A module for analyzing interactomes via APEX-MS integrated into PatternLab for proteomics

Marlon D. M. Santos[1,2*], Amanda C. Camillo-Andrade[1,2], Azalia Rodriguez[1,3], Rosario Durán[1]

[1]Analytical Biochemistry and Proteomics Unit, Instituto de Investigaciones Biológicas Clemente Estable, Institut Pasteur de Montevideo, Mataojo 2020, 11400 Montevideo, Uruguay
[2]Laboratory for Structural and Computational Proteomics, Carlos Chagas Institute, Fiocruz - Paraná, R. Professor Algacyr Munhoz Mader 3775, Curitiba, PR, Brazil
[3]Facultad de Química, Universidad de la República, Av. Gral. Flores 2124, 11800 Montevideo, Uruguay.

ABSTRACT: Proximity labeling techniques, such as APEX-MS, provide valuable insights into proximal interactome mapping; however, the verification of biotinylated peptides is not straight forward. With this as motivation, we present a new module integrated into PatternLab for proteomics to enable APEX-MS data interpretation by targeting diagnostic fragment ions associated with APEX modifications. We re-analyzed a previously published APEX-MS dataset and report a significant number of biotinylated peptides and, consequently, a confident set of proximal proteins. As the module is part of the widely adopted PatternLab for proteomics software suite, it offers users a comprehensive, easy, and integrated solution for data analysis. Given the broad utility of the APEX-MS technique in various biological contexts, we anticipate that our module will be a valuable asset to researchers, facilitating and enhancing interactome studies. PatternLab´s APEX and including a usage protocol, are available at http://patternlabforproteomics.org/apex .

KEYWORDS: APEX-MS; PatternLab for proteomics; Ascorbate Peroxidase; Interactomes; Protein-Protein Interactions.

INTRODUCTION

The dynamic and complex interplay of proteins within cellular systems underscores their fundamental roles in virtually every biological process. The study of protein-protein interactions (PPIs) is not just a matter of cataloging these relationships but also allows for discerning their functional implications, potentially unveiling regulatory pathways, identifying potential drug targets.

Proximity labeling techniques, typified by APEX-MS, are at the forefront of methods designed to explore proteomic environments within living cells[1–3]. The foundational principle of APEX-MS lies in the engineered ascorbate peroxidase (APEX) which, when fused to a specific protein, facilitates targeted and rapid promiscuous biotinylation of neighboring proteins. Such a mechanism offers a valuable temporal perspective on the proteome surrounding the protein of interest, revealing potential protein partners, whether they are stable or transient in nature[4].

Furthering this technique, APEX2 is an enhanced version of the original APEX, offering improved labeling efficiency and reduced background noise, which is pivotal for studying proteins within the crowded environment of the cell. APEX2 enhanced kinetics for biotinylation means it requires shorter labeling times, which is ideal for tracking fast-changing events within cells. The incorporation of APEX2 into proximity labeling experiments is contributing to refine our understanding of protein interactions and the dynamic proteomic architecture of various cellular compartments, making it a cutting-edge tool in cellular biology and proteomics[5].

Upon the introduction of biotin-labeled substrates, such as phenol-biotin, in the company of hydrogen peroxide, the APEX enzyme catalyzes the formation of transient radicals that preferentially modify tyrosine residues on adjacent proteins. Subsequently, these biotin-tagged proteins can be isolated and identified through tandem mass spectrometry (MS/MS). Innovations in the form of biotinylated peptide enrichment have shown great promise for the accurate identification of neighboring proteins[6].

Moreover, these methods have proven effective in pinpointing the modified tyrosine residues. This identification process is facilitated by the presence of diagnostic ions in the MS/MS spectra, which not only aids in the determination of the modification but also provides valuable spatial and topological information. Nonetheless, a persistent challenge has been the confirmation of biotin-phenol tagged peptides reported by proteomic identifications, that requires manually validation, a laborious task that often yields inconsistent interpretations and can limit reproducibility. This is where the new module, integrated with PatternLab for proteomics[7], comes into play, offering a solution that considerably streamlines this aspect of the analytical process.

EXPERIMENTAL

To this end, PatternLab for proteomics enhances the methodology by incorporating an automated system into the validation step. It integrates a series of automated processes, beginning with effective proteomic searching, filters the results at a stringent 1% False Discovery Rate (FDR), and now includes a module for subsequent validation of peptides—which indirectly validates the associated proteins—based on the presence of reporter ions. For the experiment at hand, such signature ions originated from biotin-phenol tyrosine immonium ion[8].

We demonstrate the efficacy of this module by re-analyzing the PRIDE dataset PXD007862[8]. This aforementioned study utilized APEX2-MS to map the mitochondrial intermembrane space (IMS) proteome, using APEX2 fused to the leader sequence of an IMS native protein (IMS-APEX2) and a cytoplasmic APEX (NES-APEX2). The raw data files were searched using PatternLab's default parameters[9]. The search modifications included carbamidomethylation of cysteine as fixed and oxidized methionine and biotin-phenol modification of tyrosine as variable. The sequence

database was the UniProtKB human proteome. The applicability of this tool is further confirmed by the analysis of a SILAC-APEX2 labelled dataset (MSV000079096)[4].

RESULTS AND CONCLUSION

The BioSITe study benchmarks APEX2-MS for subcellular proteome mapping, finding 1,649 and 1,236 biotin-phenol-modified peptides in IMS-APEX2 and NES-APEX2 labeled samples, respectively. This totals 1,783 peptides, with 689 unique to IMS-APEX2 and 362 to NES-APEX2, highlighting the distinct proteomes each APEX construct reveals.
Our re-analysis corroborates these results and extends the inventory of identified peptides and proteins. Utilizing the PatternLab APEX pipeline, we identified 1,224 and 1,651 biotin-phenol-modified peptides in IMS-APEX2 and NES-APEX2 samples respectively, all of them containing the diagnostic ions of and *m/z* 497.22 (phenol-biotin modified Tyr immonium ion) and *m/z* 480.19 (phenol-biotin modified Tyr immonium ion-NH_3) in MS/MS spectrum. This yielded 2,061 biotin-phenol-modified peptides, with 410 exclusives to IMS-APEX2 and 837 to NES-APEX2, as illustrated in Figure 1 (Supplementary File 1).

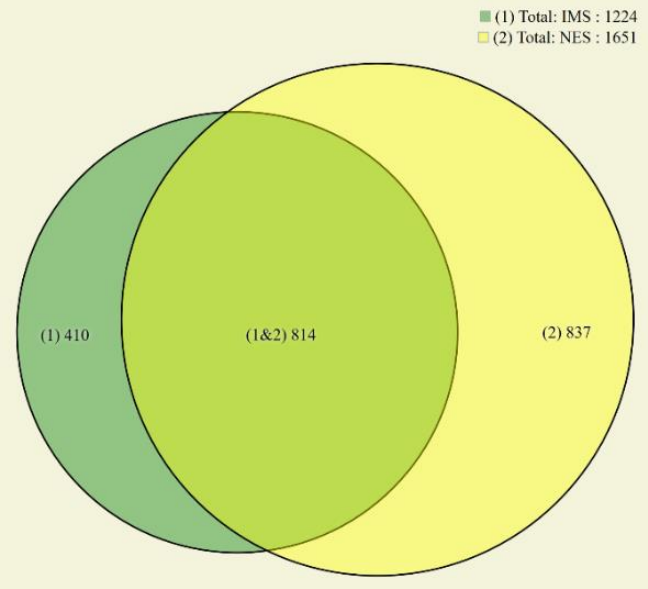

Figure 1. Venn diagram illustrating the overlap between IMS-APEX2 and NES-APEX2 in identifying proximity-labeled peptides. The PatternLab APEX identified 2,061 candidates across both samples. Specifically, 1,244 proteins were uniquely enriched by IMS-APEX2, suggesting potential mitochondrial inner membrane space subproteome interactors. NES-APEX2 labeled an additional 1,651 candidates. There were 814 peptides common to both conditions. This diagram underscores the enhanced capability of the optimized PatternLab workflow in distinguishing condition-specific proximity-labeled proteins in APEX-MS data.

The comparative analysis between the IMS-exclusive protein dataset and the MitoCarta v3[10] has indicated that 30% of the identified proteins were previously annotated as mitochondrial. This proportion closely aligns with that reported by the BioSITe study (Figure 2).

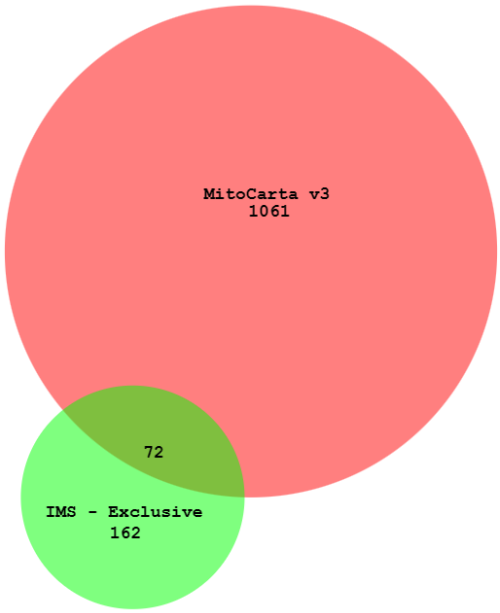

Figure 2. Venn diagram comparing mitochondrial proteins from MitoCarta v3 (1,133) with IMS-exclusive proteins (234). The overlap (72) validates the mitochondrial association of the IMS-identified proteins.

Enhanced detection capabilities in proteomics studies lead to more reliable data by robustly eliminating false positives, thus improving data integrity. Such refined methodological approaches have augmented our comprehension of the nuanced interactions within proteomic networks. Detailed accounts of the proteins identified through this improved methodological approach are provided in Supplementary File 2. A non-APEX dataset (PXD033782)[11] was used to evaluate the False Discovery Rate (FDR) and its dependence on search parameters, revealing that standard ion threshold and mass accuracy values are conservatively set in line with the customary 1% FDR (Supplementary File 3). The second dataset analyzed was downloaded from MassiVe (MSV000079096) and combines SILAC with APEX2. The authors reported 159 phenol-biotinylated peptides, of which 134 were confirmed by our pipeline, plus another 59 phenol-biotinylated peptides that mapped to 48 proteins; 40 of these were included in MitoCarta (Supplementary File 4 and 5).

The potential applications of this module are vast, ranging from elucidating protein interactions in disease mechanisms to identifying potential targets in drug discovery. Moreover, the application of the fragment ion screening strategy embodied by the PatternLab for proteomics APEX module extends its utility beyond the confines of proximity labeling. It paves the way for its adaptation in identifying other post-translational modifications (PTMs). The module's potential to be fine-tuned for the recognition of

signature ions from phosphorylation, acetylation, or glycosylation events opens new avenues for the validation of PTMs in proteomic studies. The fidelity conferred by this signature ion screening approach endows mass spectrometry-based proteomic research with an additional layer of validation, enhancing the confidence in PTM identification.

ASSOCIATED CONTENT

Supporting Information

Venn diagram of proximity-labeled peptides identified by IMS-APEX2 and NES-APEX2 (Supplementary_File_1.xlsx)

Venn diagram of proximity-labeled proteins identified by IMS-APEX2 and NES-APEX2 (Supplementary_File_2.xlsx)

Assessment of False Positives in Mass Spectrometry Data Analysis Using Non-Data APEX Dataset (Supplementary_File_3.docx)

Analysis of Phenol-Biotinylated SILAC-APEX2 Peptide (Supplementary_File_4.docx and Supplementary_File_5.xlsx)

AUTHOR INFORMATION

Corresponding Author

* Marlon D. M. Santos (marlondms@pasteur.edu.uy)

Author Contributions

MDMS was responsible for the development of the new module. MDMS and ACCA handled data analysis and Cover Art. MDMS and RD supervised the entirety of the project. MDMS, ACCA, AR and RD were actively involved in discussions during all phases of the research. MDMS, ACCA, and RD contributed to manuscript writing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank CNPq, Fiocruz INOVA Idéias and Inova Produtos, Fiocruz – PEP, Fundação Araucaria, and FOCEM-COF 03/11 and ANII, FCE_1_2019_1_155569 for financial support. AR was supported by ANII, CAP (UdelR) and PEDECIBA.

REFERENCES

(1) Hung, V.; Zou, P.; Rhee, H.-W.; Udeshi, N. D.; Cracan, V.; Svinkina, T.; Carr, S. A.; Mootha, V. K.; Ting, A. Y. `10.1007/978-1-0716-1975-9_23. *Molecular Cell* **2014**, *55* (2), 332–341. https://doi.org/10.1016/j.molcel.2014.06.003.

(2) Elmsaouri, S.; Markmiller, S.; Yeo, G. W. APEX Proximity Labeling of Stress Granule Proteins. In *The Integrated Stress Response*; Matějů, D., Chao, J. A., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2022; Vol. 2428, pp 381–399. https://doi.org/10.1007/978-1-0716-1975-9_23.

(3) Lee, S.-Y.; Kang, M.-G.; Shin, S.; Kwak, C.; Kwon, T.; Seo, J. K.; 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. https://doi.org/10.1021/jacs.6b10418.

(4) Rhee, H.-W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* **2013**, *339* (6125), 1328–1331. https://doi.org/10.1126/science.1230593.

(5) Martell, J. D.; Deerinck, T. J.; Sancak, Y.; Poulos, T. L.; Mootha, V. K.; Sosinsky, G. E.; Ellisman, M. H.; Ting, A. Y. Engineered Ascorbate Peroxidase as a Genetically Encoded Reporter for Electron Microscopy. *Nat Biotechnol* **2012**, *30* (11), 1143–1148. https://doi.org/10.1038/nbt.2375.

(6) Udeshi, N. D.; Pedram, K.; Svinkina, T.; Fereshetian, S.; Myers, S. A.; Aygun, O.; Krug, K.; Clauser, K.; Ryan, D.; Ast, T.; Mootha, V. K.; Ting, A. Y.; Carr, S. A. Antibodies to Biotin Enable Large-Scale Detection of Biotinylation Sites on Proteins. *Nat Methods* **2017**, *14* (12), 1167–1170. https://doi.org/10.1038/nmeth.4465.

(7) Santos, M. D. M.; Lima, D. B.; Fischer, J. S. G.; Clasen, M. A.; Kurt, L. U.; Camillo-Andrade, A. C.; Monteiro, L. C.; de Aquino, P. F.; Neves-Ferreira, A. G. C.; Valente, R. H.; Trugilho, M. R. O.; Brunoro, G. V. F.; Souza, T. A. C. B.; Santos, R. M.; Batista, M.; Gozzo, F. C.; Durán, R.; Yates, J. R.; Barbosa, V. C.; Carvalho, P. C. Simple, Efficient and Thorough Shotgun Proteomic Analysis with PatternLab V. *Nat Protoc* **2022**, *17* (7), 1553–1578. https://doi.org/10.1038/s41596-022-00690-x.

(8) Kim, D. I.; Cutler, J. A.; Na, C. H.; Reckel, S.; Renuse, S.; Madugundu, A. K.; Tahir, R.; Goldschmidt, H. L.; Reddy, K. L.; Huganir, R. L.; Wu, X.; Zachara, N. E.; Hantschel, O.; Pandey, A. BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *J. Proteome Res.* **2018**, *17* (2), 759–769. https://doi.org/10.1021/acs.jproteome.7b00775.

(9) Clasen, M. A.; Santos, M. D. M.; Kurt, L. U.; Fischer, J.; Camillo-Andrade, A. C.; Sales, L. A.; de Arruda Campos Brasil de Souza, T.; Lima, D. B.; Gozzo, F. C.; Valente, R. H.; Duran, R.; Barbosa, V. C.; Carvalho, P. C. PatternLab V Handles Multiplex Spectra in Shotgun Proteomic Searches and Increases Identification. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (4), 794–796. https://doi.org/10.1021/jasms.3c00063.

(10) Rath, S.; Sharma, R.; Gupta, R.; Ast, T.; Chan, C.; Durham, T. J.; Goodman, R. P.; Grabarek, Z.; Haas, M. E.; Hung, W. H. W.; Joshi, P. R.; Jourdain, A. A.; Kim, S. H.; Kotrys, A. V.; Lam, S. S.; McCoy, J. G.; Meisel, J. D.; Miranda, M.; Panda, A.; Patgiri, A.; Rogers, R.; Sadre, S.; Shah, H.; Skinner, O. S.; To, T.-L.; Walker, M. A.; Wang, H.; Ward, P. S.; Wengrod, J.; Yuan, C.-C.; Calvo, S. E.; Mootha, V. K. MitoCarta3.0: An Updated Mitochondrial Proteome Now with Sub-Organelle Localization and Pathway Annotations. *Nucleic Acids Research* **2021**, *49* (D1), D1541–D1547. https://doi.org/10.1093/nar/gkaa1011.

(11) Verissimo, D. C. A.; Camillo-Andrade, A. C.; Santos, M. D. M.; Sprengel, S. L.; Zanine, S. C.; Borba, L. A. B.; Carvalho, P. C.; Da G. Fischer, J. D. S. Proteomics Reveals Differentially Regulated Pathways When Comparing Grade 2 and 4 Astrocytomas. *PLoS ONE* **2023**, *18* (11), e0290087. https://doi.org/10.1371/journal.pone.0290087.

ABSTRACT GRAPHIC