HHS Public Access

Author manuscript

Anal Chem. Author manuscript; available in PMC 2021 September 21.

Published in final edited form as:

Anal Chem. 2020 January 21; 92(2): 2005–2010. doi:10.1021/acs.analchem.9b04396.

Automated Spatially Targeted Optical Micro Proteomics investigates inflammatory lesions in situ

Bocheng Yin[1], **Laura R Caggiano**[2], **Rung-Chi Li**[3,5], **Emily McGowan**[3], **Jeffrey W Holmes**[2,4], **Sarah E Ewald**[*,1]

[1]Department of Microbiology, Immunology and Cancer Biology and the Carter Immunology Center, University of Virginia School of Medicine, Charlottesville VA, USA, 22903.

[2]Department of Biomedical Engineering, University of Virginia School of Medicine, Charlottesville, VA, USA, 22903.

[3]Division of Allergy and Clinical Immunology, Department of Medicine, University of Virginia School of Medicine, Charlottesville, VA, USA, 22903.

[4]School of Engineering, University of Alabama at Birmingham, Birmingham, AL, USA, 35294.

[5]Department of Allergy and Immunology, Northern Light Health, Bangor Maine, Maine, USA, 04401.

Abstract

Tissue microenvironment properties like blood flow, extracellular matrix, or proximity to immune infiltrate are important regulators of cell biology. However, methods to study regional protein expression in the native tissue environment are limited. To address this need, we developed a novel approach to visualize, purify and measure proteins in situ using Automated Spatially Targeted Optical Micro Proteomics (AutoSTOMP). Here we report custom codes to specify regions of heterogeneity in a tissue section and UV biotinylate proteins within those regions. We have developed LC-MS/MS compatible biochemistry to purify those proteins and label-free quantification methodology to determine protein enrichment in target cell types or structures

[*]Corresponding Author: Phone: 434-924-1925. se2s@virginia.edu.

Author Contributions

B.Y., and S.E.E. designed the experiments. L.C and J.H performed rat cardiac infarcts. R.L. and E.M. collected clinical specimens. B.Y. performed AutoSTOMP experiments and data analysis. B.Y. and S.E.E. prepared the manuscript.

Supporting Information

The following supporting information is available free of charge at ACS website http://pubs.acs.org:

Supplementary Methods

Figure S1. AutoSTOMP accurately defines the coordinates of individual fields of view within user defined borders and targets them for re-imaging.

Figure S2. Inflammation regulatory pathways are enriched in ‘CD68’ fractions isolated by autoSTOMP and amino acid metabolism pathways are enriched in ‘flow through’ fractions of rat cardiac infarcts.

Figure S3. Phagosome maturation and inflammatory signaling pathways are enriched in the CD68 fractions and mitochondrial metabolism pathways are enriched in the flow through fractions of rat cardiac infarcts.

Figure S4. AutoSTOMP selectively biotinylates IgG4+ regions of human EoE esophagus biopsies.

Figure S5. Glycolytic metabolism and inflammatory pathways are enriched in lgG4+ regions of EoE biopsies and protein regulators of epithelial differentiation enriched in the flow through.

Table S1. The antibodies used on rat scar tissues.

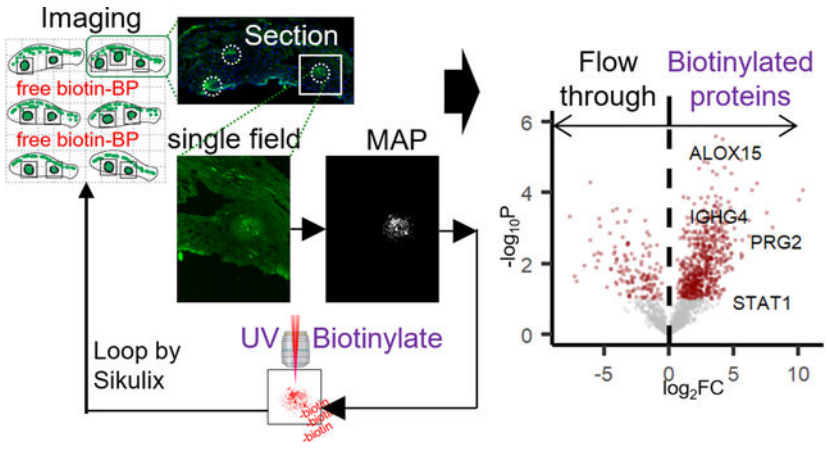
Table S2. The antibodies used on EoE samples.

Conflict of interest

The authors declare no competing financial interest.

relative to non-target regions in the same sample. These tools were applied to a) identify inflammatory proteins expressed by $CD68^+$ macrophages in rat cardiac infarcts and b) characterize inflammatory proteins enriched in $IgG4^+$ lesions in human esophageal tissue. These data indicate that AutoSTOMP is a flexible approach to determine regional protein expression in situ on a range of primary tissues and clinical biopsies where current tools and sample availability are limited.

Graphical Abstract



Keywords

AutoSTOMP; discovery proteomics; inflammation; proximity-biotinylation

Introduction

Technological advances and innovation in ‘big data’ analysis tools have revolutionized the proteomics in the last decades. These tools are still most frequently applied to studying protein expression in isolated cell-types or in bulk tissue lysates. There is still tremendous need for discovery proteomics techniques designed to study the biology of specific cell types in the context of the in situ tissue microenvironment. Laser capture microdissection, can isolate individual cells down to ~10 μm resolution for proteomics analysis.[1] However, this process is time intensive and susceptible to neighboring cell contamination. More recently, mass spectrometric imaging approaches have facilitated protein discovery using MALDI of cells or cell in situ at a 50–100 μm scale.[2] This approach provides strong regional selectivity, but it is ‘qualitative’ rather than ‘quantitative’ because MALDI-MSI cannot directly identify the protein without adequate resolution for peptide fragment sequencing.[2,3,4] We recently developed a technique called Automated Spatially Targeted Optical (AutoSTOMP) which uses standard immunofluorescence imaging to visualize structures of interest and the two-photon laser source to selectively conjugate photo-activatable-biotin tags to any protein within the structure. Biotinylated proteins are then isolated by streptavidin precipitation and identification by LC-MS/MS with label-free quantitation protocol. Previously, we demonstrated that AutoSTOMP enriches proteins from the obligate intracellular pathogen *Toxoplasma gondii* (*T. gondii)* within infected human and mouse macrophages[5]. By modifying the MAP to encompass the region surrounding but excluding *T. gondii*, host

and *T. gondii* proteins localized to the parasite vacuole membrane were identified. These studies demonstrated that AutoSTOMP can enrich proteins at a 1 μm scale and identify proteins with as little as 1 μg of protein per sample[5]. Proximity-based protein discovery tools that use a label targeting enzyme to biotinylate nearby proteins (BioID, TurboID, APEX) have an excellent resolution of ~10 nm[6,7]. However, they are limited to cell lines or animal models that have tools for genetic modification. Alternatively, in SPPLAT/BAR, biotin targeting is mediated by antibodies conjugated to a peroxidase[8,9]. Image-guided tagging is a central advantage of AutoSTOMP, which allows the user to specify biotin targeting based on co-localization stains or by thresholding, dilating, or eroding the boundaries of the image used to guide biotinylation at least ten times the resolution of laser capture micro dissection[1]. Any sample where one or more fluorescent markers (e.g., tags, probes or antibodies) are available to identify SOI is a candidate for AutoSTOMP, so this technique could be a transformative approach to perform localization-dependent protein discovery on a broad range of human clinical specimens.

In practice, however, adapting AutoSTOMP for tissue sections poses several unique challenges compared to cell culture samples. These include: identifying signal versus background in tissues with high or variable autofluorescence, automating the selection of SOI within tissue microdomains rather than elsewhere in a section (particularly for low-density SOI), and optimizing the digestion and streptavidin precipitation biochemistry to handle fixed tissues with extensive extracellular matrix networks using a protocol compatible with LC-MS[10]. Here we report the AutoSTOMP workflow to address the unique demands of in situ proteomics. This includes an updated software analysis package that defines the coordinates of multiple sections on a slide, identifies relevant microdomains of each section, and generates a tile array to automate SOI crosslinking. We have developed biochemistry protocols to examine protein enrichment in inflammatory lesions using two disease systems: 1) a rat cardiac infarct model, chosen as a tissue type with extensive extracellular matrix protein crosslinking, which poses a difficulty for protein purification; 2) human eosinophilic esophagitis (EoE), selected for the small biopsy size and low frequency of lesions in each tissue section.

Experimental Section

Note: Full protocols can be found in “Supplementary Methods” of Supporting information.

Rat cardiac infarct and Eosinophilic Esophagitis (EoE) biopsy collection and staining.

Rat myocardial infarcts were induced in 8-week-old male Sprague-Dawley rats (Envigo) by left anterior descending (LAD) coronary artery permanent ligation. 1-week post-surgery, the scar region was dissected, frozen in liquid-nitrogen-chilled isopentane, and embedded in OTC. 7 μm cryosections were methanol fixed for 20 minutes on ice and stained with an antibody specific to CD68 (clone: ED1, Bio-Rad). Animal protocols were approved by the University of Virginia Institutional Animal Care and Use Committee.

Six 1 mm biopsies were collected from a patient diagnosed with active EoE (ε15 eosinophils/hpf), according to consensus guidelines, using standard endoscopy procedures[11]. Biopsies were fixed and sectioned as described and stained with an antibody specific

to human immunoglobuin lgG4 (clone MRQ-44, Cell Marque). The human study was approved by the University of Virginia Institutional Review Board (IRB), which requires written participant consent (IRB-HSR#19562).

Tissue sections were treated with an avidin/biotin blocking kit (SP-2001, Vector Laboratories) then mounted with biotin-dPEG3-benzophenone (Biotin-BP, Quanta BioDesign) in 50/50 (v/v) DMSO/water at a concentration of 1 mM. Each slide was prepared immediately prior to AutoSTOMP imaging.

AutoSTOMP 2.0

Imaging and photo-crosslinking was performed on a LSM880 microscope (Carl Zeiss) equipped with a 25x oil immersion lens (LD LCI Plan-Apochromat 25x/0.8 Imm Korr DIC M27) and a Chameleon multiphoton light source (Coherent). AutoSTOMP 2.0, the upgraded SikuliX (version 1.1.4, http://sikulix.com/) integrated workflow was modified from the previous protocol[5] and scripted for the tissue to facilitate SOI selection on multiple tissue sections per microscopic slide. Step-by-step instruction and source codes are deposited at https://github.com/boris2008/AutoSTOMP_2.0.git.

Following photo-labeling, each sample was detached from the coverslip. Excess, unconjugated Biotin-BP was rinsed with 50/50 (v/v) DMSO/water 3 times, then with water 3 times. The slides were stored at −80 °C before processing replicates in tandem. Rat cardiac sections were lysed in the hydroxylamine lysis buffer[10] (1 M NH_2OH−HCl, 8 M urea, 0.2 M K_2CO_3, pH = 9.0) at 45 °C for 17 h to extract proteins from the insoluble extracellular matrix. EoE biopsy sections were lysed in DTT/SDS buffer[12,13] (0.1 M Tris-HCl, 0.1 M DTT, 4% SDS, pH=8.0) at 99 °C for 1 h. Tissue lysates were then diluted 1:10 in TBS-0.1% SDS, then incubated with streptavidin (SA) magnetic beads (Pierce #88817) at room temperature for 1 h. Biotinylated proteins were precipitated by a magnet. The unbound proteins were collected as the ‘flow through’ fraction and precipitated with 100% trichloroacetic acid. The biotinylated proteins were eluted from the magnetic beads in laemmli buffer at 96 °C for 5 min and collected as the AutoSTOMP fraction. The protein pellet of the ‘flow through’ fraction was resuspended in laemmli buffer at 96 °C for 5 min. The fractions were resolved in the SDS-PAGE gel at 70 V for 12 min. For each lane, a 1 cm gel fragment was excised and submitted for mass spectrometry analysis at the University of Virginia Biomolecular Analysis Facility. The samples were run on a Thermo Orbitrap Exploris 480 mass spectrometer system with an Easy Spray ion source connected to a Thermo 75 μm × 15 cm C18 Easy Spray column (trap column first). 6 μL of the extract was injected and the peptides were eluted from the column by an acetonitrile/0.1 M formic acid gradient at a flow rate of 0.3 μL/min over 2.0 hours. The nanospray ion source was operated at 1.9 kV. The digest was analyzed using the rapid switching capability of the instrument acquiring a full scan mass spectrum to determine peptide molecular weights followed by product ion spectra (Top10 HCD) to determine amino acid sequence in sequential scans.

The raw mass spectra data were parsed by MaxQuant[14] (versions 1.6.14.0, Max Planck Institute of Biochemistry). The MS raw and MaxQuant search data ($CD68^+$data in Figure 2: accession: PXD026818; and EoE data in Figure 5: accession: PXD026819) have been deposited at PRIDE (https://www.ebi.ac.uk/pride/). The MaxQuant results were then

analyzed following the label-free quantification (LFQ) data analysis protocol[14]. Student's t-test (permutation-based FDR < 0.05) and t-SNE clustering[15] were applied in Perseus[16] (versions 1.6.14.0, Max Planck Institute of Biochemistry). The resulting data were plotted in R (www.r-project.org) with the installed packages "ggplot2", "ggrepel", "heatmap.2" or using GraphPad Prism (version 8.2.1).

Immunofluorescence and streptavidin fluorescence staining

To validate Biotin-BP cross-linking, following AutoSTOMP2.0 cross linking, some samples were washed and stained with AlexaFluor-594 Streptavidin (#016-580-084, Jackson ImmunoResearch) in TBST. Samples were reimaged to co-localize with the CD68 or lgG4 signal. For immunofluorescence co-staining experiments, samples were fixed and sectioned as described for AutoSTOMP then blocked in 5% BSA and 1:200 Fc block (human FC Clone 3070, BD or rodent FC block clone 93, Affymetrix) for 1 hour in TBST prior to staining. A full list of antibodies for validation is provided in the Supporting Information.

Results and Discussion

Cardiac infarct macrophages are selectively biotinylated by AutoSTOMP

Automated spatially targeted optical micro proteomics (AutoSTOMP)[5] is a proximity-based protein labeling tool that uses standard fluorescence microscopy to visualize structures of interest (SOI). The fluorescence signal is used to identify the pixel coordinates of the SOI and generate a MAP file. The MAP file then guides two-photon excitation of the SOI with UV energy light, which conjugates benzophenone-biotin (Biotin-BP) present in the mounting media to any nearby carbon or nitrogen via the benzophenone moiety. Imaging, MAP generation, and Biotin-BP conjugation are repeated for every field of view and automated using SikuliX icon recognition software. Once biotinylation is complete, unconjugated Biotin-BP is washed away. The samples are digested off of the slide. Biotinylated proteins are streptavidin precipitated, then digested for identification by liquid chromatography-mass spectrometry (LC-MS)[5].

To test the ability of the AutoSTOMP protocol to selectively biotinylate structures of interest within tissue sections we first examined a rat myocardial infarction model. In this model, trauma caused by ligation and infiltrating immune cells causes fibroblast activation and deposition of scar tissue that ultimately impairs cardiac function. Macrophages are thought to play a role in inflammatory regulation and damaged cell turnover in the tissue. One week after surgical ligation on the left anterior descending (LAD) coronary artery, the infarct region was dissected and cryosectioned for immunofluorescence staining. The infarct region or scar is defined by loss of organized cardiac muscle structure, regions of extracellular matrix, and fibroblast expansion as well as infiltrating immune cells. To differentiate between the scar and neighboring healthy tissue, low-resolution tile scans were performed on adjacent serial sections stained with Hematoxylin and Eosin (H&E) or the macrophage marker CD68 (Figure 1A, yellow border). Using the AutoSTOMP software module the scar region was tiled into individual fields of view with defined pixel coordinates (Figure 1B). Field of view segmentation and the accuracy of the automated re-imaging program were validated experimentally (Figure S1).

To biotinylate the $CD68^+$ SOI proteins within the scar borders, each tile was imaged at 488 nm (Figure 1B). Each image of the CD68 signal was thresholded, and the pixel coordinates were defined in a MAP file. The MAP file then guided the two-photon at 720 nm wavelength, which selectively conjugated Biotin-BP to proteins within the $CD68^+$ region. This process was fully automated across each field of view in the scar region tile array (Figure 1B). To validate the accuracy of Biotin-BP targeting to the $CD68^+$ regions, some sections were washed, stained with streptavidin-594 and reimaged (Figure 1C, note photobleaching of Post-STOMP CD68 signal). These data indicate that the AutoSTOMP software allows the user to define regions of a tissue section, tile this region into fields of view, and accurately image and biotinylate SOI in an automated fashion.

AutoSTOMP enriches macrophage endolysosomal and inflammatory signaling proteins in $CD68^+$ regions of cardiac infarcts.

To enrich the CD68 SOI proteins, excess unconjugated Biotin-BP was washed off of the slide, and a sample lysate was prepared. Biotinylated SOI proteins were streptavidin precipitated, eluted as the ‘CD68’ fraction. To measure protein levels of the rest part of the entire scar sample, the unbound, ‘flow through’ fractions were also collected (Figure 2A). Of note, the ‘flow through’ fractions are expected to contain $CD68^-$ regions as well as any $CD68^+$ regions of the section that were deeper than the focal excitation volume of the two-photon (approximately 2.4 μm in the Z axis at 25x magnification[17]). Peptides were identified by LC-MS/MS and MaxQuant LFQ[18] method which measures peak volume normalized across samples to limit artifacts of run-to-run variability on the LC-MS.

1,671 rat proteins were identified across the ‘CD68’ and ‘flow through’ replicates. Relative expression of each protein across the samples was evaluated by z-score and hierarchical clustering, which indicated that the majority of proteins identified were enriched in ‘CD68’ (Figure 2B, left green bar) or lower in ‘CD68’ fractions (Figure 2B, left red bar) relative to ‘flow through’ fractions. T-distributed stochastic neighborhood embedding (t-SNE)[15] supported the conclusion that AutoSTOMP effectively enriched SOI proteins as there was more similarity within fractions than within each sample (Figure 2C).

Of the 1,671 proteins identified, 28.2% of proteins were more abundant in the ‘CD68’ fractions, and 33.7% of proteins were less abundant compared to the ‘flow through’ fractions (Figure 2D, red dots FDR < 0.1). As expected, macrophage markers CD68 (used to guide tagging), Lysozyme 2 (LYZ2 or LYZM), and Lysophosphatidylcholine acyltransferase 2 (LPCAT2) were enriched in the ‘CD68’ fractions (Figure 2D)[19–21]. The LYZ2 signal is co-localized with CD68 (Figure 3A). Consistent with the large phagocytic capacity of macrophages, the lysosomal proteins, Lysosome-associated membrane glycoprotein 1 (LAMP1), acid phosphatase 2 (ACP2), vacuolar ATP-dependent proton pumps (ATP6V1B2, ATP6V1C1, ATP6V1H) and cathepsin (CTSB, CTSZ) were enriched in ‘CD68’ fractions (Figure 2D). The $CD68^+$ signal is partially co-localized with LAMP1 (Figure 3B). The complement proteins C4 and C1Q, which are synthesized by macrophages in response to inflammatory stimuli and modulate phagocytic uptake of cargo, were enriched in ‘CD68’ fractions (Figure 2D)[22]. C1QC is expressed by cardiac resident macrophage and co-localized with CD68 (Figure 3C)[23].

To identify broader signaling networks associated with $CD68^+$ macrophages, the proteins that were significantly different between the 'CD68' and 'flow through' fractions (Figure 2D, red) were analyzed using David Bioinformatic Resources to annotate Gene Ontology (GO) terms (Figure S2)[24]. Consistent with inflammatory tissue remodeling, proteins involved in Type 1 interferon cytokine signaling and Glycerol 3-phosphate signaling and metabolism, a pathway that generates lipid signaling mediators of wound healing, were among the most represented 'biological process' GO terms in 'CD68' fractions (Figure S2A). Proteins regulating amino acid metabolism (aspartate, glutamate), ammonium compound (carnitine), RNA export[25], and extracellular matrix synthesis were enriched in 'flow through' fractions, consistent with muscle regenerative functions of stromal cells (Figure S2B). The central regulator of carnitine metabolism, acyl-coA dehydrogenase (ACADL) was one of the most significantly enriched proteins in the flow through (Figure 2D). Extracellular matrix proteins included Collagens (COL1A1 and COL18A1), fibronectin 1 (FN1) and periostin (POSTN) and elastin microfibril interfacer 1 (EMILIN1) are also enriched (Figure 2D).

Gene set enrichment analysis (GSEA) was also performed on all 1,671 proteins using the REACTOME database (Figure S3)[26,27]. The most highly enriched gene sets in the 'CD68' fractions were components of the Eph-ephrin and Fcγ receptor pathways, which signal through Rho GTPases (e.g. RhoA, Rac1, and Cdc42) to facilitate actin remodeling and the phagocytic uptake of cargo (Figure S3, red)[28–30]. Interleukin 12 (IL-12), the main macrophage product necessary for IFN-γ expression was also enriched in 'CD68' fractions (Figure S3, red)[31,32]. Similar to the results of the GO search, the 'flow through' fractions were enriched for proteins belonging to mitochondrial biogenesis and respiration, muscle contraction and extracellular matrix regulation (Figure S3, blue). In summary, these data show that AutoSTOMP is an effective tool to enrich and measure macrophage proteins in infarcted cardiac tissue sections.

AutoSTOMP enriches granulocyte proteins, eicosanoid inflammatory mediators and glycolytic metabolism machinery from $IgG4^+$ inflammatory lesions in esophageal biopsies.

We next asked if AutoSTOMP could selectively enrich proteins associated with discrete regions of human tissue biopsies. Eosinophilic esophagitis (EoE) is a disease driven by dietary allergens that leads to focal inflammatory lesions within the esophagus, which are characterized by infiltration of eosinophils and mast cells, and increased levels of Th2 cytokines[33]. The immunoglobulin G isotype IgG4 has recently been identified in the esophageal tissue and is increasingly recognized as a relevant feature of this disease[34,35]. However, progress towards understanding disease pathogenesis has been hindered by a lack of well-established animal models or the extremely limited access to samples from the primary site of inflammation and heterogeneity in the biopsy tissue[36]. To determine if AutoSTOMP was amenable to study EoE pathology, six 1 mm esophagus biopsies were isolated by endoscopy from a patient diagnosed with active EoE (Figure 5). Immunofluorescence staining for lgG4 showed that the lesions measured between 50–300 μm in diameter or approximately 10% of each section (Figure 5A). Using AutoSTOMP, the boundary of all 6–8 sections per slide was defined by the user to facilitate a low-resolution tile scan (Figure 4A–C). To avoid the high background fluorescence signal from the apical

epithelium, each lgG4$^+$ SOI was selected by the user (Figure 4D, Figure 5B 'threshold' signal vs. 'MAP'). A tile array of the SOI pixel coordinates (Figure 4E) was then generated to automate Biotin-BP tagging (Figure 4F–H). To validate Biotin-BP targeting, some slides were reserved for streptavidin-594 staining and reimaged (Figure S4).

To identify the biotinylated proteins enriched in the lgG4$^+$ lesions (Figure 5A–B), the sample lysate was incubated with streptavidin beads ('lgG4' fraction) and the unbound, 'flow through' fractions were reserved as a total protein control as described in Figure 2A. In total, 2,007 human proteins were identified across 3 samples where each sample (1–3) represented paired 'lgG4' or 'flow through' fractions pooled from 2 biopsies. (Figure 5C). When the level of each protein was evaluated across all samples, hierarchical clustering revealed three major groups enriched in 'IgG4' fractions (Figure 5C, left green bar) or enriched in 'flow through' fractions (Figure 5C, left red & blue bars). Variability in the 'flow through' samples accounted for the bimodal clustering of the red and blue groups. This indicated that the lgG4$^+$ regions were more similar in protein identity than the stromal cells from each biopsy pair and underscored the selective enrichment of the AutoSTOMP procedure. The selectivity of AutoSTOMP was also evaluated by t-SNE, which showed that despite the variability within 'flow through' fractions, there was more similarity within fractions than within each sample (Figure 5D).

Of the 2,007 proteins identified in the patient's biopsies, 27.9% were significantly enriched in the 'lgG4' fractions, and 12.3% were depleted in 'lgG4' fractions relative to the 'flow through' (Figure 5E). Granulocyte secretory proteins were among the most highly enriched proteins in the 'lgG4' fractions, including the proteoglycans PRG2 and PRG3, Defensin DEFA3 and eosinophil peroxidase (EPX) (Figure 5E)[37]. Enzymes regulating the synthesis of eicosanoid lipid inflammatory regulators were also enriched in the 'IgG4' fractions, including arachidonate-15 lipoxygenase (ALOX15), Leukotriene A-4 hydrolase (LTA4H), and glutathione S-transferase P (GSTP1), which modifies prostaglandin A2 upstream of eicosanoid synthesis[38–40]. Mediators of inflammatory cytokine production were also enriched in the IgG4 regions, including the transcription factors STAT1 (Signal Transducer and Activator of Transcription 1) and STAT3, the interferon induced effector MX1, complement protein C4A, and the Type I IL-1 receptor antagonist (IL-1RA) consistent with local activation of the inflammatory response. There was a partial overlap in the staining of PRG2, ALOX15, and IL-1RA with the IgG4 lesions, confirming their presence in the patient's EoE lesions as determined by LC-MS/MS (Figure 6).

To identify signaling networks associated with the IgG4+ lesions, we searched the significantly differentially expressed proteins (Figure 5E, red) against the 'biological process' GO term library. Most of the enriched pathways were characterized by the abundance of proteasome components detected in the IgG4 lesions (Figure S5A, asterisks). Independent of this proteasome signature, glycolytic metabolic pathways were the most highly represented GO-terms in the 'IgG4' fraction, consistent with metabolic demands needed to drive an inflammatory response (Figure S5, red)[41]. By contrast, the 'flow through' fractions were enriched in proteins regulating epithelial turnover and differentiation (Figure S5, blue) including fatty acid binding protein (FABP5), calmodulin like 5 (CALML5), small proline rich protein 3 (SPRR3), the calcium binding defensin S100A7, beta-actin (ACTB)

and Wiskott-Aldrich Syndrome Protein family member 2 (WASF2) (Figure 5E)[42]. Although future studies with an expanded number of patients will be needed to draw definitive conclusions about the mediators of EoE, these data indicate that AutoSTOMP is an effective protein discovery tool to identify immune effectors from discrete inflammatory foci in human tissue biopsies.

Conclusions

This work meets a long-standing need for spatial proteomics tools that can be applied to discover mechanisms of human disease. Our work is impactful in four areas. First, we present two novel, automated image parsing and biotinylation strategies. Second, we report two new biochemical purification protocols for highly modified samples. Third, we have applied a label-free quantification methodology to normalize protein expression across samples. Finally, we report a direct translational application to primary human patient biopsies.

In theory, image-guided tagging means that almost any structure that can be visualized is a potential target for AutoSTOMP. However, in practice, heterogeneity in SOI size, shape, and frequency were barriers to performing STOMP in tissues. This was complicated by the requirement to use multiple software platforms to tile array (python), acquire images (Zen Black), threshold images (FIJI) and perform cross linking (Zen Black) for each field of view. The icon recognition software SikuliX overcomes these limitations by allowing the user to automate image capture, image processing, and tiling discrete SOI across a tissue section. In addition, SikuliX automation facilitates the experimental timeline since AutoSTOMP-mediated biotinylation can take several days per replicate for rare structures with a high degree of complexity. This time frame is similar to what would be required to accumulate samples by laser capture microdissection (LCM).[1] The spatial resolution of AutoSTOMP (~1 μm) is better than LCM (~10 μm), affording cleaner margins. Although LCM has been used with single cell RNA sequencing, most proteomics readouts require pooled material, similar to AutoSTOMP.[5] In contrast, mass spectrometry imaging (MSI) can be performed on cells or clusters of cells as small as 10–50 μm without pooling so that finite spatial information is retained.[2] However, use of MALDI as a read out means that protein detection is more limited by MSI.[2] In pooling individual structures for protein identification by LC-MS/MS analysis, AutoSTOMP affords detection of a wider range of proteins and quantification of their relative expression. This comes at the cost of loss of spatial resolution for individual cells, however, as discussed below, this can be compensated for by selecting co-localization criteria and repeating the autoSTOMP experiment. All three of these techniques will likely prove particularly powerful when paired with new tools for multi-parameter imaging such as imaging mass cytometry[43], multiplexed ion bean imaging (MIBI)[44] or CHIP cytometry[45] that will facilitate validation of protein expression across cell types in heterogenous tissues.

One benefit of AutoSTOMP is that it is performed on fixed samples. This facilitates banking of samples that can be sectioned, stained and processed at later dates. This flow-through also means that, as new structural markers are identified by AutoSTOMP, the map file can be refined in subsequent experiments using new co-localization markers. However, fixation

also introduces challenges for tissue digestions and streptavidin precipitation. The cardiac infarct model was specifically selected because extensive extracellular matrix is noted to be problematic for tissue digestion and protein purification. Here we describe a hydroxylamine tissue lysis protocol that allows us to effectively enrich cellular proteins and ECM components for streptavidin precipitation and identification by LC-MS. Hydroxylamine interrupts protein quaternary structure by cleaving peptide bonds in particular at asparagine-glycine cites[10]. For tissues that require less stringent lysis procedures, like esophageal biopsy tissue, denaturation in DTT and SDS was sufficient to enrich granulocyte-specific proteins and immune signaling molecules associated with IgG4.

Finally, developing a protocol to analyze protein levels in the ‘flow through’ has maximized the use of each sample, which is important for rare samples or in studies where biopsy-to-biopsy variability is expected to be high. A central challenge in making this comparison is that protein abundance in the AutoSTOMP fraction is very low, typically less than or equal to 1 μg. By contrast, the protein abundance and diversity in the ‘flow through’ portion of the sample is typically far higher. For low abundance samples normalizing protein levels across replicates is important to limit artifacts or run-to-run variability on the mass spectrometer. Isotopic post-labeling and sample pooling for a single LC-MS/MS run is one solution. However, this methodology would increase competition for detection of rare proteins in the autoSTOMP sample relative to the flow-through fraction. To circumvent these problems, we employed a MaxQuant LFQ label-free quantification method that normalizes protein abundance across individual samples. This method has been validated to compare samples where protein abundance differs by 10-fold which is well suited to address these data sets[14,18]. In summary, these data indicate that AutoSTOMP is a robust and flexible tool to perform protein discovery in heterogenous tissues and primary tissue biopsies with wide implications for our understanding of human disease pathophysiology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Dr. Nicholas E. Sherman and Dr. Jeong-Jin Park for LC-MS/MS services at the W.M. Keck Biomedical Mass Spectrometry Laboratory. This work was supported by NIH R21AI156153, R35GM138381 (SEE), AHA/Allen Institute 31315 (JH and SEE) and NIH K22 AI116727 start-up funds from the University of Virginia SOM and the Emily Couric Cancer Center.

References

(1). Espina V; Wulfkuhle JD; Calvert VS; VanMeter A; Zhou W; Coukos G; Geho DH; Petricoin EF; Liotta LALaser-Capture Microdissection. Nat. Protoc2006, 1 (2), 586–603. [PubMed: 17406286]

(2). Buchberger AR; DeLaney K; Johnson J; Li LMass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. Anal. Chem2018, 90 (1), 240. [PubMed: 29155564]

(3). Ly A; Longuespée R; Casadonte R; Wandernoth P; Schwamborn K; Bollwein C; Marsching C; Kriegsmann K; Hopf C; Weichert W; Kriegsmann J; Schirmacher P; Kriegsmann M; Deininger S-OSite-to-Site Reproducibility and Spatial Resolution in MALDI–MSI of Peptides from Formalin-Fixed Paraffin-Embedded Samples. PROTEOMICS – Clin. Appl2019, 13 (1), 1800029.

(4). Théron L; Centeno D; Coudy-Gandilhon C; Pujos-Guillot E; Astruc T; Rémond D; Barthelemy J-C; Roche F; Feasson L; Hébraud M; Béchet D; Chambon CA Proof of Concept to Bridge the Gap between Mass Spectrometry Imaging, Protein Identification and Relative Quantitation: MSI~LC-MS/MS-LF. Proteomes 2016, Vol. 4, Page 322016, 4 (4), 32.

(5). Yin B; Mendez R; Zhao X-Y; Rakhit R; Hsu K-L; Ewald SEAutomated Spatially Targeted Optical Microproteomics (AutoSTOMP) to Determine Protein Complexity of Subcellular Structures. Anal. Chem2020, 92 (2), 2005–2010. [PubMed: 31869197]

(6). Chen C-L; Perrimon NProximity-Dependent Labeling Methods for Proteomic Profiling in Living Cells. Wiley Interdiscip. Rev. Dev. Biol2017, 6 (4), 10.1002/wdev.272.

(7). Branon TC; Bosch JA; Sanchez AD; Udeshi ND; Svinkina T; Carr SA; Feldman JL; Perrimon N; Ting AYEfficient Proximity Labeling in Living Cells and Organisms with TurboID. Nat. Biotechnol2018, 36 (9), 880–887. [PubMed: 30125270]

(8). Trinkle-Mulcahy LRecent Advances in Proximity-Based Labeling Methods for Interactome Mapping [Version 1; Referees: 2 Approved]. F1000Research 2019, 8 (0).

(9). Bar DZ; Atkatsh K; Tavarez U; Erdos MR; Gruenbaum Y; Collins FSBiotinylation by Antibody Recognition—a Method for Proximity Labeling. Nat. Methods2017, 15, 127. [PubMed: 29256494]

(10). Barrett AS; Wither MJ; Hill RC; Dzieciatkowska M; D'Alessandro A; Reisz JA; Hansen KCHydroxylamine Chemical Digestion for Insoluble Extracellular Matrix Characterization. J. Proteome Res2017, 16 (11), 4177–4184. [PubMed: 28971683]

(11). Dellon ES; Gonsalves N; Hirano I; Furuta GT; Liacouras CA; Katzka DAACG Clinical Guideline: Evidenced Based Approach to the Diagnosis and Management of Esophageal Eosinophilia and Eosinophilic Esophagitis (EoE). Am. J. Gastroenterol2013, 108 (5), 679–692; quiz 693. [PubMed: 23567357]

(12). Ostasiewicz P; Zielinska DF; Mann M; Wisniewski JRProteome, Phosphoproteome, and N-Glycoproteome Are Quantitatively Preserved in Formalin-Fixed Paraffin-Embedded Tissue and Analyzable by High-Resolution Mass Spectrometry. J. Proteome Res2010, 9 (7), 3688–3700. [PubMed: 20469934]

(13). Wiśniewski JRProteomic Sample Preparation from Formalin Fixed and Paraffin Embedded Tissue. J. Vis. Exp2013, No. 79, 5–9.

(14). Tyanova S; Temu T; Cox JThe MaxQuant Computational Platform for Mass Spectrometry-Based Shotgun Proteomics. Nat. Protoc2016, 11 (12), 2301–2319. [PubMed: 27809316]

(15). van der Maaten L; Hinton GVisualizing Data Using T-SNE. J. Mach. Learn. Res2008, 9 (86), 2579–2605.

(16). Tyanova S; Temu T; Sinitcyn P; Carlson A; Hein MY; Geiger T; Mann M; Cox JThe Perseus Computational Platform for Comprehensive Analysis of (Prote)Omics Data. Nat. Methods2016, 13 (9), 731–740. [PubMed: 27348712]

(17). Benninger RKP; Piston DWTwo-Photon Excitation Microscopy for the Study of Living Cells and Tissues. Curr. Protoc. cell Biol2013, Chapter4, Unit-4.11.24.

(18). Cox J; Hein MY; Luber CA; Paron I; Nagaraj N; Mann MAccurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol. Cell. Proteomics2014, 13 (9), 2513–2526. [PubMed: 24942700]

(19). Chistiakov DA; Killingsworth MC; Myasoedova VA; Orekhov AN; Bobryshev YVCD68/Macrosialin: Not Just a Histochemical Marker. Lab. Investig2017, 97 (1), 4–13.

(20). Clausen BE; Burkhardt C; Reith W; Renkawitz R; Förster IConditional Gene Targeting in Macrophages and Granulocytes Using LysMcre Mice. Transgenic Res. 1999, 8 (4), 265–277. [PubMed: 10621974]

(21). Abate W; Alrammah H; Kiernan M; Tonks AJ; Jackson SKLysophosphatidylcholine Acyltransferase 2 (LPCAT2) Co-Localises with TLR4 and Regulates Macrophage Inflammatory Gene Expression in Response to LPS. Sci. Rep2020, 10 (1), 10355. [PubMed: 32587324]

(22). Lubbers R; van Essen MF; van Kooten C; Trouw LAProduction of Complement Components by Cells of the Immune System. Clin. Exp. Immunol2017, 188 (2), 183–194. [PubMed: 28249350]

(23). Dick SA; Macklin JA; Nejat S; Momen A; Clemente-Casares X; Althagafi MG; Chen J; Kantores C; Hosseinzadeh S; Aronoff L; Wong A; Zaman R; Barbu I; Besla R; Lavine KJ; Razani B;

Ginhoux F; Husain M; Cybulsky MI; Robbins CS; Epelman SSelf-Renewing Resident Cardiac Macrophages Limit Adverse Remodeling Following Myocardial Infarction. Nat. Immunol2019, 20 (1), 29–39. [PubMed: 30538339]

(24). Huang DW; Sherman BT; Lempicki RASystematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. Nat. Protoc2009, 4 (1), 44–57. [PubMed: 19131956]

(25). Burnap SA; Mayr MExtracellular Vesicle Crosstalk between the Myocardium and Immune System upon Infarction. Circulation Research. Lippincott Williams and Wilkins2018, pp 15–17.

(26). Subramanian A; Tamayo P; Mootha VK; Mukherjee S; Ebert BL; Gillette MA; Paulovich A; Pomeroy SL; Golub TR; Lander ES; Mesirov JPGene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. Proc. Natl. Acad. Sci2005, 102 (43), 15545 LP – 15550. [PubMed: 16199517]

(27). Fabregat A; Sidiropoulos K; Viteri G; Marin-Garcia P; Ping P; Stein L; D'Eustachio P; Hermjakob HReactome Diagram Viewer: Data Structures and Strategies to Boost Performance. Bioinformatics2018, 34 (7), 1208–1214. [PubMed: 29186351]

(28). Darling TK; Lamb TJEmerging Roles for Eph Receptors and Ephrin Ligands in Immunity. Front. Immunol2019, 10 (JULY), 1–15. [PubMed: 30723466]

(29). Massol P; Montcourrier P; Guillemot JC; Chavrier PFc Receptor-Mediated Phagocytosis Requires CDC42 and Rac1. EMBO J. 1998, 17 (21), 6219–6229. [PubMed: 9799231]

(30). Hanna SJ; McCoy-Simandle K; Miskolci V; Guo P; Cammer M; Hodgson L; Cox DThe Role of Rho-GTPases and Actin Polymerization during Macrophage Tunneling Nanotube Biogenesis. Sci. Rep2017, 7 (1), 1–16. [PubMed: 28127051]

(31). Eriksson U; Kurrer MO; Sebald W; Brombacher F; Kopf MDual Role of the IL-12/IFN-γ Axis in the Development of Autoimmune Myocarditis: Induction by IL-12 and Protection by IFN-γ. J. Immunol2001, 167 (9), 5464 LP – 5469. [PubMed: 11673566]

(32). Tait Wojno ED; Hunter CA; Stumhofer JSThe Immunobiology of the Interleukin-12 Family: Room for Discovery. Immunity2019, 50 (4), 851–870. [PubMed: 30995503]

(33). O'Shea KM; Aceves SS; Dellon ES; Gupta SK; Spergel JM; Furuta GT; Rothenberg MEPathophysiology of Eosinophilic Esophagitis. Gastroenterology2018, 154 (2), 333–345. [PubMed: 28757265]

(34). Clayton F; Fang JC; Gleich GJ; Lucendo AJ; Olalla JM; Vinson LA; Lowichik A; Chen X; Emerson L; Cox K; O'Gorman MA; Peterson KAEosinophilic Esophagitis in Adults Is Associated with IgG4 and Not Mediated by IgE. Gastroenterology2014, 147 (3), 602–609. [PubMed: 24907494]

(35). Schuyler AJ; Wilson JM; Tripathi A; Commins SP; Ogbogu PU; Kruzsewski PG; Barnes BH; McGowan EC; Workman LJ; Lidholm J; Rifas-Shiman SL; Oken E; Gold DR; Platts-Mills TAE; Erwin EASpecific IgG(4) Antibodies to Cow's Milk Proteins in Pediatric Patients with Eosinophilic Esophagitis. J. Allergy Clin. Immunol2018, 142 (1), 139–148.e12. [PubMed: 29678750]

(36). Mishra ASignificance of Mouse Models in Dissecting the Mechanism of Human Eosinophilic Gastrointestinal Diseases (EGID). J. Gastroenterol. Hepatol. Res2013, 2 (11), 845–853. [PubMed: 25866707]

(37). Acharya KR; Ackerman SJEosinophil Granule Proteins: Form and Function. J. Biol. Chem2014, 289 (25), 17406–17415. [PubMed: 24802755]

(38). Spiteri MA; Bianco A; Strange RC; Fryer AAPolymorphisms at the Glutathione S-Transferase, GSTP1 Locus: A Novel Mechanism for Susceptibility and Development of Atopic Airway Inflammation. Allergy2000, 55Suppl 6, 15–20.

(39). Singh NK; Rao GNEmerging Role of 12/15-Lipoxygenase (ALOX15) in Human Pathologies. Prog. Lipid Res2019, 73, 28–45. [PubMed: 30472260]

(40). Fourie AMModulation of Inflammatory Disease by Inhibitors of Leukotriene A4 Hydrolase. Curr. Opin. Investig. Drugs2009, 10 (11), 1173–1182.

(41). Soto-Heredero G; Gómez de Las Heras MM; Gabandé-Rodríguez E; Oller J; Mittelbrunn MGlycolysis - a Key Player in the Inflammatory Response. FEBS J2020, 287 (16), 3350–3369. [PubMed: 32255251]

(42). Rosekrans SL; Baan B; Muncan V; Van Den Brink GREsophageal Development and Epithelial Homeostasis. Am. J. Physiol. - Gastrointest. Liver Physiol2015, 309 (4), G216–G228. [PubMed: 26138464]

(43). Chang Q; Ornatsky OI; Siddiqui I; Loboda A; Baranov VI; Hedley DWImaging Mass Cytometry. Cytom. Part A2017, 91 (2), 160–169.

(44). Ptacek J; Locke D; Finck R; Cvijic M-E; Li Z; Tarolli JG; Aksoy M; Sigal Y; Zhang Y; Newgren M; Finn JMultiplexed Ion Beam Imaging (MIBI) for Characterization of the Tumor Microenvironment across Tumor Types. Lab. Investig. 2020 10082020, 100 (8), 1111–1123.

(45). Hümmert MW; Alvermann S; Gingele S; Gross CC; Wiendl H; Mirenska A; Hennig C; Stangel MImmunophenotyping of Cerebrospinal Fluid Cells by Chipcytometry. J. Neuroinflammation2018, 15 (1).

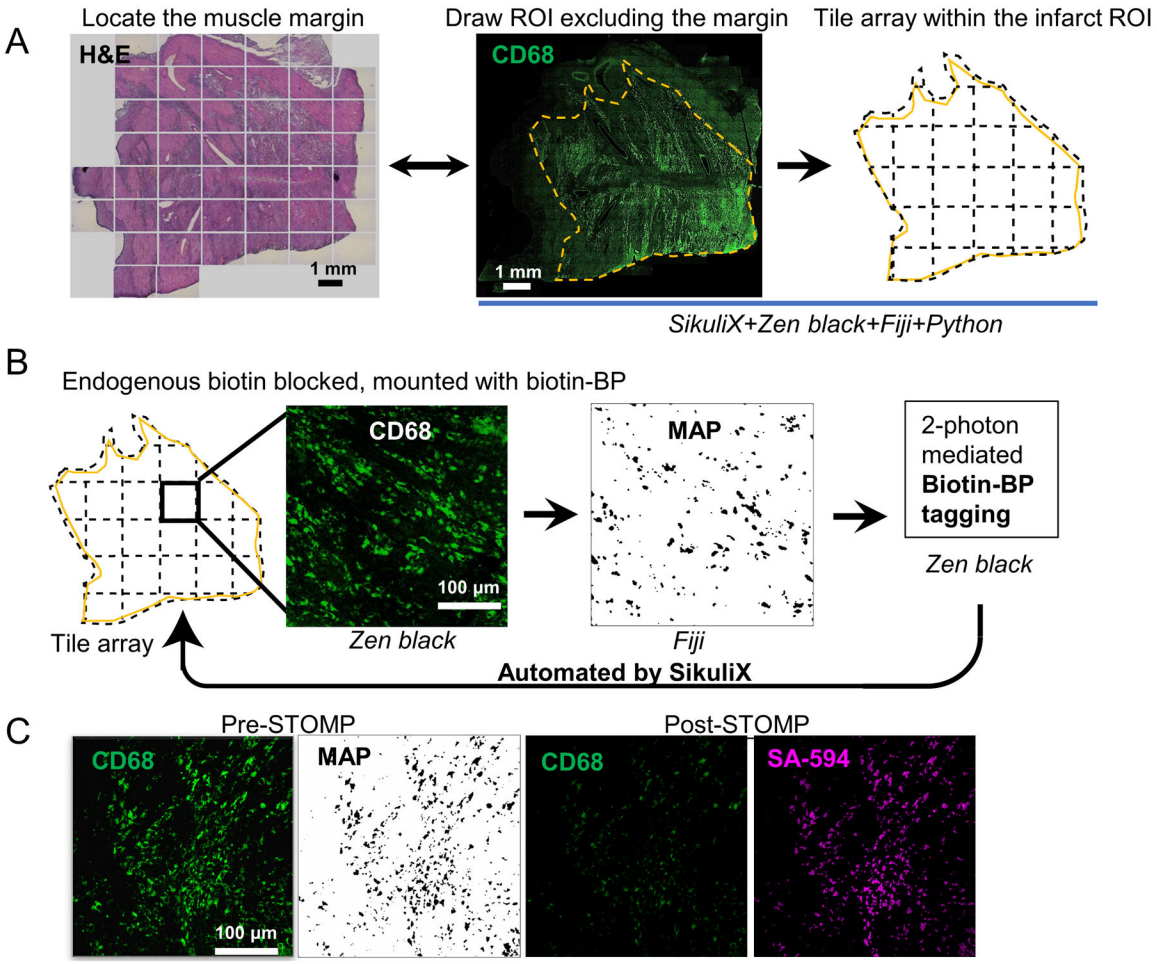


Figure 1. AutoSTOMP identifies rat cardiac infarct borders and biotinylates CD68-associated proteins in tissue sections.

A, Serial sections of cardiac infarcts were stained with Hematoxylin and Eosin (H&E, left) to identify the general border of the infarct relative to health tissue or the macrophage marker CD68 (green). Infarct borders were defined (yellow) and divided into tiles. **B**, The AutoSTOMP workflow integrates the jobs repeated in tiles. In Zeiss Zen Black the CD68 signal is imaged on each tile, exported to FIJI where it is thresholded and used to generate a MAP file identifying the pixel coordinates of the $CD68^+$ positive SOI. The MAP file is imported into Zen Black and directs the two-photon to target biotin-BP to the CD68+ SOI. This is automated across the scar. Each field of view measures 340 μm × 340 μm. A typical section contains approximately 550 tiles. **C**, to validate the selective biotinylation of $CD68^+$ SOI (Pre-STOMP) slides were washed, stained with streptavidin-594 then reimaged (Post-STOMP) to assess CD68 (note photobleaching) and streptavidin co-localization.

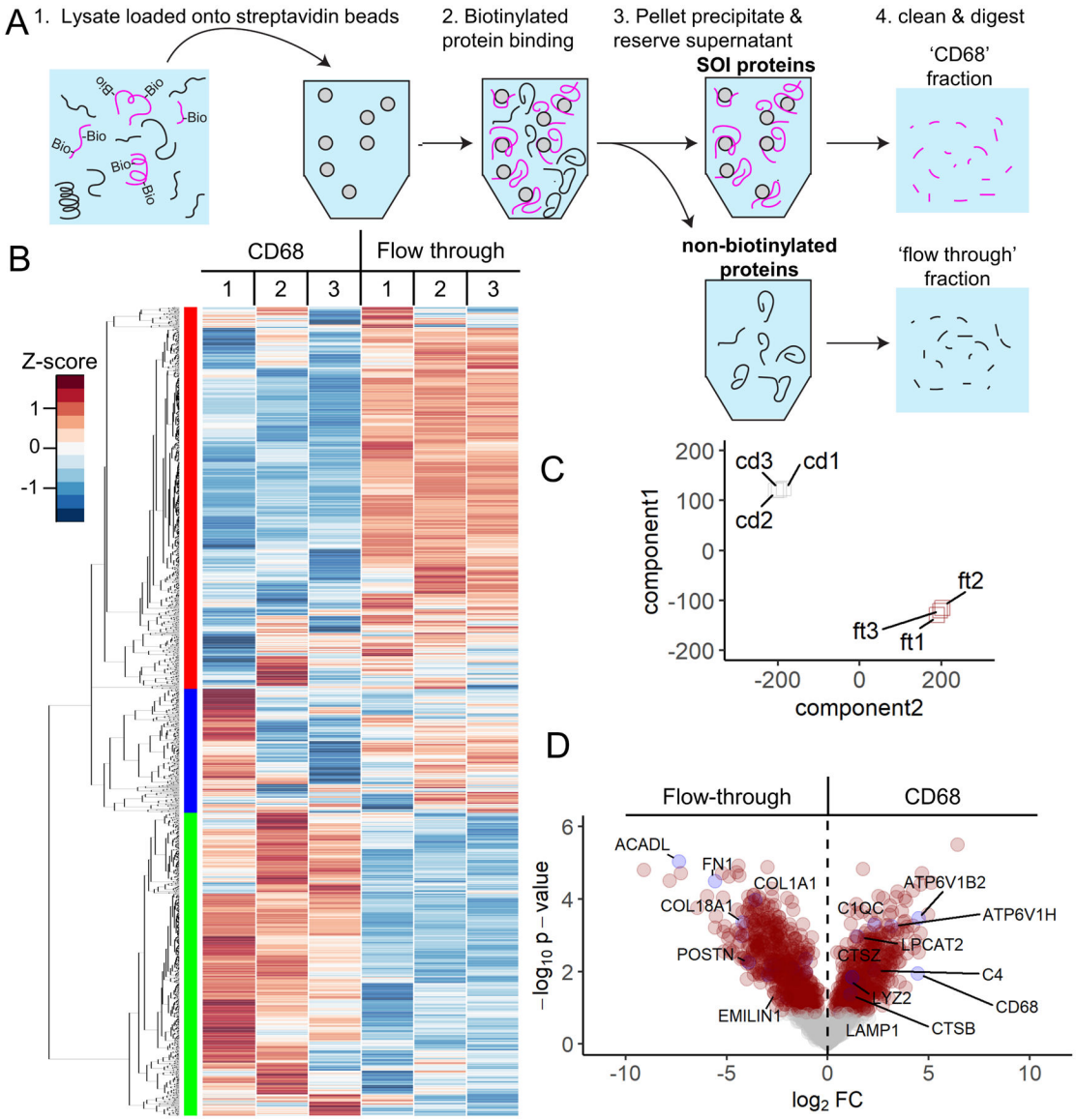


Figure 2. AutoSTOMP selectively enriches macrophage-associated proteins in the CD68$^+$ regions of rat cardiac infarcts.

CD68$^+$ regions of the cardiac infarcts are biotinylated by AutoSTOMP as described in Figure 1. **A**, After AutoSTOMP crosslinking, coverslips were washed of unconjugated Biotin-BP and lysed (1). Biotinylated proteins ('CD68' fraction) were bound to streptavidin beads (2) and pelleted (3). Unbound proteins were reserved as a 'flow through' control (3). Both fractions are trypsin/LysC digested and analyzed by LC-MS(4). **B-D** Protein abundance was determined by MaxQuant label free quantification (MaxLFQ). **B**, Of the 1,671 rat proteins identified, 94.2% were observed with one or more valid readouts in each of the three 'CD68' and/or 'flow through' fraction replicates. Heat map represents z-score for each protein across all the samples. Hierarchical clustering (left) indicates three enrichment patterns. **C**, T-distributed stochastic neighborhood embedding (t-SNE) analysis of variation among the 'CD68' fractions (cd1, cd2, cd3) and 'flow through' fractions (ft1, ft2, ft3) belonging to 3 paired replicates. **D**, Of the 1,671 proteins identified, 28.2% of were

significantly enriched and 33.7% were significantly lower in the ‘CD68’ fractions relative to the ‘flow through’ (red circles $p < 0.05$). Plotted as $-\log_{10}$ p-value (y axis) versus the $\log_2$ fold-change (x axis) of the protein abundance averaged between replicates with a false discovery of < 0.1. Some significantly expressed proteins annotated are in blue dots.

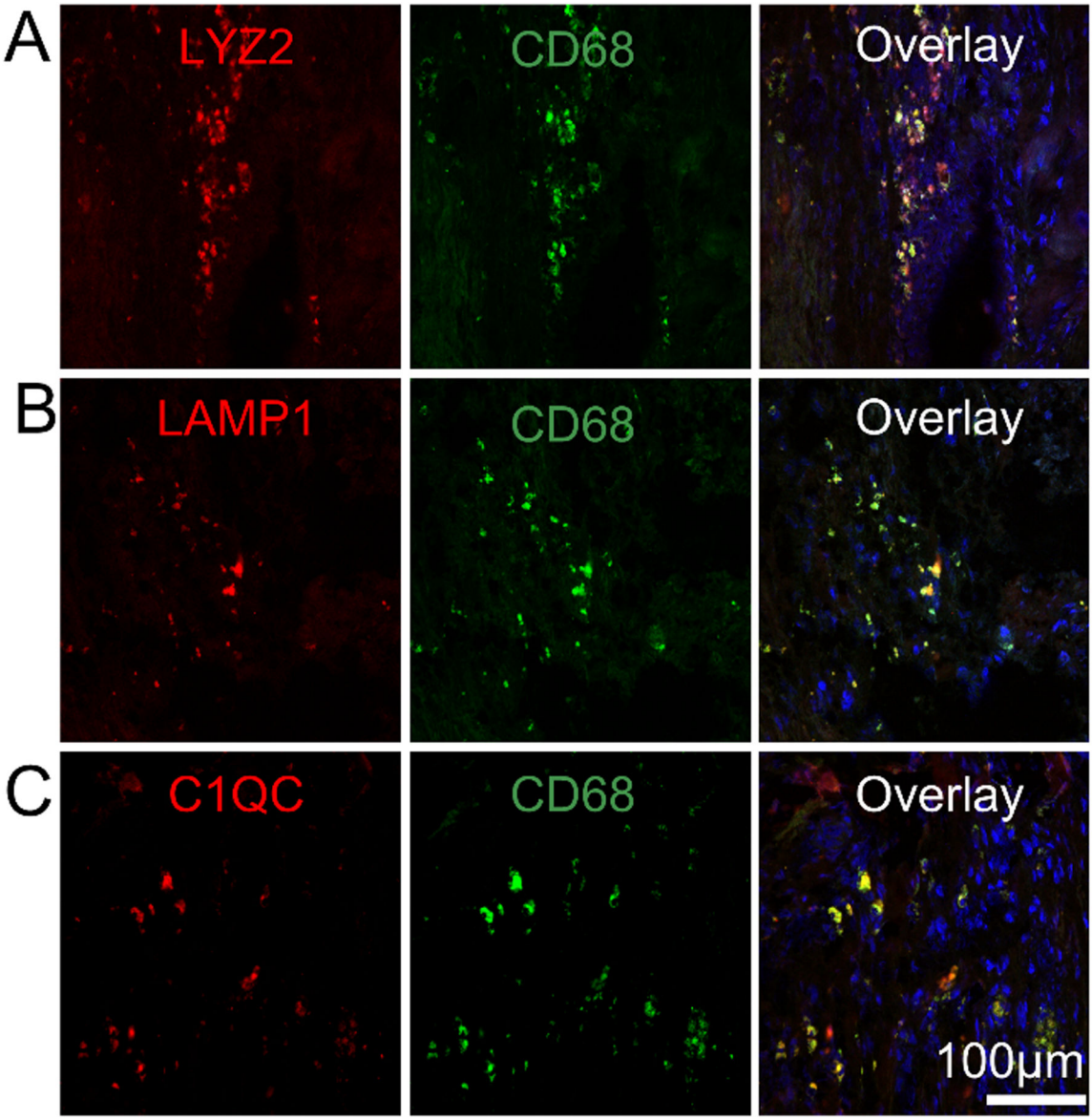


Figure 3. CD68^{+} macrophage partially co-localize with LYZ2, LAMP1 and C1QC in rat cardiac infarcts.

1 week rat cardiac infarcts were stained for CD68 (green) as described in Figure 1 and antibodies specific to the macrophage marker LYZ2 (A, red), the lysosomal protein LAMP1 (B, red) and the complement protein C1QC (C, red). Samples were counter stained with DAPI. N=3.

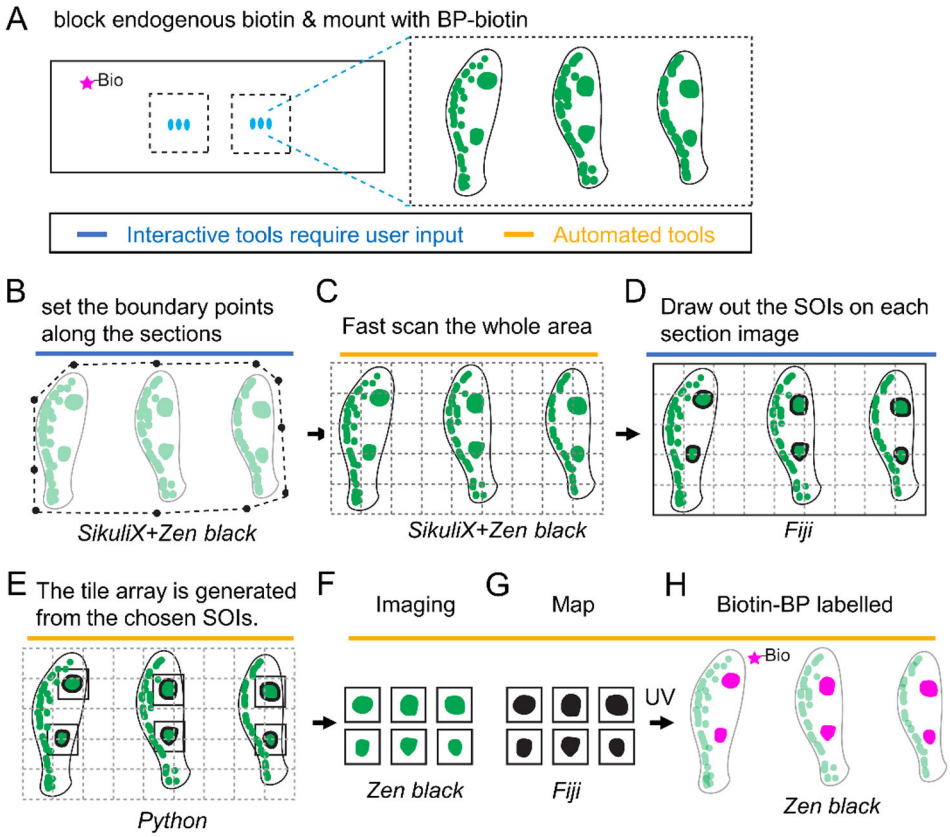


Figure 4. Schematic of autoSTOMP mediated targeting of discrete IgG4$^+$ lesions across multiple esophageal biopsy sections.

1 mm esophagus punch biopsies are isolated from a patient with active Eosinophilic Esophagitis (EoE). **A**, 6–8 sections per slide are stained for IgG4 (green), endogenous biotins are blocked and mounted with BP-biotin. A typical section contains 5 tiles at 340 μm × 340 μm per field of view using a 25x magnification objective lens. **B–H**, The SikuliX script allows the user to set the boundary points for all sections in Zen Black (**B**). **C**, The boundary points then direct low resolution tile scan in Zen Black. **D**, IgG4$^+$ structures of interest (SOI) are identified by the user on each section in Fiji. **E**, A Python script maps the coordinates of each SOI to generate a tile array of SOI-containing fields of view. **F–H**, A SikuliX script automates SOI imaging in Zen black (**F**, green), MAP file generation in Fiji (**G**, black) and BP-biotin crosslinking in Zen Black (**H**, magenta) as described in Figure 1.

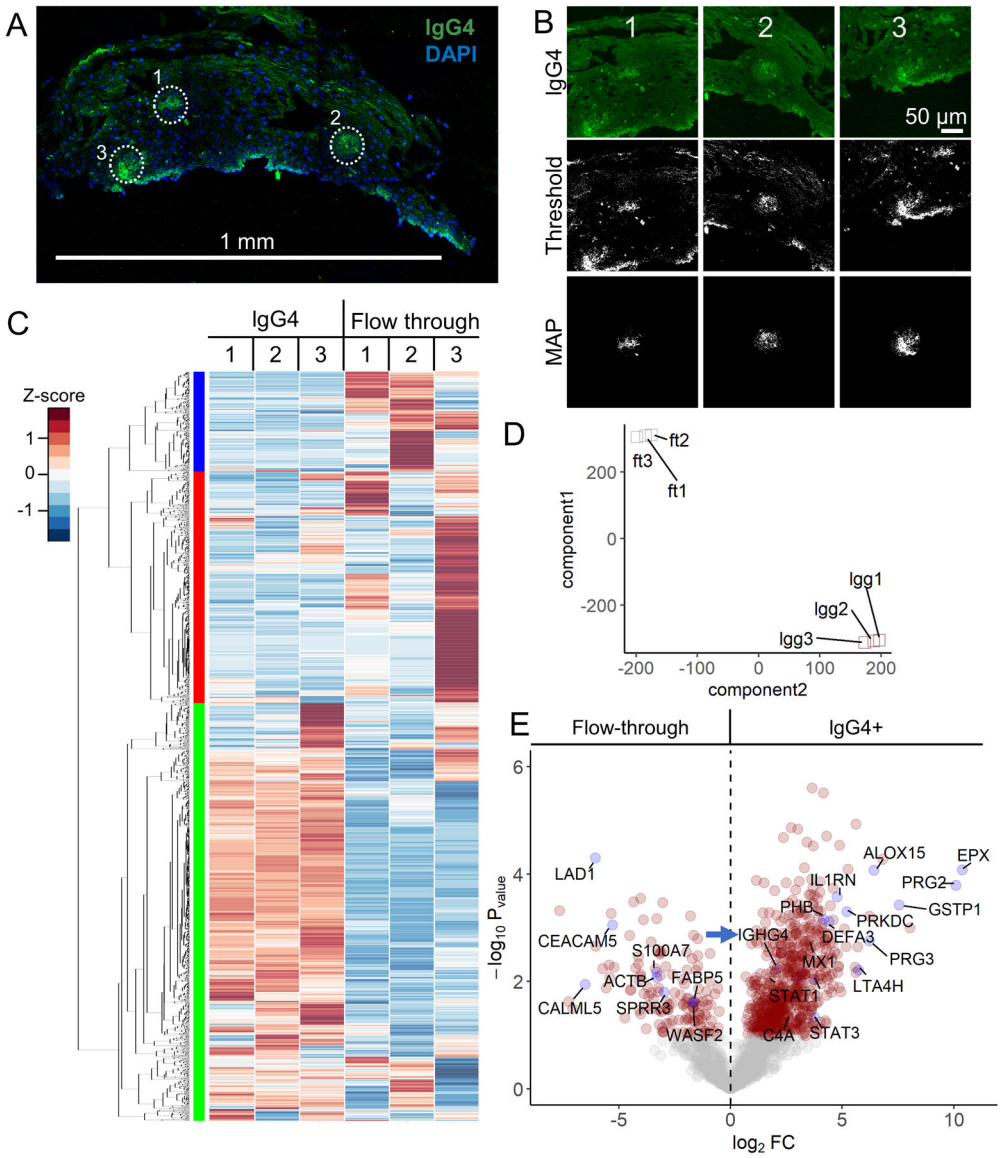


Figure 5. AutoSTOMP identifies eosinophil granule proteins associated with IgG4$^+$ lesions in EoE patient esophagus biopsies.

A, Representative section of an EoE patient biopsy stained with an antibody specific to lgG4 (green) and counterstained with DAPI (blue). IgG4$^+$ SOI are indicated by white dots, scale bar = 1 mm. **B**, Three representative SOI tile images, thresholded images and MAP files used to guide biotinylation of each SOI. **C–D**, Biotinylated proteins, the ‘lgG4’ fraction, or the unlabeled ‘flow through’ were isolated and identified as described in Figure 2A. Each sample represents sections pooled from two biopsies. **C**, 2,007 human proteins were identified and plotted by row z-score normalized across all the samples. Hierarchical clustering (left) indicates three enrichment groups. **D**, T-distributed stochastic neighborhood embedding (t-SNE) account for variation among the ‘lgG4’ fractions (lgg1, lgg2, lgg3) and ‘flow through’ fractions (ft1, ft2, ft3) across the three samples. **E**, Of the 2,007 proteins identified, 27.9% of were significantly enriched and 12.3% of the proteins were significantly lower in the ‘lgG4’ fractions relative to the ‘flow through’ fractions (red circles, $p < 0.05$,

FDR < 0.1). Plotted as $-\log_{10}$ p-value (y axis) versus the $\log_2$ fold-change (x axis) of the protein abundance averaged between replicates with a false discovery of < 0.1. IgG4 positive control (IGHG4, blue arrow indicated) is noted. Some significantly expressed proteins annotated are in blue dots.

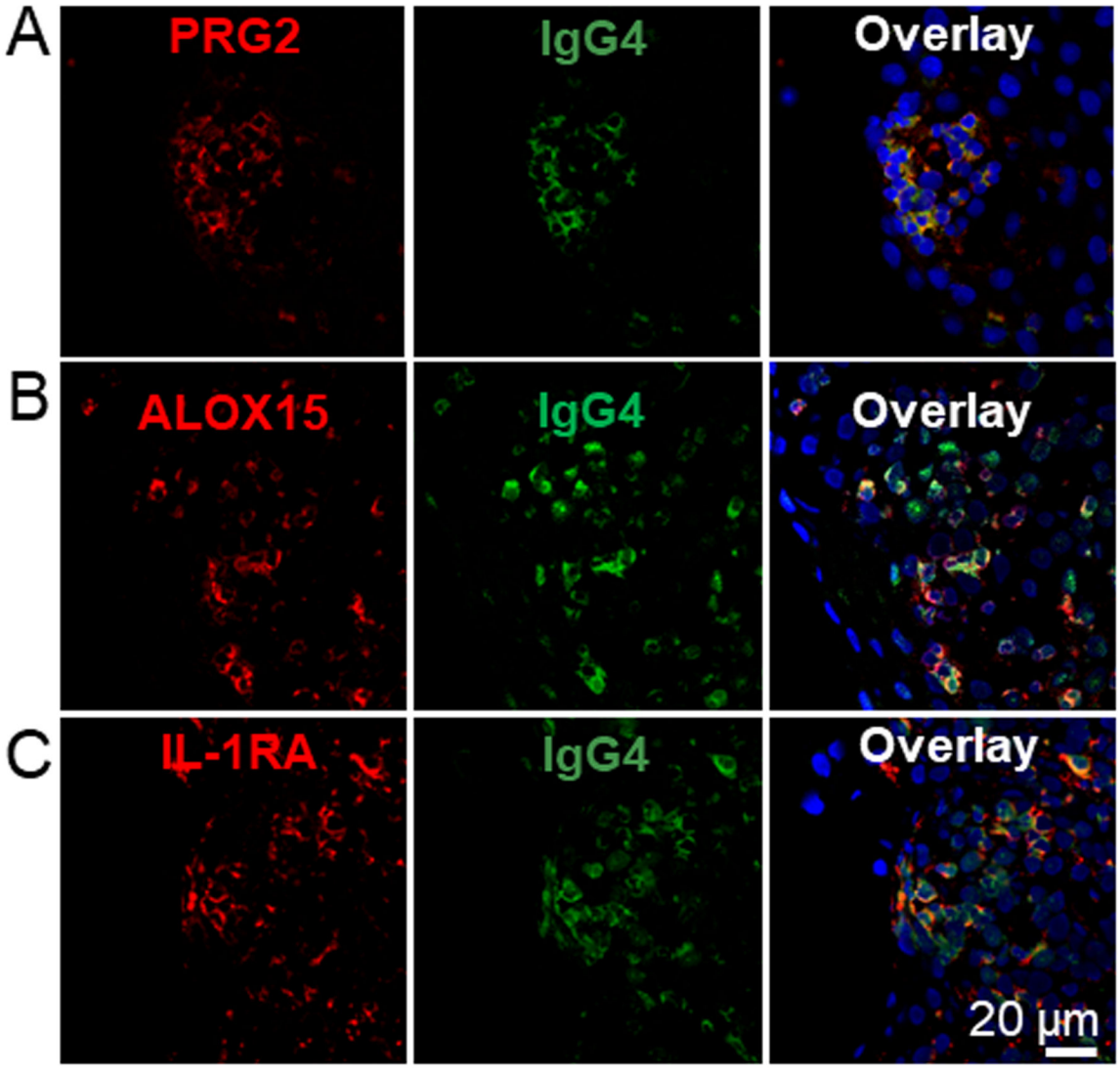


Figure 6. IgG4+ EoE lesions partially co-localize with inflammatory markers.

EoE patient biopsy sections were stained for IgG4 (green) as described in Figure 5 and antibodies specific to the proteins PRG2 (A, red), ALOX15 (B, red) and IL-1RA (C, red). Samples were counter stained with DAPI. N=3.