

1 **Title:** In-gel protein digestion using acidic methanol produces a highly selective methylation of glutamic
2 acid residues.

3 **Authors:** Marta Lozano-Prieto^{a,1}, Emilio Camafeita^{b,c,1}, Inmaculada Jorge^{b,c} Andrea Laguillo-Gómez^b,
4 Rafael Barrero-Rodríguez^b, Cristina A. Devesa^b, Clara Pertusa^a, Enrique Calvo^{b,c}, Francisco Sánchez-
5 Madrid^{a,c,d,e,2}, Jesús Vázquez^{b,c,2} and Noa B. Martín-Cofreces^{a,c,f,2}

6
7 **Affiliations:**

8 ^aImmunology Service, Hospital Universitario de la Princesa, Universidad Autónoma de Madrid (UAM),
9 Instituto Investigación Sanitaria Princesa (IIS-IP), Madrid, Spain.

10 ^bCardiovascular Proteomics Laboratory, Centro Nacional de Investigaciones Cardiovasculares Carlos
11 III (CNIC), Madrid, Spain.

12 ^cCentro de Investigación Biomédica en Red, Enfermedades Cardiovasculares (CIBERCV), Madrid,
13 Spain.

14 ^dIntercellular Communication in the Inflammatory Response. Vascular Pathophysiology Area. Centro
15 Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain.

16 ^eUniversidad Autónoma de Madrid, Madrid, Spain.

17 ^fVideomicroscopy Unit, Instituto de Investigación Sanitaria La Princesa (IIS-Princes), Madrid, Spain.

18 ¹These authors contributed equally to this work.

19 ²Corresponding authors: fsmadrid@salud.madrid.org, jvazquez@cnic.es, noa.martin@salud.madrid.org

20

21 **Address:** Noa Beatriz Martín-Cofreces. Immunology Service. Instituto de Investigación Sanitaria La
22 Princesa. Planta 1. Diego de Leon 62, 28006 Madrid, Spain. Phone: +34915202000-Ext 11607.

23

24

25 **Abstract**

26 Mass-tolerant open search methods allow the high-throughput analysis of modified peptides by mass
27 spectrometry. These techniques have paved the way to unbiased analysis of post-translational
28 modifications (PTMs) in biological contexts, as well as of chemical modifications produced during the
29 manipulation of protein samples. In this work, we have analyzed in-depth a wide variety of samples of
30 different biological origin, including cells, extracellular vesicles, secretomes, centrosomes and tissue
31 preparations, using Comet-ReCom, a recently improved version of the open search engine Comet-
32 PTM. Our results demonstrate that glutamic acid residues undergo intensive methyl esterification when
33 protein digestion is performed using in-gel techniques, but not using gel-free approaches. This effect
34 was highly specific to Glu and was not found for other methylable residues such as Asp.

35

36 PTMs are known to exert a pivotal role in the control of signaling cascades and protein structure and
37 functionality by regulating protein localization, activity, folding, interaction and stability, and in recent
38 years have been implicated in pathological processes as tumorigenesis [1], neurodegeneration [2, 3]
39 and atherosclerosis [4]. While mass spectrometry (MS) is the method of choice for PTM identification
40 and quantitation, conventional (also termed *closed*) database search approaches followed in
41 proteomics research rely on matching exact mass differences, therefore restricting PTM analyses to a
42 few modifications (typically less than four) to keep a reasonably sized search space. This limitation has
43 been circumvented in recent years by mass-tolerant database searching methods [5] such as
44 MSFragger [6], Comet-PTM [7] or Comet-ReCom, a recently improved version of the former [8].

45 Re-analyses with Comet-ReCom of the LC-MS/MS data from centrosomal-enriched fractions obtained
46 from human T lymphocytes [9] to study the presence of PTMs revealed a strikingly high proportion of
47 the total peptide-spectrum matches (PSMs) bearing a +14.01565 Da modification (corresponding to
48 methylation, (see Suppl. Fig. 1) at Glu residues (3,48% with this modification) (see 'proteome identifier
49 5' in Fig. 1B). These samples were processed using the in-gel digestion procedure, a technique that is
50 advantageous in some contexts because it uses the gel matrix as a reaction chamber to eliminate
51 contaminants or for the addition and washing of reagents [10]. There is previous preliminary evidence,
52 based on MS analysis of a very limited number of peptides (5) that the acidified methanol solutions
53 used for gel staining/destaining may produce artifactual methylation of Glu and, to a lesser extent, of
54 other residues such as Asp [11]. This is a well-known reaction whereby a carboxylic acid undergoes
55 esterification by an alcohol in acidic medium. However, in our case the methylation appeared to be
56 highly specific to Glu, being virtually undetectable in Asp (<0.1%, 35-fold lower than Glu) or other
57 residues amenable to methylation (*i.e.* Lys, Arg, Glu, His, Asn, Gln, and Cys) ('proteome identifier 5' in
58 Fig. 1B). This was particularly remarkable given the high similarity between Glu and Asp, which differ
59 only in the presence of an extra methylene group in the side chain of the former. We therefore decided

60 to take advantage of open search methods to analyze this effect in more detail in a wide range of
61 samples of different origins.

62 We applied Comet-ReCom to the analysis of distinct cell, extracellular vesicle, secretome, centrosome
63 and tissue preparations digested either in-gel (using acidified methanol in the staining/destaining
64 procedure) [12] or on-filter (without methanol) [13] (Fig. 1A, see Supplementary Table 1). We found a
65 considerably higher prevalence of the +14.01565 Da modification at Glu residues in the proteomes
66 digested in-gel as compared to on-filter digestion (Fig. 1B), but not at Asp or other methylable residues,
67 strongly suggesting that the modification produced by in-gel treatments was highly specific for Glu. Of
68 note, the prevalence of Asp residues bearing the +14.01565 Da modification was very low in
69 comparison and was not affected by the digestion method (Fig. 1C). Note also that the relative
70 abundance of Glu and Asp residues was quite similar across all the proteomes analyzed (Fig. 1D),
71 ruling out the possibility that this effect was due to a higher Glu abundance in these samples. To further
72 verify the Glu specificity towards the in-gel digestion artefact, we applied Comet-Recom to LC-MS/MS
73 data obtained from the same set of mouse embryonic fibroblast samples digested using both in-gel and
74 on-filter protocols [14]. The results confirmed that the in-gel protocol produces a highly specific methyl
75 esterification of Glu but not of other methylable residues (see 'proteome identifiers 6 and 30' in Fig. 1B).
76 The markedly different reactivity of Glu and Asp residues towards methanol in acidic media may be
77 attributed to the distinct interaction of the side-chain carboxy group with the peptide backbone [11]. It is
78 noteworthy that, despite their chemical similarity, significant differences between Glu and Asp reactivity
79 have been recently reported using molecular dynamics simulations [15]. Regarding the other
80 methylable residues, the reaction of methanol with His, Lys and Arg residues is hindered in acidic
81 media due to the protonation of their side chains. These results were also observed in other proteomic
82 data generated by other research groups (Fig. 1B protein identifiers 7, 8 and 31) (16,17). Our
83 observation that in-gel digestion methods produce a highly specific methylation of Glu will serve the

84 proteomics community to better interpret open search-based PTM analysis performed using this
85 technique.

86 **Figure Legends**

87 **Figure 1: (A)** Schematic representation of the unbiased PTM analysis based on Comet-ReCom open
88 database search, performed with different cell, extracellular vesicle (EV), secretome, centrosome and
89 tissue preparations digested in-gel or on-filter. **(B)** Amino acid distribution in methylable residues of the
90 +14.01565 Da modification across the cell, extracellular vesicle, secretome, centrosome and tissue
91 samples digested in-gel or on-filter (see Supplementary Information). **(C)** Percentage Glu and Asp
92 residues bearing the +14.01565 Da modification across the cell, extracellular vesicle, secretome,
93 centrosome and tissue samples digested in-gel or on-filter. **(D)** Percentage of PSMs containing Asp or
94 Glu across samples.

95

96 **Acknowledgments.** Fig. 1A was created with Biorender. This study was supported by the Spanish
97 Ministry of Science and Innovation, Agencia Estatal de Investigación by competitive grants PID2021-
98 122348NB-I00, PID2022-141890B-I00, PID2020-120412RB-I00, PDC2021-121797-I00 and PGC2018-
99 097019-BI00 funded by MICIU/AEI/ 10.13039/501100011033 and by “ERDF A way of making Europe”,
100 PLEC2022-009298, PLEC2022-009235 and EQC2021-007053-P funded by
101 MICIU/AEI/10.13039/501100011033 and by “European Union NextGenerationEU/ PRTR”, and
102 S2022/BMD-7333-CM (INMUNOVAR-CM) and P2022/BMD7209 (INTEGRAMUNE) funded by
103 Comunidad de Madrid. CIBER Cardiovascular (CB16/11/00272, CB16/11/00277) Fondo de
104 Investigación Sanitaria del Instituto de Salud Carlos III; co-funding by Fondo Europeo de Desarrollo
105 Regional (FEDER). The project leading to these results has received funding from “La Caixa”
106 Foundation under the project codes LCF/PR/HR22/52420019 and LCF/PR/HR23/52430018. MLP is
107 supported by a FPI fellowship (PRE2021-097478). ALG is supported by a FPU fellowship
108 (FPU18/03882). RBR is supported by a FPU fellowship (FPU20/03365). CAD is supported by a FPI

109 fellowship (PRE2019-090019). The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the
110 Ministerio de Ciencia, Innovación Y Universidades (MICIU) and the Pro CNIC Foundation), and is a
111 Severo Ochoa Center of Excellence (grant CEX2020-001041-S funded by
112 MICIU/AEI/10.13039/501100011033).

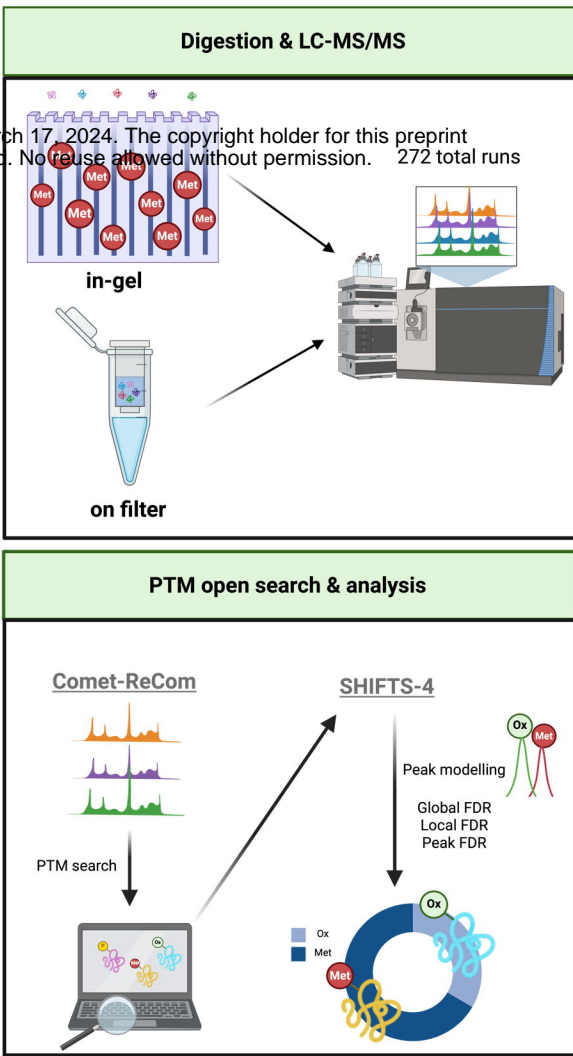
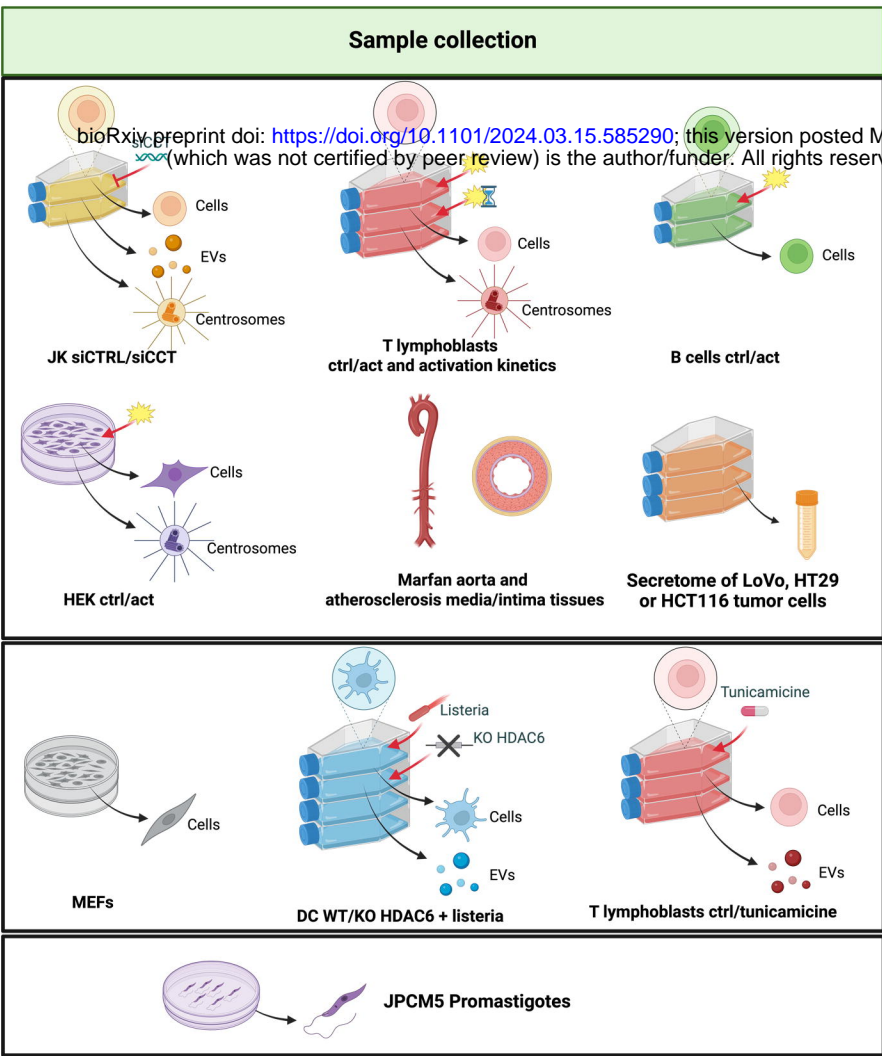
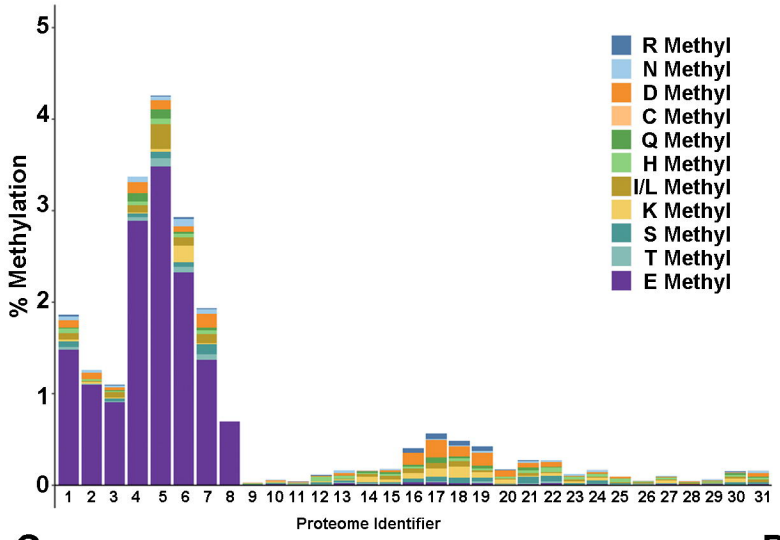
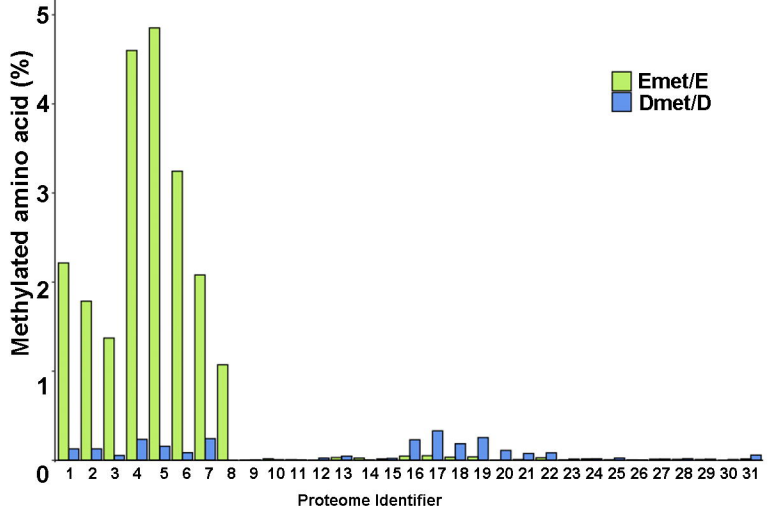
113

114 References

- 115 1. Liu, R., et al., mTORC1 activity regulates post-translational modifications of glycine
116 decarboxylase to modulate glycine metabolism and tumorigenesis. *Nat Commun*, 2021. **12**(1):
117 p. 4227.
- 118 2. Alquezar, C., S. Arya, and A.W. Kao, Tau Post-translational Modifications: Dynamic
119 Transformers of Tau Function, Degradation, and Aggregation. *Front Neurol*, 2020. **11**: p.
120 595532.
- 121 3. Junqueira, S.C., et al., Post-translational modifications of Parkinson's disease-related proteins:
122 Phosphorylation, SUMOylation and Ubiquitination. *Biochim Biophys Acta Mol Basis Dis*, 2019.
123 **1865**(8): p. 2001-2007.
- 124 4. Pirillo, A., et al., Impact of protein glycosylation on lipoprotein metabolism and atherosclerosis.
125 *Cardiovasc Res*, 2021. **117**(4): p. 1033-1045.
- 126 5. Chick, J.M., et al., A mass-tolerant database search identifies a large proportion of unassigned
127 spectra in shotgun proteomics as modified peptides. *Nat Biotechnol*, 2015. **33**(7): p. 743-9.
- 128 6. Yu, F., et al., Identification of modified peptides using localization-aware open search. *Nat*
129 *Commun*, 2020. **11**(1): p. 4065.
- 130 7. Bagwan, N., et al., Comprehensive Quantification of the Modified Proteome Reveals Oxidative
131 Heart Damage in Mitochondrial Heteroplasmy. *Cell Rep*, 2018. **23**(12): p. 3685-3697 e4.
- 132 8. Laguillo-Gomez, A., et al., ReCom: A semi-supervised approach to ultra-tolerant database
133 search for improved identification of modified peptides. *J Proteomics*, 2023. **287**: p. 104968.
- 134 9. Martin-Cofreces, N.B., et al., The chaperonin CCT controls T cell receptor-driven 3D
135 configuration of centrioles. *Sci Adv*, 2020. **6**(49).
- 136 10. Martinez-Acedo, P., et al., A novel strategy for global analysis of the dynamic thiol redox
137 proteome. *Mol Cell Proteomics*, 2012. **11**(9): p. 800-13.
- 138 11. Haebel, S., et al., Electrophoresis-related protein modification: alkylation of carboxy residues
139 revealed by mass spectrometry. *Electrophoresis*, 1998. **19**(5): p. 679-86.
- 140 12. Bonzon-Kulichenko, E., et al., A robust method for quantitative high-throughput analysis of
141 proteomes by ¹⁸O labeling. *Mol Cell Proteomics*, 2011. **10**(1): p. M110 003335.
- 142 13. Wisniewski, J.R., et al., Universal sample preparation method for proteome analysis. *Nat*
143 *Methods*, 2009. **6**(5): p. 359-62.
- 144 14. Bonzon-Kulichenko, E., et al., Improved integrative analysis of the thiol redox proteome using
145 filter-aided sample preparation. *J Proteomics*, 2020. **214**: p. 103624.
- 146 15. Lemke, T., et al., Three Reasons Why Aspartic Acid and Glutamic Acid Sequences Have a
147 Surprisingly Different Influence on Mineralization. *J Phys Chem B*, 2021. **125**(36): p. 10335-
148 10343.
- 149 16. Almeida-Marques, C., et al., Secretome processing for proteomics: A methods comparison.
150 *Proteomics*, 2024: p. e2300262.
- 151 17. Sanchiz, A., et al., The Experimental Proteome of *Leishmania infantum* Promastigote and Its
152 Usefulness for Improving Gene Annotations. *Genes (Basel)*, 2020. **11**(9).
- 153

154 **Author contribution.** MLP, data curation, formal analysis, validation, writing of original draft, Fig 1; EC,
155 data curation, formal analysis, validation, writing of original draft; IJ, data curation, formal analysis,
156 validation, writing of original draft; ALG, software; RBR, software; CAD, software; CP, data curation;
157 EC, data curation, formal analysis; FSM, JV and NBMC: conceptualization, funding, supervision, writing
158 of original draft.

159 **Data availability.** All data included in this study will be available (see Suppl. Material).

A**B****C****D**