V. Ions were accumulated for 0.085 seconds per scan and transferred to a paracell. Forty 1.2 second transients were summed at 2 megawords per scan. N-terminal sequence and modifications were determined using DataAnalysis and BioTools software from Bruker with an RMS error of 2.8 ppm at 1,000 m/z.

Matrix Deposition and Protein Extraction.

Three different methods of matrix deposition were tested with sinapic acid (SA). For the TM Sprayer (HTX Technologies, Carrboro, NC), SA was deposited at 20 mg/mL in 50:50 HPLC grade ACN:HPLC grade water + 0.2% TFA. The nozzle temperature was set at 30 °C with 50:50 ACN:HPLC grade water + 0.2% TFA as the pushing solvent (flow rate 0.200 mL/min). The sprayer velocity was set at 900 mm/min with 2 passes, 10 psi gas flow, resulting in a density of 0.0044 mg/mm². After coating, the slide was rehydrated at 80 °C for 3.5 minutes with 1 mL of 5% acetic acid.

MALDI-MSI.

Imaging experiments were conducted using an ultrafleXtreme MALDI TOF/TOF with a 1 kHz Smartbeam laser (Bruker Daltonics, Billerica, MA, USA). Intact proteins were analyzed from 4,000 – 40,000 *m/z* with the global attenuator offset at 25%, and the lens adjusted to 7 kV. The digitizer was set to 0.08 Gs/s and the pulsed ion extraction was set to 50 ns. MSI was collected at 25 and 50 µm spatial resolution at 30 shots/spot, 1000 Hz, on medium energy distribution set to 75% laser power with hexagon measuring raster. Analysis of MALDI-MSI data was conducted in flexImaging, and flexAnalysis (Bruker Daltonics, Billerica, MA, USA) and SCiLS 2016b (Bremen, Germany). These are described in further detail on page S4 in the Supplementary Information.

RESULTS AND DISCUSSION

Most published MALDI-MSI studies employ Bruker's software, which limits the number of fiducials to three and employs a rigid image registration. Considering the work of Sweedler and colleagues⁴⁵, these conditions extrapolate to a maximum (expert) registration accuracy of 185 µm.⁴³ This accuracy is insufficient for assigning mass spectra to cell-defining features and, in many cases, to an anatomical region (*e.g.* distinguishing the adjacent, functionally distinct cerebellum and locus coeruleus). Even when restricted to a single point pair (*i.e.* simulating the image-guided analysis of one cell), cellular-resolution registration

accuracy could not be achieved using instrument manufacturer's software (Figure S-1). An additional systematic registration error is introduced by laser misalignment, which for example, was $349 \pm 19 \mu\text{m}$ and $265 \pm 17 \mu\text{m}$ in the x and y axes respectively in our as-installed MALDI-FTICR source. Our previous studies worked around these limitations by the manual intensity-based (retrospective) registration of small areas (e.g. $200 \mu\text{m}^2$) or by microinjection of individual fluorescent neurons with MALDI matrix.^{37,40} The aftermarket software microMS⁴⁵ enables additional fiducials (20 or more) and fluorescent cells to be targeted for MALDI-MS acquisition, provided the cells are sufficiently dispersed (*i.e.* are not part of a thin tissue section). We reasoned the sources of error in registration, the laborious workarounds, and even the need for an optical cell-identifying image, could all be avoided by using the *in situ* mass of EYFP as a cell-specific mass marker.

EYFP as a Tool to Define Neuron Cohorts.

The objective of this study is to identify the spectra within a MALDI-MS image that are derived from a specific neuronal cohort, using the detection of a fluorescent protein's molecular ion. Fluorescent proteins derived from the green fluorescent protein (GFP) of *Aequorea victoria*,⁴⁷ are well-established tools for *in vivo* labeling.^{48,49} Sanes and colleagues established lines of mice with red, green, yellow, and cyan fluorescent proteins (together termed XFP) expressed in distinct neuronal cohorts, under control of the neuron-selective Thy1 promoter. The expression pattern of each of these lines, which has been described in detail,⁵⁰ covers many parts of the peripheral and central nervous system, and favors motor and sensory neurons.

Our interest in motor neuron physiology led to the use of two mouse lines expressing enhanced YFP (EYFP), a variant of GFP^{47,51} that includes the fluorescent-yield enhancing T203Y mutation. 'Thy-1 YFP-16' mice express EYFP in motor, sensory, and a subset of central neurons in cortical layers 2–6, along their axons. 'Thy-1 YFP-H' mice express EYFP in the cells that are selectively vulnerable to degeneration in ALS, layer 5 motor neurons (Figure 1).⁵² Laser scanning cytometry (LSC), which quantifies fluorescence throughout an entire tissue section (*i.e.* signal is not attenuated by depth in thin tissue sections), was used to identify all cells in a Hoechst stained tissue section.⁵³ Consistent with numerous qualitative histopathology studies, quantitative LSC images confirm that other cell-types vastly outnumber neurons, and that neuron-specific chemical profiles would be diluted by glial cell chemical profiles at spatial resolutions above *c.a.* $25 \mu\text{m}$.

Purification of EYFP Proteoforms.

The DNA sequence of the XFP variants, including the EYFP variants used to create the YFP-H/16 mouse lines used here, were not included in the original manuscripts and could not be obtained. Regardless, the DNA sequence would not have accounted for post-transcriptional or post-translational processing. Therefore, the purification of EYFP from Thy-1 YFP-H mice was undertaken to enable purified protein characterization. EYFP fractionation was followed throughout the purification process using fluorescence spectroscopy. Hydrophobic interaction chromatography (HIC) was used previously for YFP and was therefore attempted.⁵⁴ However, with mouse brain preparations, HIC exhibited insufficient retention of EYFP, and we therefore developed an alternative purification

method. This method is described in detail above in the experimental section. Briefly, high resolution (4% w/v intervals) ammonium sulfate fractionation (Figure 2A) was followed by anion exchange chromatography (AEX) (Figure 2B). Special care was taken to avoid low pH because at pH 6.5 (the isoelectric point of YFP) 50% of total fluorescence activity is lost.^{55,56} Additionally, increasing the pH of the buffers from 6.5 to 7 resulted in improved AEX retention and fractionation. EYFP proteoforms were present in five of the 41 AEX fractions. Following buffer-exchange, these fractions were analyzed by LC-ESI-MS to determine the intact mass, and two proteoforms were detected at 26,882 and 26,754 Da (Figure 2C). This intact mass difference and peptide mass fingerprinting (PMF, see below) indicate that the less abundant proteoform (*c.a.* 25% of total YFP by MS or fluorescence intensity) results from the loss of the C-terminal lysine from the more abundant proteoform, likely during purification by carboxypeptidase activity. Only the more abundant proteoform was detected *in situ* (see below).

Characterization of EYFP.

To date, the purification and fluorometric analysis of YFP has been from bioluminescent bacteria *Vibrio fischeri*.⁵⁷ GFP from *A. victoria* and several of its variants have been characterized by mass spectrometry,⁵⁸ but EYFP had not. In addition, there are multiple gene products referred to as EYFP as well as sequence ambiguity within certain gene products. Employing various mass spectrometry methods, we determined that the EYFP present in these mice contained the following amino acid insertions and substitutions: M1_S2insV, S65G, V68L, S72A, T203Y, and H231L as compared to GFP (Figure 3), consistent with a published EYFP DNA sequence.^{51,56} Primary structure and modification state of EYFP were determined through a combination of top-down protein characterization, peptide mass fingerprinting, and LC-MS/MS peptide analysis (described in supporting information, page S4) of the EYFP expressed in the studied mouse neurons. The theoretical (26,882.1Da) and experimentally-determined (26,881.7 Da, Waters Q-ToF) average masses of the proposed EYFP proteoforms differed by 15 ppm.

Tissue Washing to Maintain Fluorescence and Minimize Protein Delocalization.

Tissue washing is an essential step for protein MALDI-MSI sample preparation. Washing removes lipids, salts and other interfering substances that can suppress protein ionization allowing the matrix crystals to coalesce with the proteins *in situ*.⁵⁹ We determined that commonly applied washing procedures, such as Carnoy's wash, as well as the "6-step" wash developed by Yang *et. al.*,⁶⁰ removed all traces of EYFP fluorescence. Experiments with mixtures of washing solvents indicated optimal wash conditions for maintaining EYFP localization and fluorescence as 70% ethanol (twice for 30 seconds each) followed by 30 seconds of 90% ethanol (Figure S-2).

Optimization for High Mass Protein Detection.

The sensitivity of MALDI-MS analysis decreases in proportion to protein size.⁶¹ New detectors, such as the CovalX high mass HM1 detector, are being developed to address this need.⁶² Technical challenges associated with imaging of intact proteins of high mass are significant enough that many studies resort to on-tissue digestion.²⁶ For proper desorption/ionization, proteins need to be sufficiently extracted from the tissue microenvironment to

interact with the matrix, without becoming delocalized.⁶³ The major factors known to affect protein detection *in situ*⁶⁴ were explored in preliminary experiments and critical parameters are discussed. These included: detergent enhancement of signal (Figure S-3),⁶⁵ temperature (Figure S-4), the ratio of aqueous to organic solvents in the matrix solution (*i.e.* wetness), matrix layer thickness and homogeneity, solvent saturation of the deposition chamber, automated matrix deposition methods, and droplet size (Figure S-5). ACN concentration was determined to be critical for manual and automated deposition techniques (Figures S-6 and S-7).^{59,66} A subsequent recrystallization or “rehydration” step⁶⁷ was also found to improve signal-to-noise. Optimized matrix deposition parameters for the ImagePrep and TM Sprayers result in comparable protein spectra, including EYFP detection (Figure 4). In control experiments with non-transgenic animals the signal identified as EYFP was not observed.

Mass Spectrometry Imaging of EYFP.

Following the optimization described above, using the TM Sprayer to deposit matrix (described in detail in the Experimental section above and the Supplemental section) and the Bruker ultrafleXtreme MALDI TOF/TOF to acquire MSI, detection of EYFP was achieved at both 50 and 25 μm spatial resolution (Figure 5). In addition to EYFP, other proteins such as Purkinje cell protein 4 and myelin basic protein (and many others not depicted) were localized with MSI in relevant brain regions (Figure 5). To improve mass accuracy we performed external calibration using a high mass protein calibrant pre-acquisition, and internal calibration using known mouse brain proteins post-acquisition (Table S-1). To account for solvent-induced tissue deformation, a post-wash fluorescence microscopy image of the ITO slide was obtained prior to MSI.

The localization of fluorescent cell bodies (neuronal soma) and the EYFP molecular ion are in general agreement (Figure 5). Further analysis (*e.g.* changing laser spot size and raster area) indicated that at the spatial resolutions required for single-neuron analysis, this technique is sample-limited and approached the detection limit of current instrumentation. One shortcoming of our method was that the EYFP molecular ion was not detected as well in white matter (*e.g.* corpus callosum and striatum) as it is detected in gray matter. This is evident in the MS images and is depicted in the segmentation map produced via bisecting *k*-means clustering of the MALDI-MSI data (Figure S-8). Attempts to improve EYFP detection in white matter were successful, but required wet-deposition or rehydration that resulted in delocalization of smaller (*e.g.* Ub) proteins.

The differential EYFP intensity between gray and white matter was not explored in detail in the present study, which focused on optimizing signal from cortical neurons (gray matter). Further examination of this would address whether EYFP is being preferentially extracted from lipid-rich white matter during the delipidating tissue wash, or if EYFP signal is suppressed by components of white matter. Despite these shortcomings, the EYFP molecular ion could be detected in thousands of neurons per brain section, and this signal generally correlated with a fluorescent foci (Figure 5). To determine whether EYFP-expressing cells happen to express a protein with a mass similar to EYFP, which could be mis-identified in MS images, EYFP-expressing cells were isolated and their proteome was elucidated. From a

thin tissue section, within an area of the motor cortex with high EYFP expression, a region containing approximately 100 cells was micro-dissected. This tissue sample was prepared and LC-MS/MS was employed as previously described⁶⁸, demonstrating EYFP was the predominant protein in its mass range (Figure S-9).

CONCLUSIONS

We demonstrate techniques that allow the detection of EYFP—and presumably, related XFP variants—using MALDI-MSI. When combined with transgenic mice with cell-selective YFP expression, this overcomes some of the challenges facing cell-cohort characterization using MSI by eliminating the needs for an optical image and for image registration. If preceded by tryptic digestion, this technique has the potential to be applied to additional MALDI-MS workflows²⁶ and other MSI techniques such as LAESI⁶⁹ and DESI.⁷⁰ Spectra generated by this technique could serve as a basis for generating cell-type classifier that could serve to extract cell-cohorts from unlabelled tissues.^{71,72} MSI at high spatial resolution occurs near the detection limit of a number of proteins, but is expected to increase as instrumentation advances. Promoters with alternative expression patterns have been employed to create numerous XFP animal lines, labeling additional cell-types⁷³, and have been combined, resulting in XFP mosaics exemplified by the Brainbow mouse.^{73,74} MSI could complement this technique, or with further instrumentation and method advancement, replace it, if cell-type-specific mass markers are employed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by ALSA grant 18-ILAO-420. ECR is in receipt of an NIH R25 (R25 CA-89017) Fellowship in partnership with the Ferenc Jolesz National Center for Image Guided Therapy at BWH (P41 EB015898). The work in the Ivanov laboratory was supported by the NIH awards 1R01GM120272 and R01CA218500. The authors would like to thank Alexandra Hendrickson, Somak Ray, Daniel P. Donnelly, David Calligaris, Amanda Clark, Michael Regan, and Begona Gimenez-Cassina Lopez for technical assistance and Kerry J. Ressler for the EYFP RNA sequence from transgenic mice.

REFERENCES

- (1). Armbrrecht L; Dittrich PS *Analytical chemistry* 2017, 89, 2–21. [PubMed: 28105840]
- (2). Galler K; Brautigam K; Grosse C; Popp J; Neugebauer U *The Analyst* 2014, 139, 1237–1273. [PubMed: 24495980]
- (3). Wang Z; Gerstein M; Snyder M *Nature Reviews Genetics* 2009, 10, 57.
- (4). Altelaar AFM; Heck AJR *Current opinion in chemical biology* 2012, 16, 206–213. [PubMed: 22226769]
- (5). Bandura DR; Baranov VI; Ornatsky OI; Antonov A; Kinach R; Lou X; Pavlov S; Vorobiev S; Dick JE; Tanner SD *Analytical chemistry* 2009, 81, 6813–6822. [PubMed: 19601617]
- (6). Holzlechner M; Strasser K; Zareva E; Steinhauser L; Birnleitner H; Beer A; Bergmann M; Oehler R; Marchetti-Deschmann M *Journal of proteome research* 2017, 16, 65–76. [PubMed: 27755872]
- (7). Yamanaka K; Chun SJ; Boillee S; Fujimori-Tonou N; Yamashita H; Gutmann DH; Takahashi R; Misawa H; Cleveland DW *Nature neuroscience* 2008, 11, 251–253. [PubMed: 18246065]
- (8). Zhang L; Vertes A *Angewandte Chemie* 2018, 57, 4466–4477. [PubMed: 29218763]

- (9). Qi M; Philip MC; Yang N; Sweedler JV ACS chemical neuroscience 2018, 9, 40. [PubMed: 28982006]
- (10). Winograd N Annual review of analytical chemistry 2016.
- (11). McDonnell LA; Piersma SR; Altelaar AFM; Mize TH; Luxembourg SL; Verhaert PDEM; van Minnen J; Heeren RMA Journal of Mass Spectrometry 2005, 40, 160–168. [PubMed: 15706616]
- (12). Giesen C; Wang HA; Schapiro D; Zivanovic N; Jacobs A; Hattendorf B; Schuffler PJ; Grolimund D; Buhmann JM; Brandt S; Varga Z; Wild PJ; Gunther D; Bodenmiller B Nature methods 2014, 11, 417–422. [PubMed: 24584193]
- (13). Gode D; Volmer DA The Analyst 2013, 138, 1289–1315. [PubMed: 23314100]
- (14). Chughtai K; Heeren R In Chem. Rev, 2010, pp 3237–3277. [PubMed: 20423155]
- (15). Körsge M; Pelster A; Dreisewerd K; Arlinghaus H The official journal of The American Society for Mass Spectrometry 2016, 27, 277–284.
- (16). Duan-Sun Z; Valeria P; Benjamin JP; Agnieszka KR; Poczek JC; Mei W; Haydn MP; James ME; David PC; Claude PL Nature 2012, 481, 520. [PubMed: 22246323]
- (17). Kraft ML; Weber PK; Longo ML; Hutcheon ID; Boxer SG Science (New York, N.Y.) 2006, 313, 1948.
- (18). Lovri J; Dunevall J; Larsson A; Ren L; Andersson S; Meibom A; Malmberg P; Kurczy ME; Ewing AG ACS nano 2017, 11, 3446. [PubMed: 27997789]
- (19). Monroe EB; Jurchen JC; Lee J; Rubakhin SS; Sweedler JV Journal of the American Chemical Society 2005, 127, 12152–12153. [PubMed: 16131155]
- (20). Passarelli MK; Pirkel A; Moellers R; Grinfeld D; Kollmer F; Havelund R; Newman CF; Marshall PS; Arlinghaus H; Alexander MR; West A; Horning S; Niehuis E; Makarov A; Dollery CT; Gilmore IS Nature methods 2017, 14, 1175–1183. [PubMed: 29131162]
- (21). Sheng L; Cai L; Wang J; Li Z; Mo Y; Zhang S; Xu J-J; Zhang X; Chen H-Y Int J Mass Spectrom 2017, 421, 238–244.
- (22). Sogawa K; Watanabe M; Sato K; Segawa S; Ishii C; Miyabe A; Murata S; Saito T; Nomura F Analytical and bioanalytical chemistry 2011, 400, 1905–1911. [PubMed: 21442367]
- (23). Dilillo M; Ait-Belkacem R; Esteve C; Pellegrini D; Nicolardi S; Costa M; Vannini E; Graaf E. L. d.; Caleo M; McDonnell LA Scientific reports 2017, 7, 603. [PubMed: 28377615]
- (24). Berkenkamp S; Kirpekar F; Hillenkamp F Science 1998, 281, 260–262. [PubMed: 9657721]
- (25). Mohammadi AS; Phan NT; Fletcher JS; Ewing AG Analytical and bioanalytical chemistry 2016, 408, 6857–6868. [PubMed: 27549796]
- (26). Heijs B; Tolner EA; Bovee JV; van den Maagdenberg AM; McDonnell LA Journal of proteome research 2015.
- (27). Zavalin A; Yang J; Hayden K; Vestal M; Caprioli RM Analytical and bioanalytical chemistry 2015, 407, 2337–2342. [PubMed: 25673247]
- (28). Anderson DM; Carolan VA; Crosland S; Sharples KR; Clench MR Rapid communications in mass spectrometry: RCM 2009, 23, 1321–1327. [PubMed: 19337978]
- (29). Porta T; Grivet C; Kraemer T; Varesio E; Hopfgartner G Analytical chemistry 2011, 83, 4266–4272. [PubMed: 21510611]
- (30). Lanni EJ; Masyuko RN; Driscoll CM; Aerts JT; Shrout JD; Bohn PW; Sweedler JV Analytical chemistry 2014, 86, 9139–9145. [PubMed: 25133532]
- (31). Rauser S; Marquardt C; Balluff B; Deininger SO; Albers C; Belau E; Hartmer R; Suckau D; Specht K; Ebert MP; Schmitt M; Aubele M; Hofler H; Walch A Journal of proteome research 2010, 9, 1854–1863. [PubMed: 20170166]
- (32). Eberlin LS; Norton I; Dill AL; Golby AJ; Ligon KL; Santagata S; Cooks RG; Agar NY Cancer research 2012, 72, 645–654. [PubMed: 22139378]
- (33). Groseclose MR; Castellino S Analytical chemistry 2013, 85, 10099–10106. [PubMed: 24024735]
- (34). Jonas O; Calligaris D; Methuku KR; Poe MM; Francois JP; Tranchese F; Changelian A; Sieghart W; Ernst M; Krummel DA; Cook JM; Pomeroy SL; Cima M; Agar NY; Langer R; Sengupta S Journal of biomedical nanotechnology 2016, 12, 1297–1302. [PubMed: 27319222]
- (35). Kim AJ; Basu S; Glass C; Ross EL; Agar N; He Q; Calligaris D Pain Practice, 0.

- (36). Spraggins JM; Rizzo DG; Moore JL; Noto MJ; Skaar EP; Caprioli RM *Proteomics* 2016, 16, 1678–1689. [PubMed: 27060368]
- (37). Rawlins CM; Salisbury JP; Feldman DR; Isim S; Agar NY; Luther E; Agar JN *Methods in molecular biology* 2015, 1346, 133–149. [PubMed: 26542720]
- (38). Longuespee R; Alberts D; Pottier C; Smargiasso N; Mazzucchelli G; Baiwir D; Kriegsmann M; Herfs M; Kriegsmann J; Delvenne P; De Pauw E *Methods* 2016, 104, 154–162. [PubMed: 26690073]
- (39). Zimmerman TA; Rubakhin SS; Romanova EV; Tucker KR; Sweedler JV *Analytical chemistry* 2009, 81, 9402–9409. [PubMed: 19835365]
- (40). Boggio KJ; Obasuyi E; Sugino K; Nelson SB; Agar NY; Agar JN *Expert review of proteomics* 2011, 8, 591–604. [PubMed: 21999830]
- (41). Wiegelmann M; Dreisewerd K; Soltwisch J *Journal of the American Society for Mass Spectrometry* 2016, 27, 1952–1964. [PubMed: 27549394]
- (42). Kompauer M; Heiles S; Spengler B *Nature methods* 2017, 14, 90. [PubMed: 27842060]
- (43). Jansson ET; Comi TJ; Rubakhin SS; Sweedler JV *ACS chemical biology* 2016, 11, 2588–2595. [PubMed: 27414158]
- (44). Rabe JH; D AS; Schulz S; Munteanu B; Ott M; Ochs K; Hohenberger P; Marx A; Platten M; Opitz CA; Ory DS; Hopf C *Scientific reports* 2018, 8, 313. [PubMed: 29321555]
- (45). Comi TJ; Neumann EK; Do TD; Sweedler JV *Journal of the American Society for Mass Spectrometry* 2017, 28, 1919–1928. [PubMed: 28593377]
- (46). Abdelmoula WM; Škrášková K; Balluff B; Carreira RJ; Tolner EA; Lelieveldt BPF; van der Maaten L; Morreau H; van den Maagdenberg AMJM; Heeren RMA; McDonnell LA; Dijkstra J *Analytical chemistry* 2014, 86, 9204–9211. [PubMed: 25133861]
- (47). Tsien RY *Annual review of biochemistry* 1998, 67, 509–544.
- (48). Zhang L; Sevinsky CJ; Davis BM; Vertes A *Analytical chemistry* 2018, 90, 4626–4634. [PubMed: 29505244]
- (49). Cristea IM; Williams R; Chait BT; Rout MP *Molecular & cellular proteomics: MCP* 2005, 4, 1933–1941. [PubMed: 16155292]
- (50). Feng G; Mellor RH; Bernstein M; Keller-Peck C; Nguyen QT; Wallace M; Nerbonne JM; Lichtman JW; Sanes JR *Neuron* 2000, 28, 41–51. [PubMed: 11086982]
- (51). Ormo M; Cubitt AB; Kallio K; Gross LA; Tsien RY; Remington SJ *Science* 1996, 273, 1392–1395. [PubMed: 8703075]
- (52). Porrero C; Rubio-Garrido P; Avendano C; Clasca F *Brain Res* 2010, 1345, 59–72. [PubMed: 20510892]
- (53). Luther E; Kametsky L; Henriksen M; Holden E *Methods in cell biology* 2004, 75, 185–218. [PubMed: 15603427]
- (54). Noor SSM; Tey BT; Tan WS; Ling TC; Ramanan RN; Ooi CW *J Liq Chromatogr R T* 2014, 37, 1873–1884.
- (55). Griesbeck O; Baird GS; Campbell RE; Zacharias DA; Tsien RY *The Journal of biological chemistry* 2001, 276, 29188–29194. [PubMed: 11387331]
- (56). Wachter RM; Elsliger MA; Kallio K; Hanson GT; Remington SJ *Structure* 1998, 6, 1267–1277. [PubMed: 9782051]
- (57). Daubner SC; Astorga AM; Leisman GB; Baldwin TO *Proceedings of the National Academy of Sciences of the United States of America* 1987, 84, 8912–8916. [PubMed: 3480518]
- (58). Cannon JR; Kluwe C; Ellington A; Brodbelt JS *Proteomics* 2014, 14, 1165–1173. [PubMed: 24596159]
- (59). Seeley EH; Oppenheimer SR; Mi D; Chaurand P; Caprioli RM *Journal of the American Society for Mass Spectrometry* 2008, 19, 1069–1077. [PubMed: 18472274]
- (60). Yang J; Caprioli RM *Analytical chemistry* 2011, 83, 5728–5734. [PubMed: 21639088]
- (61). Park J; Qin H; Scalf M; Hilger RT; Westphall MS; Smith LM; Blick RH *Nano letters* 2011, 11, 3681–3684. [PubMed: 21806063]

- (62). van Remoortere A; van Zeijl RJ; van den Oever N; Franck J; Longuespee R; Wisztorski M; Salzet M; Deelder AM; Fournier I; McDonnell LA *Journal of the American Society for Mass Spectrometry* 2010, 21, 1922–1929. [PubMed: 20829063]
- (63). Franck J; Longuespee R; Wisztorski M; Van Remoortere A; Van Zeijl R; Deelder A; Salzet M; McDonnell L; Fournier I *Medical science monitor: international medical journal of experimental and clinical research* 2010, 16, BR293–299. [PubMed: 20802405]
- (64). Cohen SL; Chait BT *Analytical chemistry* 1996, 68, 31–37. [PubMed: 8779435]
- (65). Mainini V; Angel PM; Magni F; Caprioli RM *Rapid communications in mass spectrometry: RCM* 2011, 25, 199–204. [PubMed: 21157864]
- (66). Agar NYR; Yang HW; Carroll RS; Black PM; Agar JN *Analytical chemistry* 2007, 79, 7416–7423. [PubMed: 17822313]
- (67). Hankin JA; Barkley RM; Murphy RC *Journal of the American Society for Mass Spectrometry* 2007, 18, 1646–1652. [PubMed: 17659880]
- (68). Li S; Plouffe BD; Belov AM; Ray S; Wang X; Murthy SK; Karger BL; Ivanov AR *Molecular & cellular proteomics: MCP* 2015, 14, 1672–1683. [PubMed: 25755294]
- (69). Shrestha B; Patt JM; Vertes A *Analytical chemistry* 2011, 83, 2947–2955. [PubMed: 21388149]
- (70). Eberlin LS; Liu X; Ferreira CR; Santagata S; Agar NY; Cooks RG *Analytical chemistry* 2011, 83, 8366–8371. [PubMed: 21975048]
- (71). Ong TH; Kissick DJ; Jansson ET; Comi TJ; Romanova EV; Rubakhin SS; Sweedler JV *Analytical chemistry* 2015, 87, 7036–7042. [PubMed: 26076060]
- (72). Comi TJ; Do TD; Rubakhin SS; Sweedler JV *Journal of the American Chemical Society* 2017, 139, 3920–3929. [PubMed: 28135079]
- (73). Livet J; Weissman TA; Kang H; Draft RW; Lu J; Bennis RA; Sanes JR; Lichtman JW *Nature* 2007, 450, 56–62. [PubMed: 17972876]
- (74). Weissman TA; Pan YA *Genetics* 2015, 199, 293–306. [PubMed: 25657347]

Figure 1. Genetically-encoded fluorescence enables the detection of cell cohorts *in situ*. Quantitative laser scanning cytometry (LSC) images of a 12 μm thick section from a ‘Thy-1 YFP-H’ mouse strain shown at varying magnifications illustrate neurons (yellow, entire neuron labeled) and all cells (blue, nuclei labeled with Hoechst stain) using 408 and 488 nm fluorescent filters. A) Entire coronal tissue section with (B) and (C) showing magnified regions within the motor cortex that contain Layer V motor neurons labeled with EYFP.

Figure 2.

Isolation of EYFP proteoforms from YFP-16 mouse brains. A) Fluorimetry at 527 nm confirms successful ammonium sulfate fractionation of EYFP. B) The EYFP-enriched ammonium sulfate fraction was subjected to anion exchange chromatography. Subsequent fluorimetry at 527 nm indicated that fractions 7–11 contained EYFP. C) LC-ESI-MS analysis (top: raw spectra, bottom: deconvoluted spectra) of fractions 8 (exemplar of fractions 7–9) and 11 (exemplar of fractions 10 and 11) assessed the intact mass and purity of EYFP. Two EYFP proteoforms were detected at intact masses of 26,882 and 26,754 Da.

The lighter proteoform was found to be the product of carboxypeptidase activity during purification and was not used for subsequent analysis.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3.

Characterization of EYFP. Top-down protein characterization, peptide mass fingerprinting, and LC-MS/MS analysis, together with the intact masses determined above, were used to confirm the primary structure and modifications of EYFP. A) ESI-FTICR-MS using funnel-skimmer dissociation determined that the EYFP expressed in studied mice underwent N-terminal methionine excision followed by N-terminal acetylation, and also confirmed the N-terminal sequence up to residue twelve. B) Determined sequence and modifications of EYFP expressed in mouse neurons with peptides observed by PMF of neuron-purified EYFP and

LC-MS/MS of neuron protein extract shown. G'Y'G' fixed modification of -20.026 Da (monoisotopic) was used. C) Crystal structure of EYFP (PDB ID 2YFP) and chemical structure of the G'Y'G' chromophore (5-imidazolinone ring and didehydrotyrosine formation) observed at residues 65–67.

Figure 4.

Optimized *in situ* detection of EYFP in MALDI-MS using two different automated deposition systems. A) 60% ACN + 0.1% TFA is the most commonly used solvent for SA but was not effective with the TM Sprayer. B) By adjusting the solvent ratio to 50% ACN, protein extraction improved and EYFP was detected. C) Protein sensitivity and EYFP detection can also be achieved with the ImagePrep (Bruker). Rehydration with a 5% acetic acid solution showed significant improvement in the extraction of proteins and the detection

of EYFP. The ImagePrep produced comparable results to the TM Sprayer but required nearly 10x the amount of time to coat a slide.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5.

MALDI-MS images in register with the fluorescence image of a YFP-16 mouse brain at 50 and 25 μm spatial resolution (5–100% max. peak intensity for all proteins shown in all images). A and F) Fluorescence microscopy images of mouse brain tissue sections displaying where EYFP is expressed. B and G) MSI of EYFP with a yellow mass filter ($27,000 \pm 1.2\%$) applied. C and H) MSI of Purkinje Cell Protein 4 (Teal, $6720 \pm 0.12\%$), Myelin Basic Protein (Red, $14,125 \pm .25\%$), and EYFP (Yellow) displayed simultaneously. D, E, I, and J) Magnifications of the regions from panels C (for D and E) and H (for I and J) outlined in white and magenta, showing a region of the mouse brain containing the striatum, corpus callosum and sensory cortex. K) Overall average spectrum of the MSI displayed in panels F-J, with proteins of interest and their mass filters shown. MSI data processing is described in detail in the supplemental methods on page S4.