Published in final edited form as:

J Proteome Res. 2018 February 02; 17(2): 759–769. doi:10.1021/acs.jproteome.7b00775.

BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation

Dae In Kim[†,+], **Jevon A. Cutler**[†,‡,§,+], **Chan Hyun Na**[†,‖], **Sina Reckel**[⊥], **Santosh Renuse**[†,‖], **Anil K. Madugundu**[†,#], **Raiha Tahir**[§,▽], **Hana L. Goldschmidt**[⬡], **Karen L. Reddy**[§,◆], **Richard L. Huganir**[⬡], **Xinyan Wu**[†,§], **Natasha E. Zachara**[§], **Oliver Hantschel**[⊥], and **Akhilesh Pandey**[*,†,§,‖,¶]

[†]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States [‡]Pre-Doctoral Training Program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States [§]Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States [‖]Center for Proteomics Discovery, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States [⊥]Swiss Institute for Experimental Cancer Research (ISREC), School of Life Sciences, École polytechnique fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland [#]Institute of Bioinformatics, International Technology Park, Bangalore 560066, India [▽]Biochemistry, Cellular and Molecular Biology Graduate Program, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States [⬡]Solomon H. Snyder Department of Neuroscience, Kavli Neuroscience Discovery Institute, Johns Hopkins University School of Medicine, Baltimore,

[*]**Corresponding Author:** Tel: +1-410-502-6662, pandey@jhmi.edu.
[+]D.I.K. and J.A.C. contributed equally to this work.

ORCID

Akhilesh Pandey: 0000-0001-9943-6127

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI:10.1021/acs.jproteo-me.7b00775. Supplementary Figure 1. Evaluation of anti-biotin antibodies and comparison to other strategies for biotinylated peptide enrichment. Supplementary Figure 2. Mass spectra from quantitative BioSITe experiment. Supplementary Figure 3. Comparison between IMS proteins identified by BioSITe and previous studies. (PDF)

Supplemental Table 1. List of biotinylated proteins and peptides identified by antibody 1 and antibody 2 in BirA*-BCR-ABL Ba/F3 cells. Supplemental Table 2. List of biotinylated proteins and peptides identified by conventional on-bead digestion experiment in BirA*-BCR-ABL Ba/F3 cells. Supplemental Table 3. List of biotinylated proteins and peptides identified by conventional on-bead digestion experiment in BirA*-BCR-ABL Ba/F3 cells. Supplemental Table 4. Biotinylated proteins and peptides identified by elution after on-bead digestion experiment in BirA*-BCR-ABL Ba/F3 cells. Supplemental Table 5. Biotinylated proteins and peptides identified by DiDBiT in BirA*-BCR-ABL Ba/F3 cells. Supplemental Table 6. List of quantitative analysis of biotinylated proteins and peptides identified by BioSITe in BirA*-BCR-ABL p190 and BirA*-BCR-ABL p210 Ba/F3 cells. Supplemental Table 7. List of biotin-phenol labeled proteins and peptides identified by BioSITe in HEK293 cells expressing NES-APX2. Supplemental Table 8. List of Biotin-phenol labeled proteins and peptides identified by BioSITe in HEK293 cells expressing IMS-APX2. Supplemental Table 9. List of O-GlcNAc-modified proteins and peptides identified by BioSITe. (XLSX)

Author Contributions

D.I.K., J.A.C, and A.P. designed research; D.I.K., J.A.C., S. Renuse., R.T., and H.L.G. performed experiments; D.I.K., J.A.C., C.H.N., S. Renuse., A.K.M., K.L.R, R.L.H., X.W., N.E.Z., and A.P. analyzed data; S. Reckel. and O.H. analyzed structural data; D.I.K., J.A.C., and A.P. wrote the paper. All authors edited the paper.

The authors declare no competing financial interest.

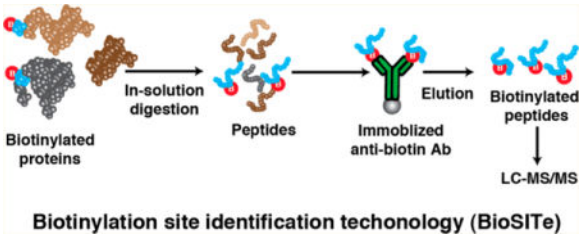
All mass spectrometry data and search results have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD007862) via the PRIDE partner repository with the data set identifier PXD007862 and project name ‘BioSITe: A novel method for direct detection of site specific biotinylation’.

Maryland 21205, United States ◆Center for Epigenetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States ¶Departments of Pathology and Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States

Abstract

Biotin-based labeling strategies are widely employed to study protein-protein interactions, subcellular proteomes and post-translational modifications, as well as, used in drug discovery. While the high affinity of streptavidin for biotin greatly facilitates the capture of biotinylated proteins, it still presents a challenge, as currently employed, for the recovery of biotinylated peptides. Here we describe a strategy designated Biotinylation Site Identification Technology (BioSITe) for the capture of biotinylated peptides for LC−MS/MS analyses. We demonstrate the utility of BioSITe when applied to proximity-dependent labeling methods, APEX and BioID, as well as biotin-based click chemistry strategies for identifying O-GlcNAc-modified sites. We demonstrate the use of isotopically labeled biotin for quantitative BioSITe experiments that simplify differential interactome analysis and obviate the need for metabolic labeling strategies such as SILAC. Our data also highlight the potential value of site-specific biotinylation in providing spatial and topological information about proteins and protein complexes. Overall, we anticipate that BioSITe will replace the conventional methods in studies where detection of biotinylation sites is important.

Graphical abstract



Keywords

protein−protein interactions; subcellular proteome; peptide; biotinylation; proximity-dependent biotinylation; BioID; APEX

INTRODUCTION

Mapping protein−protein interactions (PPIs) and subcellular proteomes is an important strategy to reveal biological phenomena and mechanisms of disease. Newly developed proximity-dependent biotinylation methods such as APEX and BioID facilitate the discovery of PPIs and subcellular proteomes within living cells. In these methods, cellular expression of a protein of interest (bait) fused with an engineered enzyme leads to proximity-dependent biotinylation of proteins of interest.[1] Both of these methods employ exogenously provided biotin or its analog to serve as molecular probes used by the enzymes. The APEX system leverages an engineered ascorbate peroxidase enzyme designated APEX2 that uses hydrogen

peroxide to catalyze the transfer of biotin-phenol to tyrosine residues in proteins.[2,3] BioID similarly relies on an engineered biotin ligase called BirA*[4,5] to generate biotinoyl-5′-AMP, which is reactive with lysine residues of proteins.[6]

The identification of biotinylated proteins that are generated using these strategies has hitherto relied upon capture using the streptavidin family of proteins, followed by either on-bead digestion or in-gel digestion, followed by LC−MS/MS analysis. While these methods have been used successfully, they both depend on the use of identified nonbiotinylated peptides as surrogates for biotinylation, as biotinylated peptides are rarely, if at all, detected. The reason biotinylated peptides are rarely identified is primarily because of the strong affinity of streptavidin for biotin. When in-gel digestion is used, elution of biotinylated proteins is required prior to electrophoresis. Several elution methods have been developed; using detergents,[7] extremely low pH[8] or solvents[9] and alternative strategies have been employed including weakened affinity of streptavidin/avidin to biotin through chemical[10] or genetic[11] means. However, even these modifications only lead to partial recovery of biotinylated proteins from streptavidin,[12] and nonbiotinylated peptides always vastly outnumber biotinylated peptides.[2] The need for good elution strategies has been circumvented by the use of on-bead digestion, which has become a convenient and widely used method in proximity-dependent biotinylation experiments. However, biotinylated peptides remain bound to beads and are undetected as the streptavidin−biotin interaction is not disrupted. In a different approach, which does not suffer from the limitations just described, Schiapparelli et al. have described a strategy designated DiDBiT (Direct Detection of Biotin-containing Tags), where they first digested proteins into peptides and then enriched biotinylated peptides using NeutrAvidin.[13] Because this method was mostly tested using proteins biotinylated in vitro, we reasoned that the strong affinity of NeutrAvidin for biotin could still limit detection of biotinylated peptides, especially in the context of in vivo biotinylation such as in proximity-dependent biotinylation methods, where the overall level of biotinylation level might be low.

To improve the detection of biotinylated peptides, we have developed a strategy designated **Bio**tinylation **S**ite **I**denti®cation **Te**chnology (**BioSITe**), which is based on the use of anti-biotin antibodies to directly capture and identify biotinylated peptides in a single LC −MS/MS run. We show that detection of biotinylated peptides greatly increases the confidence of candidate protein identification in proximity-dependent biotinylation methods. By providing the site of biotinylation on proteins labeled in these methods, this approach also offers a new level of information about the structural aspects of PPIs and protein topology. Finally, we describe a simple approach for quantitative BioSITe experiments through the use of isotopically labeled biotin, obviating the use of metabolic-labeling-based quantitative strategies such as SILAC.[14] While this paper was being finalized for publication, a report on a similar strategy using an antibody approach for direct enrichment of biotinylated peptide was published.[15]

MATERIALS AND METHODS

Antibodies and Reagents

Anti-biotin antibody 1 (Abcam, no. ab53494), anti-biotin antibody 2 (Bethyl Laboratories, no. 150-109A), streptavidin-HRP (Abcam, no. ab7403), protein G beads (EMD Millipore, no. 16-266), high-capacity NeutrAvidin agarose (Thermo Fisher Scientific, no. 29202), biotin (Sigma-Aldrich, no. B4501), biotin-2′,2′,3′,3′-d_4 (Sigma-Aldrich, no. 809068), Lipofectamine 2000 (Thermo Fisher Scientific, no. 11668019), biotin-phenol (Iris Biotech, no. CDX-B0270-M100), sequencing-grade trypsin (Promega, no. V5113), GalT1 enzymatic labeling kit (Invitrogen, no. C33368), PNGase F (New England Biolabs, no. P0704S), Click-IT Biotin DIBO Alkyne (Thermo Fisher Scientific, no. C10412), and trypsin (Worthington Biochemical Corporation, no. LS003741) were used.

BioID and APEX Preparation and Lysis

BioID was performed essentially as described by Roux et al. with some minor modifications. After overnight incubation with 50 μM biotin, Ba/F3 cells expressing BirA*-BCR-ABL p190 were lysed in BioID lysis buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 0.4% SDS, 2% Triton X-100 with Halt protease inhibitors), followed by sonication (three rounds, duty cycle 30%, 20 s pulses). After cleaning, the samples were centrifuged at 16000*g* for 10 min, and equal volumes of 50 mM Tris were added to the BioID samples. Lysates were quantified by bicinchoninic acid (BCA) assay, and 10 mg of protein per replicate was incubated with 200 μL of high-capacity NeutrAvidin overnight at 4 C. After incubation, the bead slurry was washed with 2% SDS, 50 mM Tris three times, BioID lysis buffer three times, 50 mM Tris three times, and 50 mM triethylammonium bicarbonate (TEABC) three times. Cells used for experiments to enrich biotinylated proteins using anti-biotin antibodies were lysed in the same buffer used in the BioSITe strategy, and the biotinylated proteins were captured and eluted in the same manner as described for BioSITe experiments. The experiments were carried out in triplicate using 10 mg of proteins per replicate.

APEX was performed as described by Hung et al. APEX-IMS and APEX-NES were transiently transfected into HEK293 using Lipofectamine 2000. One day later, the transfected cells were incubated with 50 μM biotin-phenol for 30 min and then with 1 mM H_2O_2 for 1 min at room temperature. The APEX reaction was subsequently quenched by washing cells with quenching solution containing 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide in PBS. Cells were collected by scrapping and pelleted.

Peptide Preparation

For BioSITe and other peptide enrichment experiments, cells were lysed using a lysis buffer containing 50 mM TEABC and 8 M urea and sonication as described above. The protein concentration of samples was measured by BCA assay. Reduction and alkylation were performed by serial incubation with 10 mM DTT for 30 min and by 20 mM IAA for 30 min in the dark. Lysate was diluted to 2 M urea by adding three cell lysate volumes of 50 mM TEABC. The proteins were digested with trypsin (1:20 of trypsin to protein) at 37 °C overnight. The resulting tryptic peptides for both experiments were desalted using a Sep-PAK C_{18} and subsequently lyophilized.

For on-bead digest, NeutrAvidin beads were centrifuged and resuspended in 100 μL of 50 mM TEABC and 2 M UREA. Reduction and alkylation of captured proteins was performed by incubation with 10 mM dithiothreitol (DTT) for 60 min and by 20 mM iodoacetamide (IAA) in the dark for 30 min. The volume was then increased to 200 μL with 50 mM TEABC pH 8.0 containing 3 μg of sequencing grade trypsin. Overnight digestion carried out on a shaker maintained at 37 °C. The supernatant was collected and the beads were washed once with 50 μL of 50 mM TEABC, which was collected and pooled. The final 450 μL was acidified to 1% trifluoroacetic acid, spun at 18000*g*, and desalted with a homemade reversed-phase C_{18} column following stage-protocol previously published methods. In experiments designed to enrich biotinylated proteins using anti-biotin antibodies, after elution, the samples were dried and resuspended in 2 M urea in 50 mM TEABC and digested with 3 μg of trypsin and desalted using a C_{18} column.

Direct Detection of Biotin-Containing Tags

This procedure was carried out as described by Schiapparelli et al. Ten mg of peptides generated from cells expressing BirA*-BCR-ABL p190 was dissolved in 1 mL of PBS and incubated with 90 μL of high-capacity NeutrAvidin beads for 1 h at room temperature. Bead slurry was sequentially washed with PBS three times, 5% acetonitrile three times, and ultrapure water three times. Captured peptides were eluted by adding the elution buffer containing 0.2% trifluoroacetic acid, 0.1% formic acid, and 80% acetonitrile. Peptides were eluted with and without boiling for 5 min for a total of 10 elutions, which were pooled and desalted using a desalting column as previously described.

BioSITe

For each replicate, 10 mg of total protein from BirA*-BCR-ABL cells was digested into peptides as described above. For APEX samples, total protein amount was estimated by cell number and digested into peptides. The amount of total peptide was then determined by peptide BCA assay to be 8 mg per replicate. The bead slurry was sequentially thoroughly washed. Biotinylated peptides were eluted four times. Anti-biotin antibodies (100 μg) were coupled to 120 μL of protein G bead slurry overnight at 4 °C. Antibody-coupled beads were further washed with PBS once and BioSITe capture buffer (50 mM Tris, 150 mM NaCl, 0.5% Triton X-100) twice. Peptides were dissolved in 1 mL of BioSITe capture buffer. After dissolving peptides, pH was adjusted to neutral (7.0 to 7.5) and peptide BCA was performed to estimate peptide concentration. Peptides were subsequently incubated with anti-biotin antibody-bound protein G beads for 2 h at 4 °C. The bead slurry was sequentially washed two times with BioSITe capture buffer, two times with 50 mL of Tris, and two times with ultrapure water. Biotinylated peptides were eluted four times using elution buffer (80% acetonitrile and 0.2% trifluoroacetic acid in water). The eluent was further cleaned up using C_{18} reversed-phase column as previously described.

BioSITe for Quantitative Proteomics

Peptides from the combined samples were subjected to BioSITe analysis. Cells expressing BirA*-BCR-ABL p190 or BirA*-BCR-ABL p210 were cultured overnight with 50 mM heavy or light biotin, respectively. The cells were lysed in equal amounts from each

condition and were mixed to generate a total of 10 mg of sample, which was digested, desalted, lyophilized, and subjected to BioSITe as described above.

Biotinylation of O-GlcNAc Sites Using Click Chemistry

HEK293 cells were treated with 500 nm thiamet-G for 4 h to enhance the overall O-GlcNAcylation levels by inhibiting O-GlcNAcase (OGA), an enzyme that removes O-GlcNAc. Peptides were generated, as previously described, and in-solution Lys-C digestion was carried out overnight at 37 °C. The peptide digest was then cleaned up using C_{18} solid-phase extraction cartridge. The cleaned peptides were then vacuum-dried and labeled with tandem mass tags. O-GlcNAc sites on these peptides were modified in vitro using GalT1 enzymatic labeling kit. In brief, peptides were reconstituted in 1.1 mL of 20 mM HEPES (pH 7.9) and 80 μL of 100 mM $MnCl_2$, 75 μL of 0.5 mM UDP-GalNAz, 62 μL of GalT1, and 5 μL of PNGase F. The mixture was then incubated overnight at 4 °C. The peptides were then incubated with 10 μM biotin DIBO alkyne at room temperature for 1 hour. The peptides were cleaned using C_{18} solid phase extraction (SPE) cartridge followed by strong-cation exchange SPE and vacuum dried. The peptides with biotinylated O-GlcNAc-modification were then enriched using BioSITe.

Mass Spectrometric Analysis

The fractionated peptides were analyzed on an Orbitrap Fusion Lumos Tribrid Mass spectrometer coupled to the Easy-nLC 1200 nanoflow liquid-chromatography system (Thermo Fisher Scientific). The peptides from each fraction were reconstituted in 20 μL of 0.1% formic acid and loaded onto an Acclaim PepMap 100 Nano-Trap Column (100 μm × 2 cm, Thermo Fisher Scientific) packed with 5 μm C_{18} particles at a flow rate of 4 μL per minute. Peptides were separated at 300 nL/min flow rate using a linear gradient of 7 to 30% solvent B (0.1% formic acid in 95% acetonitrile) over 95 min on an EASY-Spray column (50 cm × 75 μm ID, Thermo Fisher Scientific) packed with 2 μm C_{18} particles, which was fitted with an EASY-Spray ion source that was operated at a voltage of 2.3 kV.

Mass-spectrometry analysis was carried out in a data-dependent manner with a full scan in the mass-to-charge ratio (m/z) range of 300−18000 in the “Top Speed” setting, three seconds per cycle. MS and MS/MS were acquired for the precursor ion detection and peptide fragmentation ion detection, respectively. MS scans were measured at a resolution of 120000 at an m/z of 200. MS/MS scans were acquired by fragmenting precursor ions using the higher energy collisional dissociation (HCD) method and detected at a mass resolution of 30000, at an m/z of 200. Automatic gain control for MS was set to one million ions and that for MS/MS was set to 0.05 million ions. A maximum ion injection time was set to 50 ms for MS and 100 ms for MS/MS. MS was acquired in profile mode and MS/MS was acquired in centroid mode. Higher energy collisional dissociation was set to 32 for MS/MS. Dynamic exclusion was set to 35 seconds, and singly charged ions were rejected. Internal calibration was carried out using the lock mass option (m/z 445.1200025) from ambient air. Data acquisition of click-chemistry-modified O-GlcNAc modified peptides were carried out using alternate HCD/ETD (electron-transfer dissociation) method.

Post-Processing and Bioinformatics

Proteome Discoverer (v 2.1; Thermo Scientific) suite was used for quantitation and identification using all three replicate LC–MS/MS runs per experiment searched together. Spectrum selector was used to import spectrum from raw file. During MS/MS preprocessing, the top 10 peaks in each window of 100 m/z were selected for database search. The tandem mass spectrometry data were then searched using SEQUEST algorithm against protein databases (for BioID experiments: mouse NCBI RefSeq 73 (58039 entries) with the addition of fasta file entries for BCR-ABL p190 and the DH and PH domain of BCR-ABL p210; for APEX experiments: human NCBI RefSeq (73198 entries) with the addition of fasta file entries of IMS-APEX2 and NES-APEX2 constructs) with common contaminant proteins. The search parameters for identification of biotinylated peptides were as follows: (a) trypsin as a proteolytic enzyme (with up to three missed cleavages); (b) peptide mass error tolerance of 10 ppm; (c) fragment mass error tolerance of 0.02 Da; and (d) carbamido-methylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da) and biotinylation of lysine (+226.07759 Da) as variable modifications.

The search parameters for the identification of biotin-phenol modified peptides were as follows: (a) trypsin as a proteolytic enzyme (with up to two missed cleavages); (b) peptide mass error tolerance of 10 ppm; (c) fragment mass error tolerance of 0.02 Da; and (d) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da). Biotinylation of lysine (+226.07759 Da), biotin-phenol modification of tyrosine (+361.14601 Da), and oxidized-biotin-phenol modification of tyrosine (+377.141 Da) were all used as variable modifications. For the identification and quantification of the peptides modified by light or heavy biotin, all of the raw files from the three replicates were searched together. The search parameters for identification of either light or heavy biotinylated peptides were as follows: (a) trypsin as a proteolytic enzyme (with up to three missed cleavages); (b) peptide mass error tolerance of 10 ppm; (c) fragment mass error tolerance of 0.02 Da; and (d) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da), light biotinylation of lysine (+226.07759 Da), and heavy biotinylation of lysine (+230.103 Da) as variable modifications. The minimum peptide length was set to six amino acids. For the identification of click-chemistry-modified O-GlcNAcylated peptides, apart from oxidation of methionine and carbamido-methylation of cysteine, variable modification of click label (993.36 Da, i.e., HexNAc+GalNAz+DIBO alkyne biotin) on serine and threonine residues was included in the database search. Peptides and proteins were filtered at a 1% false-discovery rate (FDR) at the PSM level using percolator node and at the protein level using protein FDR validator node, respectively.

The protein quantification was performed with the following parameters and methods. The heavy-to-light ratios of the biotinylated peptides were measured by the Precursor Ions Quantifier node. Unique and razor peptides both were used for peptide quantification, while protein groups were considered for peptide uniqueness. Precursor ion abundance was computed based on intensity, and the missing intensity values were replaced with the minimum value. Protein grouping was performed with strict parsimony principle to generate

the final protein groups. All proteins sharing the same set or subset of identified peptides were grouped, while protein groups with no unique peptides were filtered out. The Proteome Discoverer iterated through all spectra and selected PSM with the highest number of unambiguous and unique peptides.

Identified protein and peptide spectral match (PSM) level data were exported as tabular files from Proteome Discoverer 2.1. We used an in-house Python script to compile the peptide level site information mapped to UniProt or RefSeq databases. The summary count on number of supported peptides and PSMs is then calculated at the protein level. In the case of the heavy-biotin experiment, the ratios are calculated from total channel intensities of identified peptides.

Protein Modeling

The degree and location of biotinylation as well as the protein domain organization were assigned based on information in UniProt (uniprot.org) taking into account the protein size, number of total lysines, and domain annotation according to PROSITE and InterPro databases. To map biotinylation sites on 3D structures, structural models were obtained from the Protein Data Bank (PDB, rcsb.org) as far as available or homology models were created using Swiss-Model (swissmodel.expasy.org).

Data Availability

All mass spectrometry data and search results have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD007862) via the PRIDE partner repository with the data set identifier PXD007862 and project name ‘BioSITe: A novel method for direct detection of site specific biotinylation’.

Experimental Design and Statistical Rationale

All experiments described in this study were performed as three independent replicates from the same starting material (process replicates). Because this is a new method, this replicate design was employed to account for sampling bias. A qualitative comparison of the identifications from the methods being compared was done in all cases except in the experiment involving heavy biotin where the median fold ratios were plotted.

RESULTS AND DISCUSSION

Detection of Biotinylated Peptides Using BioSITe

Proteins that are biotinylated by proximity-dependent biotinylation methods (Figure 1A) are generally captured by streptavidin-conjugated beads, followed by on-bead digestion (hereafter referred to as the conventional method) (Figure 1B, top). We reasoned that the strong affinity between biotin and streptavidin (K_d = ~10^{-15})[16] is a critical factor that limits elution and subsequent detection of biotinylated peptides. To overcome this we decided to use a capture reagent with a lower affinity, that is, a biotin-specific antibody, which would allow for dissociation of biotin and identification of biotinylated sites (Figure 1B, bottom). To evaluate the potential of BioSITe in improving the analysis of proximity labeling interactome studies, we leveraged our previously characterized Ba/F3 cell system that was

engineered to stably express the BCR-ABL oncogene (p190 variant) cloned in-frame with BirA*.[17] On the basis of our testing of a panel of commercially available antibodies against biotin, we chose two polyclonal antibodies (from Abcam and Bethyl Laboratories) that yielded high signal intensities from immunoprecipitation of biotinylated proteins, followed by Western blotting of whole cell lysates of BirA*-p190 (Supplementary Figure 1A). Tryptic peptides generated from these samples were separately enriched with each antibody and analyzed by LC–MS/MS (see Materials and Methods for details regarding the enrichment and MS analysis). Peptides containing biotinylated lysines were readily detected including a biotinylated peptide of SHC1 adapter protein 1 (SHC1), a direct binding partner of BCR-ABL, and its spectrum is shown in Figure 1C.[9,10] Lysine biotinylation was evidenced by MS/MS spectrum with three signature ion peaks as well as a series of fragmentation ions harboring biotin. The signature ions, m/z 227, m/z 310, and m/z 314, correspond to fragmented biotin, an immonium ion harboring biotin with a loss of NH_3 and an immonium ion harboring biotin, respectively.[4,5] Because we observed a partially overlapping pattern in terms of site identification, we used both antibodies together in subsequent BioSITe experiments (Supplemental Table 1) (Supplementary Figure 1B).

We next sought to compare the performance of BioSITe with other methods available for biotinylated protein or peptide capture. We analyzed triplicate samples from (i) whole protein capture followed by on-bead digestion, (ii) biotinylated peptides eluted from beads subsequent to on-bead digestion, (iii) NeutrAvidin-based biotinylated peptide capture method (DiDBiT), and (iv) whole protein capture with anti-biotin antibodies followed by elution and in-solution digestion. MS data from BioSITe experiments identified 3403 biotinylated peptides corresponding to 1193 proteins (Supplemental Table 2). On-bead digestion identified 11 biotinylated peptides (Supplemental Table 3), with no overlap of those biotinylated peptide identified by BioSITe (Figure 1D). Similarly, elution after on-bead digestion identified 55 biotinylated peptides (Supplementary Figure 1C and Supplemental Table 4), DiDBiT detected 924 biotinylated peptides (Supplementary Figure 1D and Supplemental Table 5), and whole protein capture with anti-biotin antibodies followed by elution and in-solution digestion identified only 44 biotinylated peptides (data not shown). Importantly, BioSITe led to identification of the largest number of biotinylated peptides of all methods tested in our study. On the basis of these findings, we conclude that ease of dissociating biotinylated peptides from anti-biotin antibodies combined with the peptide level enrichment is responsible for a higher number of identification using BioSITe. However, given that DiDBiT and BioSITe outperform both whole protein enrichment methods (21- to 85- and 77- to 309-fold increase in biotinylated peptide identification, respectively), we conclude that the peptide-level enrichment step provides the greatest increase in identification, while the choice of capture reagent (avidin analog or antibody) provides an additional ~4-fold increase in the number of biotinylated peptides that are identified.

When we compared the list of proteins whose biotinylated sites were detected by BioSITe with proteins identified from conventional on-bead digestion, we found a substantial overlap (718 proteins) that included well-documented interacting proteins such as SHC1, ABL interactor 1 (ABI1), ubiquitin-associated and SH3 domain-containing protein B (UBASH3B), phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1 (SHIP1), inositol

polyphosphate phosphatase like 1 (SHIP2), and growth factor receptor-bound protein 2 (GRB2) (Figure 1E).[17,18] There were 467 biotinylated proteins identified only by BioSITe, representing a subset of biotinylated proteins that were missed by conventional on-bead digestion (Figure 1E). Importantly, 23% (110/467) of these proteins have more than two sites of biotinylation; however, the majority, 77% (359/467), of these proteins are represented by only one site of biotinylation, indicating that these could represent low abundant proteins or transient interactions only observable by BioSITe. In addition, we observed that 807 proteins were uniquely identified by the conventional method, and we hypothesize that many of these represent nonspecifically bound proteins (46% (378/807)) are represented by ≤2 PSMs) (Supplemental Table 2).

Mapping of Biotinylation Sites in Proteins

To study the pattern of protein biotinylation and to determine if any interesting features could be revealed by the degree of biotinylation and the location of biotinylation sites, we generated a curated list of core interactors of BCR-ABL. We then grouped these proteins according to the proportion of lysines that were detected to be biotinylated (Figure 2A). Among the proteins with the greatest degree of biotinylation, we found that established direct BCR-ABL interactors such as GRB2,[19,20] GRB2-associated binding protein (GAB2,)[21,22] and SHC1[23] contained >23% of lysines that were biotinylated. The bait protein, BCR-ABL, exhibited even higher biotinylation as 75 of 95 lysine residues were biotin-modified, whereas other proteins known to interact with BCR-ABL via an adaptor subunit, for example, PIK3CA/B, exhibited a lower extent of biotinylation (3%). Thus we hypothesize that a high degree of biotinylation could reflect proximity and high residence time between BCR-ABL and its interactors or substrates. Mapping of biotinylated lysines onto the structures of corresponding proteins revealed that the majority of biotinylated lysines were in regions with low secondary structures (i.e., located within N- or C-termini or interdomain linkers outside of folded domains). As expected in this case, a majority of biotinylated sites were positioned near known functional domains involved in protein–protein interactions with BCR-ABL such as src homology 2 (SH2) domain, phosphotyrosine binding (PTB) domain, and src homology 3 (SH3) domain (Figure 2B). For instance, GRB2, which binds BCR-ABL via its SH2 domain, is highly biotinylated in both of its SH3 domains that are located adjacent to the SH2 domain, while the adapter molecule, CRK proto-oncogene adaptor protein (CRK), is additionally biotinylated on the SH2 domain itself (Figure 2B). Two biotinylated lysine residues (K689, K696) of signal transducer and activator of transcription 5A (STAT5A), a prominent substrate of BCR-ABL, are also found adjacent to the SH2 domain (Figure 2B), which mediates dimerization subsequent to phosphorylation of STAT5A by BCR-ABL.[24,25] Taken together, our observations suggest that the site-level information generated by BioSITe could be potentially useful for gaining spatial information on the architecture of protein complexes.

Application of BioSITe for Quantitative Proteomics

Given the utility of identifying site-specific biotinylation sites, we wished to adopt a quantitative approach that would allow us to investigate relative abundance of biotinylated peptides under different biological conditions. Traditionally, accurate relative quantitation approaches have relied upon SILAC, which is based on metabolic incorporation of stable

isotope containing amino acids into proteins in living cells.[14] Here, instead of using heavy amino acids, we decided to use heavy biotin in the BioID experiment that would serve the same purpose. Again, we used our previously established system for interrogating the differential interactomes of two variants of BCR-ABL oncogene.[17] The p190 and p210 versions of BCR-ABL differ from each other only by the presence or absence of DH and PH domains but have distinct phenotypes in humans. We previously tagged both variants with BirA* (hereafter referred to as BirA*-p190 and BirA*-p210)[17] and used SILAC-based quantitation to characterize their differential interactomes.[17,18] We repeated the experiment by culturing the cells overnight in "heavy ($^2H_4$)" or "light (H_4)" biotin (Figure 3A). LC–MS/MS analysis of mixed cell lysates from BirA*-p190 and BirA*-210 cells using BioSITe revealed the expected mass differences at the MS as well as MS/MS levels on biotinylated peptides (Supplementary Figure 2A). As expected, the biotinylation sites from the DH and PH domain of BCR-ABL p210 were highly enriched in BirA*-p210 harboring cells (Figure 3B, Supplemental Table 6). To compare this BioSITe experiment with labeled biotin with our previous SILAC-based BioID study, we averaged all biotinylation site ratios in the BioSITe data for specific proteins. We found many differential interactors determined by our previous studies to be relevant for BCR-ABL p190- and p210-specific signaling pathways; these included the p210-enriched interactors of SHIP1, SHIP2, CBL, and UBASH3B (Supplementary Figure 2B) and the p190-enriched interactors of Wiskott–Aldrich Syndrome protein (WAS) and WAS/WASL-interacting protein family member 1 (Wipf1).[17,18] Our quantitative BioSITe analysis also led to the identification of potentially novel interactors that preferentially interacted either with p190 (e.g., SRSF protein kinase 2 (SRPK2) and BRCA1 associated protein (BRAP)) or with p210 (e.g., C-terminal src kinase (CSK) and PDGFA-associated protein 1 (PDAP1) (Figure 3B)). Interestingly, we did not previously identify CSK as a p210-specific interactor, although it was found to be hyperphosphorylated in cells expressing p210. Thus the preferential interaction between CSK and p210 BCR-ABL, which was observed by quantitative BioSITe, further implicates its role in p210-specific signaling pathways.[17,18]

Our quantitative BioSITe approach has potential advantages over traditional SILAC for proximity-based labeling methods. While SILAC requires multiple passages to ensure complete metabolic labeling, which has its limitations with difficult to culture and slowly dividing cells, the addition of heavy-labeled biotin can be completed with an overnight incubation in the setting of BioID. Using heavy biotin-phenol in the APEX system coupled to BioSITe is another potential of the method as APEX requires a very short labeling time (<1 min).[2,26]

Applying BioSITe to the APEX System

To demonstrate the broad applicability of BioSITe, we applied our method to APEX, another proximity-dependent biotinylation method.[3] We chose an established system of mapping the mitochondrial subcellular proteome in which we expressed APEX2 fused to signal sequences, targeting it to either the intermembrane space in the mitochondria (IMS-APEX2) or the cytoplasm (NES-APEX2) in HEK293T cells (Figure 4A).[3] We carried out BioSITe experiments in triplicate for both IMS-APEX2 and NES-APEX2, and mass spectrometry analysis confirmed the enrichment of biotin-phenol-modified peptides (Figure 4B). Peptides

containing biotin-phenol-modified tyrosines were readily detected by LC–MS/MS analysis in both experiments. Biotin-phenol modified tyrosines were confirmed by MS/MS spectrum with three signature ion peaks as well as a series of fragmentation ions harboring biotin. The signature ions, m/z 227.08, m/z 480.19 and m/z 497.22, correspond to fragmented biotin, biotin-phenol modified tyrosine immonium ions with loss of NH_3, and biotin-phenol modified tyrosine immonium ions, respectively (Figure 4B).[2,26]

BioSITe identified a total of 1454 biotin-phenol modified peptides derived from 656 proteins from IMS-APEX2 labeled lysates (Supplemental Table 7) biotin-phenol modified peptides derived from 786 proteins from NES-APEX2 labeled lysates (Supplemental Table 8). After filtering out modified peptides common to both NES-APEX2 and IMS-APEX2 samples, we identified 384 biotin-phenol modified peptides representing 244 proteins uniquely identified by IMS-APEX2 (Figure 4C). When we compared our data with a previously published study using the same system in which 135 IMS proteins were identified,[3] we observed that 62 proteins were identified by both methods. Of the 182 proteins that were only identified by the BioSITe method, 74 proteins are already annotated in MitoCarta, a mitochondrial protein database (Figure 4C),[27] suggesting that these proteins were likely bona fide proteins that were missed by the previous study. Furthermore, BioSITe identified 345 biotinylated sites, while the previous study found 80 sites, of which 45 biotinylated peptides were identified by both methods and the large majority were unique to the BioSITe method (Figure 4D).

Recently, a method based on desthiobiotin-phenol, a synthetic biotin-phenol derivative, has been applied to the APEX system, which takes advantage of weaker affinity of desthiobiotin-phenol to streptavidin.[26] In this study, desthio-biotin-phenol-containing peptides were enriched and used to define the IMS proteome. When we compared proteins and peptides identified by all three methods (in-gel digestion based methods by Hung et al., desthiobiotin-phenol-based study by Lee et al., and BioSITe), BioSITe yielded more labeled peptides and proteins (Supplementary Figure 3A,B). We have noticed that there are many proteins and peptides that are uniquely identified by each method, suggesting that perhaps mitochondrial proteome is still insufficiently cataloged, there are issues with sampling, there are inherent biases of each method, or a combination of these factors (Supplementary Figure 3A,B). As an additional caveat, we cannot rule out the possibility that each data set includes false-positives owing to porous mitochondrial outer membrane, which could itself lead to labeling of cytosolic proteins.[3]

Biotinylation Site-Dependent Topology Prediction

A promising attribute of site-specific biotinylation data in the context of subcellular proteome mapping is the ability to infer topology of proteins with transmembrane domains. For example, using a construct that localizes APEX2 to the mitochondrial matrix (Matrix-APEX2) in conjunction with the IMS-APEX2 construct, one could map the parts of transmembrane proteins that have domains in the matrix and the IMS. In some cases, BioSITe was able to provide biotinylation sites that definitively allowed us to predict the topology of 29 mitochondrial proteins. For instance, for translocase of inner mitochondrial membrane 50 (TIMM50) anchored in inner mitochondria membrane, BioSITe provided biotinylation sites of Y122, Y255, Y259, and Y301 at its C-terminus (Figure 4E, top). In

agreement with our data, the C-terminus of TIMM50 has been previously shown to be exposed to the IMS.[28,29] Biotinylation of Y301 was also described by a different study by Lee et al. using desthiobiotin-phenol as a molecular probe.[26] In the case of NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8 (NDUFB8), BioSITe identified biotinylation of Y161 at the C-terminus, thus confirming the previously resolved structure.[30] Biotinylation of the IMS side in NDUFB8 was not identified by desthiobiotin-phenol-based analysis.[26] Finally, in the case of a known mitochondrial protein, coiled-coil domain containing protein 51 (CCDC51), whose topology has not yet been well characterized, BioSITe led to the identification of a novel site at Y275, which is positioned between the two putative trans-membrane domains, thereby establishing its topology in the inner membrane of the mitochondria.[31] As labeled by IMS-APEX2, we hypothesized that the region containing Y275 is located in the IMS. We also noticed that Y143 located at the N-terminus of CCDC51 has been annotated by Matrix-APEX,[26] suggesting the CCDC51 is embedded in the mitochondrial inner membrane and the N terminus of CCDC51 is located inside the mitochondria matrix. Taken together, the observed site-level information indicates an IMS linker bridging the two transmembrane domains and the C terminus is also projecting into the matrix (Figure 4E, left).

Applying BioSITe to Biotin-Based Click Chemistry

To test the potentially broad applicability of BioSITe, we applied our method to a biotin-based click chemistry strategy to identify the post-translational modification of serine and threonine residues by *N*-acetylglucosamine (O-GlcNAc).[32] O-GlcNAc is associated with a wide variety of cellular processes including cell cycle, stress response, transcription, and nutrient sensing.[33,34] To detect O-GlcNAc-modified residues, whole cell lysates from HEK293T cells were digested by the protease Lys-C and the resulting peptides were subjected to PNGase F and mutant β-1,4-galactosyltransferase (Gal-T1 (Y289L)) treatment, which lead to specific labeling of O-GlcNAc-modified peptides with an additional sugar containing an azide (GalNAz) moiety, which is compatible with click chemistry (Figure 4F). These modified peptides were subsequently biotinylated in a second reaction by click chemistry using DIBO biotin alkyne where the azide group of UDP-GalNAz recruits the alkyne group attached to biotin.[35,36] Peptides containing biotin were then captured using BioSITe for LC–MS/MS analysis.

Our LC–MS/MS analysis identified O-GlcNAc sites on many known O-GlcNAc-modified proteins including host cell factor 1 (HCFC1), nuclear pore complex protein 153 (NUP153), nuclear pore complex protein 214 (NUP214), and nuclear pore complex protein 98 (NUP98)[37,38] (Supplementary Table 9). A total of 10 known O-GlcNAc sites were identified on HCF1, while 5 known sites were identified on NUP153. NUP153 is a nuclear pore protein with zinc finger domains that are critical for the recruitment of COP1 (coatomer complex protein 1) to the nuclear envelope. NUP153 is involved in the breakdown of the nuclear envelope at the prophase of cell cycle.[39] It has been shown that the level of O-GlcNAcylation at S529 increases dramatically upon phosphorylation of NUP153 at S534 by CDK1 during cell cycle.[37] Figure 4G shows an MS/MS spectrum of an O-GlcNAc-modified peptide from NUP153. Click chemistry strategies are not limited to identification sites of glycosylated proteins but have many applications; in particular, recent studies of small-

molecule probes coupled to click chemistry are enabling high-throughput drug discovery.[40,41] We anticipate that BioSITe could benefit these experiments as many existing reagents rely on biotin for enrichment by precluding site-specific identification for the reasons described above.

CONCLUSIONS

We have shown that BioSITe permits enrichment and detection of biotinylated peptides including the localization of biotinylation sites within peptides. Using an isotopically labeled version of biotin, BioSITe allows for a quantitative analysis that is ideal for characterizing molecular differences across different biological conditions. Although we have demonstrated the use of a four Dalton heavier biotin, which is commercially available, other biotin isotopologues can be generated that will allow for higher levels of multiplexing. The ability to obtain site level information in terms of biotinylation can provide spatial information regarding the architecture of protein complexes and topology of subcellular proteomes. We have also shown the broad applicability of BioSITe by applying this strategy to a biotin-based click chemistry strategy for identifying O-GlcNAc-modified sites. Overall, BioSITe is an attractive option as a simpler strategy with a high yield than previously used methods, which relied on low yields of biotinylated peptides and, in some cases, a higher number of LC–MS/MS runs in in-gel digests or customized molecular probes (e.g., desthiobiotin-phenol), which are out of reach for most laboratories to generate. Additionally, the site-level information provided in click chemistry approaches will circumvent the need for complicated cleavable tags that are frequently employed. Taken together, we foresee the general applicability of BioSITe in most applications that involve biotinylation as a strategy to tag proteins or post-translational modifications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by NCI's Clinical Proteomic Tumor Analysis Consortium Initiative (U24CA210985) and a shared instrumentation grant (S10OD021844). J.A.C. was supported by NIGMS Training Grant 5T32GM07814. We thank the Center for Proteomics Discovery at Johns Hopkins and all the members of the Pandey laboratory.

References

1. Kim DI, Roux KJ. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. Trends Cell Biol. 2016; 26(11):804–817. [PubMed: 27667171]
2. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, Ting AY. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. Science. 2013; 339(6125):1328–1331. [PubMed: 23371551]
3. Hung V, Zou P, Rhee H-W, Udeshi ND, Cracan V, Svinkina T, Carr SA, Mootha VK, Ting AY. Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. Mol Cell. 2014; 55(2):332–341. [PubMed: 25002142]
4. Chakravartty V, Cronan JE. The Wing of a Winged Helix-Turn-Helix Transcription Factor Organizes the Active Site of BirA, a Bifunctional Repressor/Ligase. J Biol Chem. 2013; 288(50):36029–36039. [PubMed: 24189073]

5. Chapman-Smith A, Cronan JE. The Enzymatic Biotinylation of Proteins: A Post-Translational Modification of Exceptional Specificity. Trends Biochem Sci. 1999; 24(9):359–363. [PubMed: 10470036]
6. Roux KJ, Kim DI, Raida M, Burke B. A Promiscuous Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in Mammalian Cells. J Cell Biol. 2012; 196(6):801–810. [PubMed: 22412018]
7. De La Fuente EK, Dawson CA, Nelin LD, Bongard RD, McAuliffe TL, Merker MP. Biotinylation of Membrane Proteins Accessible via the Pulmonary Circulation in Normal and Hyperoxic Rats. Am J Physiol. 1997; 272(3 Pt 1):L461–470. [PubMed: 9124603]
8. Rybak J-N, Scheurer SB, Neri D, Elia G. Purification of Biotinylated Proteins on Streptavidin Resin: A Protocol for Quantitative Elution. Proteomics. 2004; 4(8):2296–2299. [PubMed: 15274123]
9. Tong X, Smith LM. Solid-Phase Method for the Purification of DNA Sequencing Reactions. Anal Chem. 1992; 64(22):2672–2677. [PubMed: 1294003]
10. Morag E, Bayer EA, Wilchek M. Reversibility of Biotin-Binding by Selective Modification of Tyrosine in Avidin. Biochem J. 1996; 316(1):193–199. [PubMed: 8645205]
11. Laitinen OH, Nordlund HR, Hytönen VP, Uotila STH, Marttila AT, Savolainen J, Airenne KJ, Livnah O, Bayer EA, Wilchek M, et al. Rational Design of an Active Avidin Monomer. J Biol Chem. 2003; 278(6):4010–4014. [PubMed: 12458212]
12. Taskinen B, Zauner D, Lehtonen SI, Koskinen M, Thomson C, Kähkönen N, Kukkurainen S, Määttä JAE, Ihalainen TO, Kulomaa MS, et al. Switchavidin: Reversible Biotin-Avidin-Biotin Bridges with High Affinity and Specificity. Bioconjugate Chem. 2014; 25(12):2233–2243.
13. Schiapparelli LM, McClatchy DB, Liu H-H, Sharma P, Yates JR, Cline HT. Direct Detection of Biotinylated Proteins by Mass Spectrometry. J Proteome Res. 2014; 13(9):3966–3978. [PubMed: 25117199]
14. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. Mol Cell Proteomics. 2002; 1(5):376–386. [PubMed: 12118079]
15. Udeshi ND, Pedram K, Svinkina T, Fereshetian S, Myers SA, Aygun O, Krug K, Clauser K, Ryan D, Ast T, et al. Antibodies to Biotin Enable Large-Scale Detection of Biotinylation Sites on Proteins. Nat Methods. 2017; 14(12):1167–1170. [PubMed: 29039416]
16. Chaiet L, Wolf FJ. The Properties of Streptavidin, a Biotin-Binding Protein Produced by Streptomycetes. Arch Biochem Biophys. 1964; 106(Supplement C):1–5. [PubMed: 14217155]
17. Cutler JA, Tahir R, Sreenivasamurthy SK, Mitchell C, Renuse S, Nirujogi RS, Patil AH, Heydarian M, Wong X, Wu X, et al. Differential Signaling through p190 and p210 BCR-ABL Fusion Proteins Revealed by Interactome and Phosphoproteome Analysis. Leukemia. 2017; 31:1513–1524. [PubMed: 28210003]
18. Reckel S, Hamelin R, Georgeon S, Armand F, Jolliet Q, Chiappe D, Moniatte M, Hantschel O. Differential Signaling Networks of Bcr–Abl p210 and p190 Kinases in Leukemia Cells Defined by Functional Proteomics. Leukemia. 2017; 31:1502–1512. [PubMed: 28111465]
19. Weisswange I, Newsome TP, Schleich S, Way M. The Rate of N-WASP Exchange Limits the Extent of ARP2/3-Complex-Dependent Actin-Based Motility. Nature. 2009; 458(7234):87–91. [PubMed: 19262673]
20. Tomas A, Futter CE, Eden ER. EGF Receptor Trafficking: Consequences for Signaling and Cancer. Trends Cell Biol. 2014; 24(1):26–34. [PubMed: 24295852]
21. Dunn GP, Cheung HW, Agarwalla PK, Thomas S, Zektser Y, Karst AM, Boehm JS, Weir BA, Berlin AM, Zou L, et al. Vivo Multiplexed Interrogation of Amplified Genes Identifies GAB2 as an Ovarian Cancer Oncogene. Proc Natl Acad Sci U S A. 2014; 111(3):1102–1107. [PubMed: 24385586]
22. Zheng Y, Zhang C, Croucher DR, Soliman MA, St-Denis N, Pasculescu A, Taylor L, Tate SA, Hardy WR, Colwill K, et al. Temporal Regulation of EGF Signalling Networks by the Scaffold Protein Shc1. Nature. 2013; 499(7457):166–171. [PubMed: 23846654]
23. Brehme M, Hantschel O, Colinge J, Kaupe I, Planyavsky M, Köcher T, Mechtler K, Bennett KL, Superti-Furga G. Charting the Molecular Network of the Drug Target Bcr-Abl. Proc Natl Acad Sci U S A. 2009; 106(18):7414–7419. [PubMed: 19380743]

24. Schaller-Schönitz M, Barzan D, Williamson AJK, Griffiths JR, Dallmann I, Battmer K, Ganser A, Whetton AD, Scherr M, Eder M. BCR-ABL Affects STAT5A and STAT5B Differentially. PLoS One. 2014; 9(5):e97243. [PubMed: 24836440]
25. Hantschel O, Warsch W, Eckelhart E, Kaupe I, Grebien F, Wagner K-U, Superti-Furga G, Sexl V. BCR-ABL Uncouples Canonical JAK2-STAT5 Signaling in Chronic Myeloid Leukemia. Nat Chem Biol. 2012; 8(3):285–293. [PubMed: 22286129]
26. Lee S-Y, Kang M-G, Shin S, Kwak C, Kwon T, Seo JK, Kim J-S, Rhee H-W. Architecture Mapping of the Inner Mitochondrial Membrane Proteome by Chemical Tools in Live Cells. J Am Chem Soc. 2017; 139(10):3651–3662. [PubMed: 28156110]
27. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: An Updated Inventory of Mammalian Mitochondrial Proteins. Nucleic Acids Res. 2016; 44(D1):D1251–1257. [PubMed: 26450961]
28. Pawlak KJ, Prasad M, Thomas JL, Whittal RM, Bose HS. Inner Mitochondrial Translocase Tim50 Interacts with 3β-Hydroxysteroid Dehydrogenase Type 2 to Regulate Adrenal and Gonadal Steroidogenesis. J Biol Chem. 2011; 286(45):39130–39140. [PubMed: 21930695]
29. de la Cruz L, Bajaj R, Becker S, Zweckstetter M. The Intermembrane Space Domain of Tim23 Is Intrinsically Disordered with a Distinct Binding Region for Presequences. Protein Sci Publ Protein Soc. 2010; 19(11):2045–2054.
30. Wu M, Gu J, Guo R, Huang Y, Yang M. Structure of Mammalian Respiratory Supercomplex I1III2IV1. Cell. 2016; 167(6):1598–1609. [PubMed: 27912063]
31. Floyd BJ, Wilkerson EM, Veling MT, Minogue CE, Xia C, Beebe ET, Wrobel RL, Cho H, Kremer LS, Alston CL, et al. Mitochondrial Protein Interaction Mapping Identifies Regulators of Respiratory Chain Function. Mol Cell. 2016; 63(4):621–632. [PubMed: 27499296]
32. Hart GW, Housley MP, Slawson C. Cycling of O-Linked β-N-Acetylglucosamine on Nucleocytoplasmic Proteins. Nature. 2007; 446(7139):1017–1022. [PubMed: 17460662]
33. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross Talk Between O-GlcNAcylation and Phosphorylation: Roles in Signaling, Transcription, and Chronic Disease. Annu Rev Biochem. 2011; 80:825–858. [PubMed: 21391816]
34. Love DC, Hanover JA. The Hexosamine Signaling Pathway: Deciphering the “O-GlcNAc Code. Sci Signaling. 2005; 2005(312):re13–re13.
35. Ning X, Guo J, Wolfert MA, Boons G-J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. Angew Chem, Int Ed. 2008; 47(12):2253–2255.
36. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. Angew Chem, Int Ed. 2002; 41(14):2596–2599.
37. Wang Z, Udeshi ND, Slawson C, Compton PD, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, Hart GW. Extensive Crosstalk Between O-GlcNAcylation and Phosphorylation Regulates Cytokinesis. Sci Signaling. 2010; 3(104):ra2–ra2.
38. Myers SA, Daou S, Affar EB, Burlingame A. Electron Transfer Dissociation (ETD): The Mass Spectrometric Breakthrough Essential for O-GlcNAc Protein Site Assignments—a Study of the O-GlcNAcylated Protein Host Cell Factor C1. Proteomics. 2013; 13(6):982–991. [PubMed: 23335398]
39. Mackay DR, Elgort SW, Ullman KS. The Nucleoporin Nup153 Has Separable Roles in Both Early Mitotic Progression and the Resolution of Mitosis. Mol Biol Cell. 2009; 20(6):1652–1660. [PubMed: 19158386]
40. Niessen S, Dix MM, Barbas S, Potter ZE, Lu S, Brodsky O, Planken S, Behenna D, Almaden C, Gajiwala KS, et al. Proteome-Wide Map of Targets of T790M-EGFR-Directed Covalent Inhibitors. Cell Chem Biol. 2017; 24(11):1388–1400. [PubMed: 28965727]
41. Bar-Peled L, Kemper EK, Suciu RM, Vinogradova EV, Backus KM, Horning BD, Paul TA, Ichu T-A, Svensson RU, Olucha J, et al. Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. Cell. 2017; 171(3):696–709. [PubMed: 28965760]

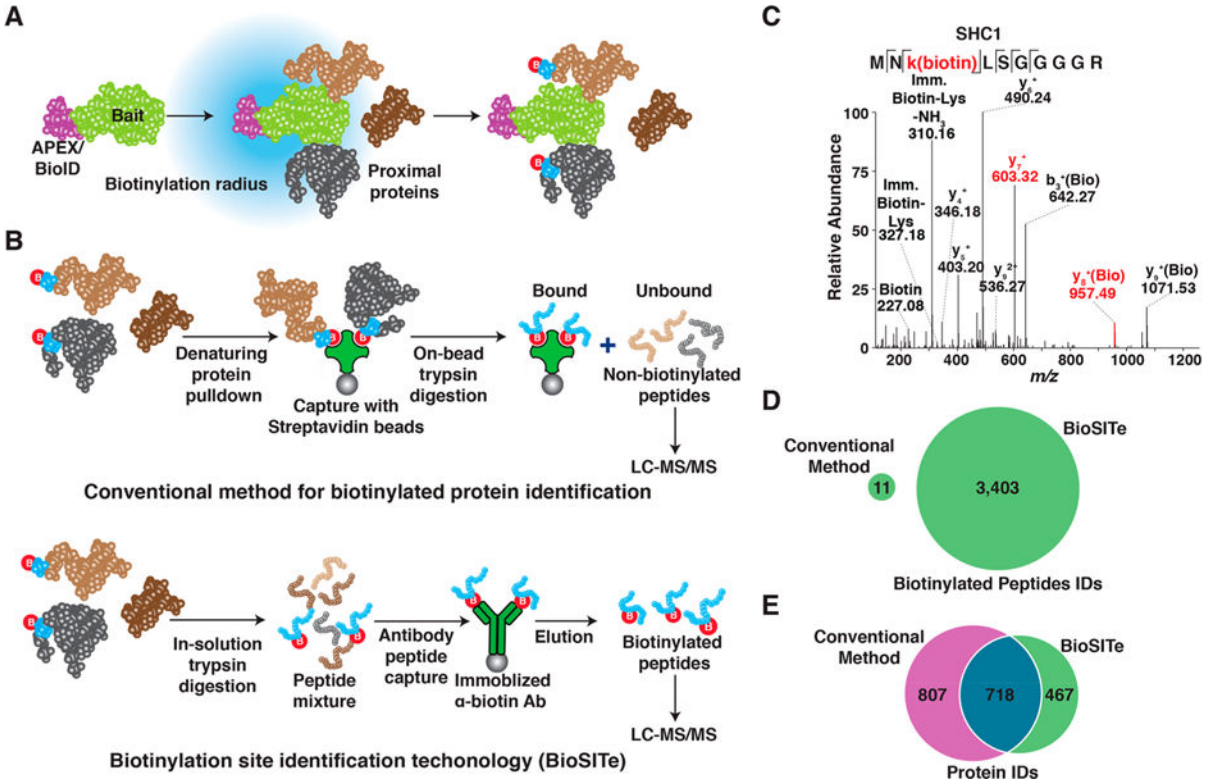


Figure 1.

Overview of BioSITe. (A) Schematic of biotinylation labeling techniques commonly used to identify protein–protein interactions in vivo. A bait protein (green) fused to an engineered biotinylation enzyme (APEX2 or BioID, purple) biotinylates proteins in the proximity of the enzymes. Biotinylation, red circles with “B”; interacting proteins, brown/black; labeling radius, blue. (B) In conventional methods for candidate protein identification, biotinylated proteins are generally captured by streptavidin (green) conjugated to beads under denaturing conditions. Proteins bound to the beads are subsequently digested, generating nonbiotinylated peptides (brown and gray), which readily elute from the beads and can be identified by LC–MS/MS, and biotinylated peptides (cyan) that remain tightly bound to the beads. In BioSITe, proteins are digested prior to enrichment and biotinylated peptides are captured using by anti-biotin antibodies coupled to beads. (C) MS/MS spectra of SHC-transforming protein 1 (SHC1) biotinylated peptide detected by LC–MS/MS. Fragment ions adjacent to the biotin modification that confirm the site of biotinylation are indicated in red. (D) Lack of overlap of biotinylated peptides identified from BirA*-BCR-ABL detected by the conventional method or BioSITe. (E) Overlap of biotinylated proteins identified by BioSITe (green) and by the conventional method of on-bead digestion (purple).

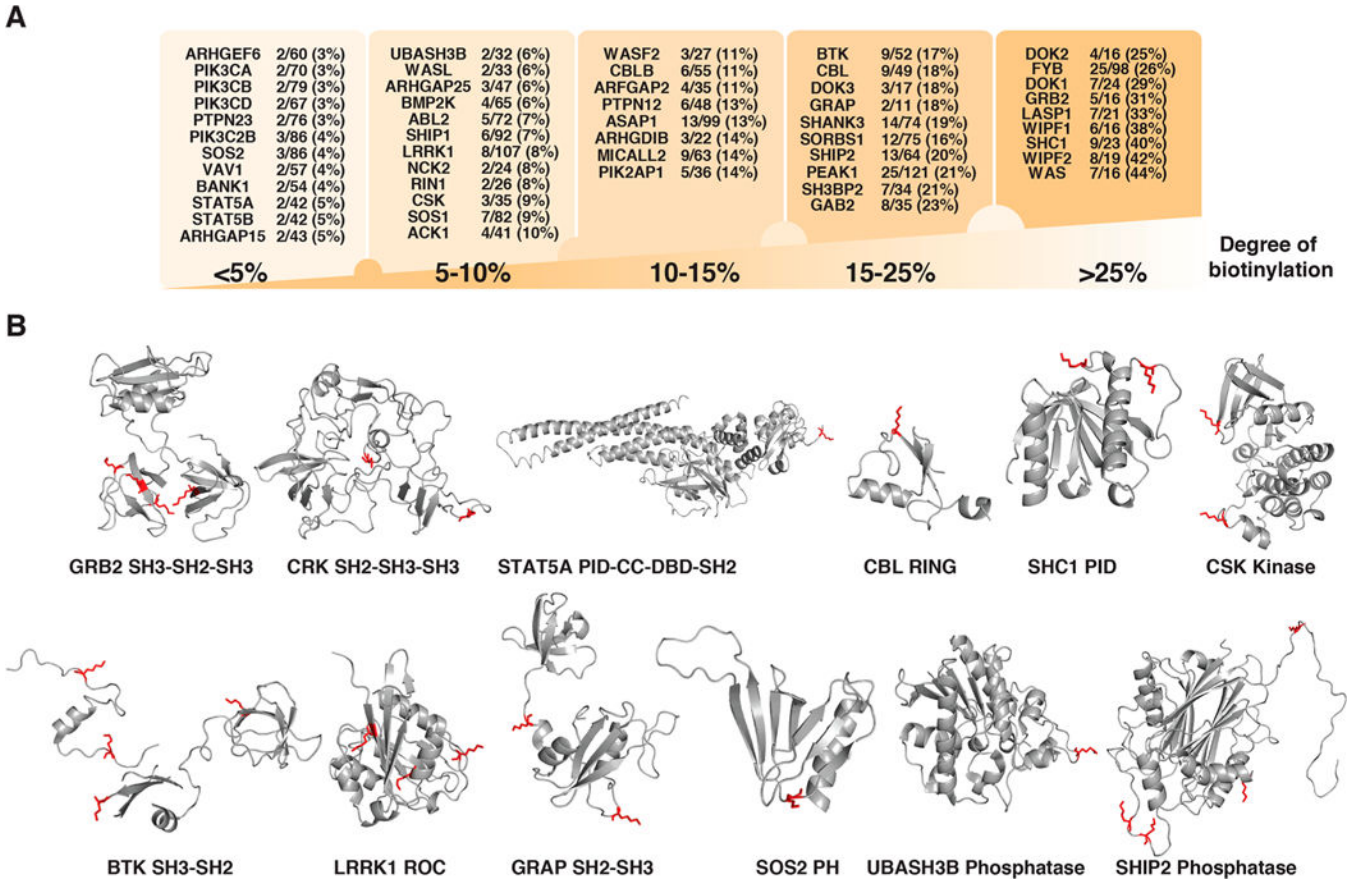


Figure 2.

Mapping of biotinylation sites. (A) Biotinylated proteins identified by BioSITe were grouped by the degree of biotinylation. (B) 3D models of representative proteins identified in the study. GRB2 and CRK are homology models based on their human homologues (PDB ID 1GRI and 2EYZ, respectively), while the structure of STAT5A was taken from PDB ID 1Y1U. The other domain structures from each protein were modeled using their human homologues from PDB. Lysine residues that are biotinylated upon interaction with BirA*-BCR-ABL are highlighted as red sticks and labeled with their position. The biotinylated lysines are colored in red and the functional domains are indicated (CC, coiled coil; ROC, Ras of complex proteins; PH, pleckstrin homology; SH3, Src-homology 3; SH2, Src-homology 2; PID, phosphotyrosine interaction domain; DBD, DNA-binding domain).

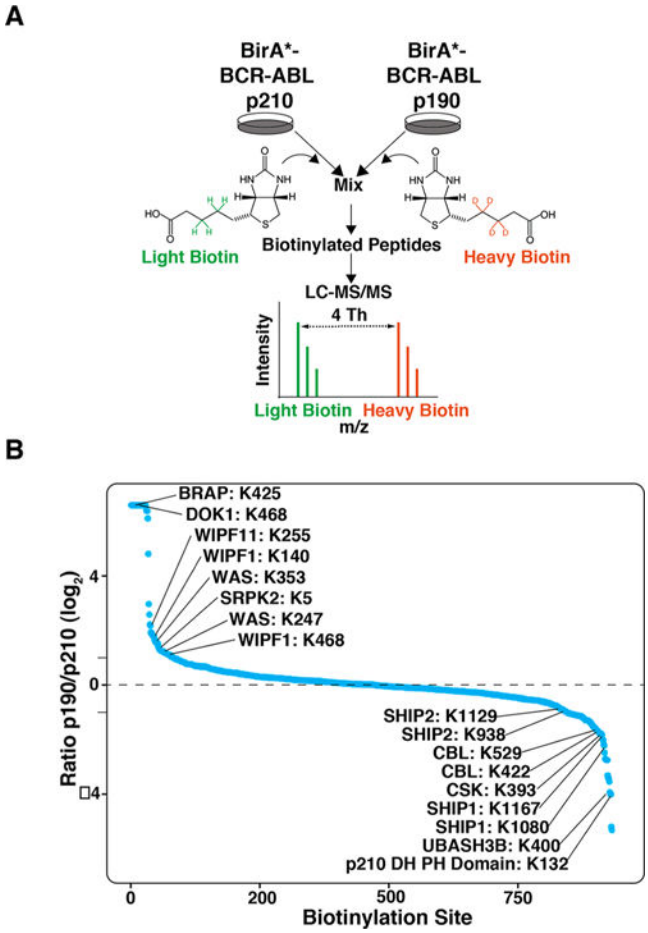


Figure 3.

Overview of quantitative BioSITe. (A) Experimental workflow for differential interactome analysis of BirA*-p210 and BirA*-p190. Ba/F3 cells expressing BirA*-p210 and BirA*-p190 were incubated overnight with media containing light (green) or heavy biotin containing four deuterium atoms (red), respectively. Equal amounts of cell lysates from each condition were mixed and digested into peptides. Biotinylated peptides were enriched using BioSITe and analyzed by LC–MS/MS. (B) Plot of relative abundance of the biotinylated sites between BirA*-p210 and BirA*-p190. Identified biotinylated sites and the corresponding proteins are plotted according to their $\log_2$ intensity ratios (BirA*-p190/BirA*-p210).

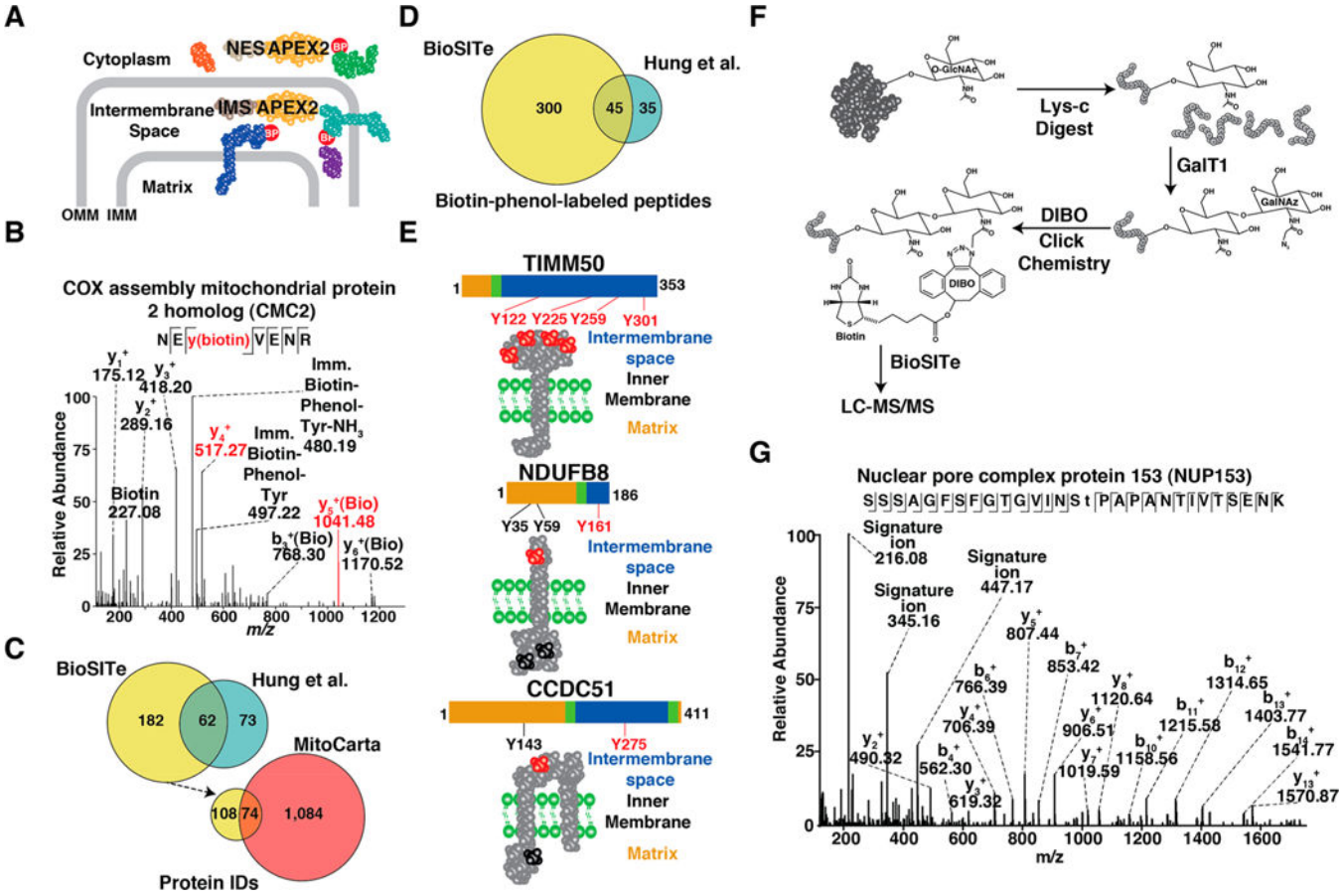


Figure 4.

Application of BioSITe to the APEX system and biotin-based click chemistry (A) APEX2 constructs targeted to either the mitochondrial intermembrane space (IMS-APEX2) or cytoplasm (NES-APEX2) were expressed in HEK293T cells. APEX2 leads to labeling of proximal proteins in the intermembrane space or the cytoplasm with biotin-phenol (BP, red) (OMM, outer mitochondrial membrane; IMM, inner mitochondrial membrane). (B) MS/MS spectrum of a biotin-phenol-labeled peptide corresponding to COX assembly mitochondrial protein 2 homologue (CMC2). Peaks labeled in red indicate fragment ions series with or without mass shift by biotin modification. All peaks containing biotin are labeled with "Bio". (C) Overlap of proteins identified by BioSITe (yellow) with that from a previous study identifying IMS proteins using a conventional in-gel digestion method (Hung et al., blue). The proteins unique to the BioSITe IMS-APEX2 experiment were compared with published mitochondrial protein database MitoCarta 2.0 (red). (D) Overlap of biotin-phenol-labeled peptides enriched by BioSITe (yellow) with that from Hung et al. (blue). (E) Predicted topology of mitochondrial transmembrane proteins of translocase of inner mitochondrial membrane 50 (TIMM50) (top), NADH: ubiquinone oxidoreductase subunit A8 (NDUFB8) (middle), and coiled-coil domain containing 51 (CCDC51) (bottom). Matrix-facing (orange), transmembrane (green), and intermembrane space domains are indicated. Cartoon depictions of these proteins embedded in the inner membrane (green) are shown in gray with biotinylation sites detected by BioSITe are colored in red and black, detected by

Lee et al. using the Matrix-APEX2 construct. (F) Schematic overview of enrichment method for O-GlcNAcylated peptides. After LysC digestion, an azide-modified monosaccharide (GalNAz) is added to the O-GlcNAc motif using galactosyltransferase1 GalT1. Biotinylation of O-GlacNAc motif subsequently mediated by click chemistry reaction between the azide group and of GalNAz and the alkyne group of biotin DIBO alkene. Biotin is used to enrich peptides with O-GlcNAc motif using BioSITe. (G) MS/MS spectra of a O-GlcNAc modified peptide corresponding to nuclear pore complex protein NUP153 (NUP153). Ion peaks at m/z 216.08, 345.15, and 447.16, labeled signature ions, correspond to fragments from click chemistry reagents.