UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR

TRABAJO FIN DE MÁSTER

Desarrollo de nuevas metodologías para el análisis global de modificaciones post-traduccionales en muestras biológicas

Máster Universitario en Bioinformática y biología computacional

Autor: Devesa Arbiol, Cristina Amparo

Tutores: Vázquez Cobos, Jesús Mª y Bonzón Kulichenko, Elena
Ponente: Martínez Muñoz, Gonzalo
Departamento de Ingeniería Informática

FECHA: junio, 2021

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR

TRABAJO FIN DE MÁSTER

Development of new methodologies for the global analysis of post-translational modifications in biological samples

Máster Universitario en Bioinformática y biología computacional

Autor: Devesa Arbiol, Cristina Amparo

Tutores: Vázquez Cobos, Jesús Mª y Bonzón Kulichenko, Elena
Ponente: Martínez Muñoz, Gonzalo
Departamento de Ingeniería Informática

FECHA: junio, 2021

General index:

Figures Index:

Tables Index:

INTRODUCTION

Post-translational modifications (PTMs) are covalent processing events that change the properties of a protein subsequently to its synthesis. Far from being mere “decorations”, PTMs of a protein can determine its activity state, localization, turnover, and interactions with other proteins [1].

Accordingly, multiple diseases could be caused or manifested by PTMs. For instance, an investigation carried out in our laboratory through the characterization of tissue-specific PTMs demonstrated that mitochondrial hereroplasmy induced oxidative modifications in the heart that mainly affects the proteins of the oxidative phosphorylation system (OXPHOS) [2]. Some studies relate oxidative post-translational modifications [3,4], citrullination, carbamylation [5], glycosylation [6,7], homocysteinylation [8], low density lipopoprotein (LDL) oxidation [9], matrix gamma-carboxyglutamic acid protein post-translational modification and cell specific histone modification with atherosclerosis disease [10,11]. These findings highlight the clinical importance of PTMs in the expansion of this significant disease, and their potential use as biomarkers.

Mass spectrometry (MS) is the most powerful tool to characterize PTMs. It is estimated that during traditional MS/MS sequence database searches a huge amount [12], about 60%, of the spectra cannot be assigned since they stem from peptides that contain unknown sequence variations or post-translational or chemical modifications. An exact match between observed and predicted fragments is required and increasingly narrow (ppm) precursor and fragment tolerances are used due to the of high-resolution instruments. However, the classical approach has several limitations since the modification to be searched must be selected in advance and that only up to eight PTMs are searched at the same time, since otherwise would compromise search time and identification reliability [13,14]. With this regard, the open search strategy *(“Open Search”)* has revolutionized the field, based on wide mass windows for the precursors and narrow for the fragments, allowing the detection of any kind of modification detectable by MS, successfully identifying modified peptides at an unprecedented scale by implementing tolerances in the order of hundreds of Daltons for the precursor mass [15]. Nonetheless, originally, the approach was only capable of identifying half of the potentially detectable modifications and without locating the position of the modified residue. In addition, there was a lack of tools for the quantitative analysis of PTM.

In the recent years, the Open Search-based tools for unbiased PTM identification have experienced a quick development in the proteomics community [1-13]. The Comet-PTM pipeline developed in our laboratory [2], constitutes a breakthrough in Open Search PTM analysis, since it calculates the mass difference between the experimental and all the possible theoretical precursors present in the database, adding this variation in mass (Δmass) to all the fragments and calculates a score for all the resulting sequence candidates, keeping the candidate with the best score, which is also the modified site. Thanks to this feature, CometPTM outperforms the traditional OpenSearch strategy,

identifying all detectable modifications and their positions within the peptide sequence. Moreover, the pipeline includes tools for the simultaneous quantitative analysis of modified peptides, proteins and functional categories on the bases of a robust statistical model [16].

Although the hypothesis-free PTM analyses constitute a radical improvement in the comprehensive study of PTMs by MS, they still face many challenges. For instance, the already high complexity of the Δmass histograms is increased due to unaccounted semi-tryptic or non-tryptic cuts, monoisotopic peak assignment errors, adduct formation and ammonia or water losses, among other factors. It is thus fundamental to be able to take into account all these peptide forms to achieve unbiased quantitative results. Additionally, in many cases the quality of the spectrum is not adequate enough to accurately locate the modified peptide position. In such cases the modification is found spread along the whole peptide sequence, increasing the variability among experiments and lowering the statistical power to detect abundance changes. Finally, the PTM analysis results in very long list of Δmasses and there are no tools to classify them into ontologies, hindering a system-wide quantitative study that would most definitely help to reveal the meaningful biological information.

Atherosclerosis and subsequent cardiovascular diseases are the leading causes of death worldwide [17]. It is a lipid-driven, chronic inflammatory disease that develops plaque formation at specific sites of the arterial tree leading to hardening and narrowing of the arterial lumen [18]. Diabetes, current smoking, hypertension, hyperlipidemia and male gender are common risk factors of this disease related with vascular inflammation [19]. PTMs also have been implicated in atherosclerosis [2-11]. With this regard, using the Low Densitty Lypoprotein receptor deficient (LDLR KO) mouse model, it was demonstrated that atherosclerosis can be attenuated by the pharmacological inhibition of tyrosine phosphatase 1B [20], and the genetic deficiency of Tyrosyl protein Sulfotransferase [21], although the molecular mechanisms involved are still unknown. PTMs also play a role in vascular calcification [10], a process that occurs in association with atherosclerotic lesions development [22]. Thus matrix gamma-carboxyglutamic acid protein (MGP) expressed by vascular smooth muscle cells (VSMC) functions as an inhibitor of vascular calcification [10]. MGP function depends on the enzymatic carboxylation of specific Glu residues converting them to γ-carboxyglutamic, a much stronger calcium chelator. Nonetheless, the specific molecular mechanisms linking altered PTMs with the development of atherosclerosis, have not been elucidated. Thereby, it is essential not only to develop bioinformatics tools that improve PTMs identification and quantitation, but also to apply them to the comprehensive PTM proteomics analysis in a model of atherosclerosis to gain a fundamental global understanding of this disease.

OBJECTIVES

This work has two main objectives. The first goal is to improve *Open-search*-based PTM identification and quantitation strategies by developing tools that would enable:

- to reduce the artefactual sources of Δmass histogram complexity;
- to increase the accuracy in the localization of the modified site;
- to carry out quantitative ontology-based PTM statistical analysis.

The second goal is to demonstrate the performance of the newly developed tools to perform a comprehensive and unbiased characterization of the plasmatic PTM changes that are associated with atherosclerosis in a mouse model. This new knowledge would improve our molecular understanding of this devastating disease.

MATERIAL AND METHODS

Dataset

The dataset used to analyze PTMs was a high-throughput proteomic study from an ongoing project in the laboratory with mice plasma labelled with TMT and searched with CometPTM [23]. Male LDLR-KO pro-atherogenic mice were and fed with a normal diet (CHOW) or a high fat diet (HFD). At 16 weeks of age, blood was obtained. Samples digestion was performed using the filter-aided sample preparation digestion kit (FASP, Expedeon) as previously described [24], considering minor modifications [23], with trypsin and labelled using 10 plex- TMT reagents according to manufacturer's instructions (Thermo Fisher Scientific TM) [23]. After MS, the raw files were searched with CometPTM [2] with some modifications: trypsin digestion with 2 missed cleavages, +-500 Da precursor ion tolerance and 0.02 Da fragment ion tolerance. TMT labeling (+229.162932 Da) at peptide N-terminal and Lys, as well as Cys carbamidomethylation (+57.021 Da) were selected as a fixed modifications [23]. CometPTM output files were used as input files for a pipeline composed by the modules developed by Andrea Laguillo and Rafael Barrero (ongoing PhD Thesis), DMCalibrator, DMModeller, PeakInspector, PeakSelector, PeakAssignator modules can be downloaded from https://github.com/CNIC-Proteomics/SHIFTS-4.

Development of Bioinformatic tools

FDRFilterer, DM0Solver, TrunkSolver, PDMTableMaker, SiteSolver and Sticker modules were implemented in Python 3 and can be used in both Windows and Linux operating systems. It is currently compatible with conventional LC-MS/MS data and can be used with common proteolytic enzymes (including trypsin and Lys-C).

FDRFilterer

FDRFilterer filters FDR according to the conditions set by the user, in the configuration file.

Input files:

- .tsv
- Configuration file (.ini). Default configuration file can be modified by user.
    - FDRFilterer parameters:
        - *GlobalThres*: global FDR threshold. Maximum Global FDR.
        - *PeakThres*: peak FDR threshold. Maximum Peak FDR.
        - *LocalThres:* local FDR threshold. Maximum Local FDR.
        - *GlobalFDR_column_name:* name of the column containing Global FDR values.
        - *PeakFDR_column_name:* name of the column containing Peak FDR values.

- *LocalFDR_column_name:* name of the column containing Local values.
- *Label_column_name:* name of the column that indicates label.
- *decoys_naming:* parameter that indicates how decoys are named.

Output files:

- FDRFilterer output (default suffix: “FDRFilterered”)
- A log file (default suffix: " FDRFilterered_logFile")

DM0Solver

DM0Solver is a module that detects whether a modified peptide, termed as pdm (a peptide form defined by peptide sequence, Δmass and position) from here on, has a Δmass belonging to a list provided by the user (Table 1), for that purpose relative error (ppm) is calculated. In such a case, the Δmass is appended at the end of the clean sequence (new output column). The corresponding label and the error generated selecting that label are added in two additional columns (Table 2). If the Δmass does not belong to the list, the module passes the modified sequence without any modification to the output columns (Figure 1A). To carry out these tests, the relative error (ppm) must be calculated:

$$Relative\ error\ (ppm) = abs(\frac{(Theoretical\ mass + Experimental\ mass\ ) * 1000000}{Theoretical\ mass + Label\ mass})$$

DM0Solver needs the following input files:

- .tsv file
- Configuration file (.ini), a default configuration file that can be modified by user, containing:
  - DM0Solver parameters:
    - *Relative_Error_ppm:* relative error (ppm) allowed.
    - *Exp_mh_column_name:* calibrated experimental mh column name.
    - *Theo_mh_column_name:* theoretical mh column name.
    - *Sequence_column_name:* sequence with Δmass column name.
    - *DM0Sequence_output_column_name:* column name of the output in which the selected sequence is annotated.
    - *DM0Label_output_column_name:* column name of the output in which the chosen label is annotated.
    - *DM0Label_ppm_output_column_name:* column name of the output in which the calculated error in ppm is annotated.
    - *output_file_suffix:* chosen suffix for output file.

  - DM0Solver list (list of masses with their corresponding label):
    An example of input list for DM0Solver configuration file can be observed in Table 1. The first column shows examples of labels (DM0 for zero Δmass, or DM0;C13 for carbon 13 Δmass), while the second column contains de mass in Da of each label. DM0Solver uses this list to assign a label, as long as it meets the error threshold.

Table 1. *Example of input list for DM0Solver configuration file*

Label	**Mass (Da)**
DM0	0
DM0;C13	1.003355
Na adduct	21.981943
Ammonium_ adduct	17.026549
H20_loss	-18.010565

It delivers two output files:

- DM0Solver output (default suffix: “DM0S”)
  - New columns (Table 2):
    - *DM0Sequence*: output column in which the reassigned sequence is annotated.
    - *DM0Label:* output column in which the selected label is annotated.
    - DM0Label_ppm: output column in which the error, that is obtained selecting the label, is annotated.
- A log file (default suffix: "DS_logFile")

Table 2. *DM0Solver output columns example.*

DM0Sequence	DM0Label	DM0Label_ppm
YELELRPTGEVEQYSATATYELLK_-0.030	DM0	9.22
ECCHGDLLECADDRAELAK_21.954	Na_adduct	9.89
LLGS[-499.369414]MLVLVLGHHLGK		226274.34

TrunkSolver

TrunkSolver detects whether the Δmas, in a pdm element may be explained by the truncation of a non-modified peptide inside the sequence of the corresponding protein or, by a truncation and the presence of a Δmass belonging to a list provided by the user (Table 3), for that purpose relative error (ppm) is calulated. In such a case, the Δmass is appended at the end of the clean sequence (“TrunkSequence” output column), while the corresponding label and the recalculated Δmass are added in additional columns named “TrunkLabel” and “TrunkDM”, respectively. If TrunkSolver is unable to explain the Δmass by a truncation, then it passes the modified sequence and its original Δmass without any modification to the output columns. In both cases five extra columns(“New_DM”,“New_Theo_mh”,“Match_number”,“Possible_option”, “Trunk_stats _mods”) will be created for subsequent **PeakAssignatior** execution, which is necessary for

the correct peak assignation (Andrea Laguillo ongoing PhD Thesis) (Table 4, Figure 1 B). Relative error will be calculated as:

$$Relative\ error\ (ppm) = abs(\frac{(Theoretical\ mass + Experimental\ mass\ ) * 1000000}{Theoretical\ mass + Label\ mass})$$

TrunkSolver needs as input files:

- .tsv file
- .fasta file
- MassMod.ini. Default configuration file can be modified by user.
- Configuration file (.ini). Default configuration file can be modified by user:
  - TrunkSolver parameters:
    - *Relative_Error_ppm:* relative error (ppm) allowed.
    - *Exp_mh_column_name:* calibrated experimental mh column name.
    - *Theo_mh_column_name:* theoretical mh column name.
    - *Sequence_column_name:* sequence with Δmass column name.
    - *Calibrated_Delta_MH_column_name:* calibrated Δmass mh column name.
    - *MasterProtein_column_name:* master Protein column name.
    - *static_modifications_column_name:* static modifications column name.
    - *Decnum:* decimals points required in TrunkSequence column.
    - *X:* number of positions to the right and left, that the TrunkSolver is allowed to extend from the original Δmass site.
    - *New_Deltamass_output_column_name:* new Δmass column name.
    - *New_Theo_mh_output_column_name:* new theoretical mh column name.
    - *TrunkSequence_output_column_name:* column name of the output where the chosen sequence is annotated.
    - *TrunkDM_output_column_name:* column name of the output where the recaulcutaed Δmass is annotated, taking in to account the label.
    - *TrunkLabel_output_column_name:* column name of the output where the chosen label is annotated.
    - *TrunkLabel_ppm_output_column_name:* column name of the output where the calculated error in ppm is annotated.
    - *Static_modifications_position_output_column_name:* column name of the output where the new fix modifications positions are annotated
    - *Matchnumber_output_column_name:* column name of the output where possible options number is annotated.
    - *Possible_option_output_column_name:* column name of the output where all possible options.
    - *output_file_suffix:* chosen suffix for output file.
  - TrunkSolver list (list of masses with their corresponding label).

An example of input list for TrunkSolver configuration file can be observed in Table 3. The first column shows examples of labels, TMT: Δmass of tandem mass tag, 2TMT; two Δmasses of tandem mass tags,-TMT: minus Δmass of tandem mass tag, ,-2TMT: minus two Δmasses of tandem mass tags. The second column contains de mass in Da of each label. TrunkSolver uses this list to assign a label, as long as it meets the error threshold.

Table 3. *Example of input list for TrunkSolver configuration file*

Label	**Mass** Da()
TMT	229.162932
2TMT	458.325864
(-)TMT	-229.162932
(-)2TMT	-458.325864

TrunkSolver delivers the following output files:

- TrunkSolver output (default suffix: "_TS")
  - New columns (Table 4):
    - *TrunkSequence:* output column name in which reassigned sequence is annotated
    - *Trunk_DM:* output column name in which recalculated Δmass taking in to account the labels.
    - *Trunk_Label:* output column name in which the selected label is saved. If the Δmass corresponds to a combination between one of the labels in the configuration file and a cut, the type of cut also will be noted (Figure1 B, Table 4). The label will be TrypticCut if the choice of label involves a tryptic cut of the sequence, or a Truncation if it involves a non-tryptic cut.
    - *Trunk_Label_ppms*: output column in which the error, that is obtained selecting the label, is annotated.
    - *New_Theo_Mh:* output column name in which the recalculated theoretical mass is annotated.
    - *New_DM:* output column name in which the recalculated Δmass is saved.
    - *Trunk_stats_mods:* output column name in which the new static modifications positions are saved, since with the changes in the sequence this could vary.
    - *Match_number:* output column name in which the number of possible options allowed by TrunkSolver is annotated.
    - *Possible_options:* output column name in which all possible options allowed by TrunkSolver are annotated
- A log file (default suffix: "TS_logFile")

Table 4. *Example of TrunkSolver output file columns, considering input file*

Input

DM0Sequence	Theoretical_mh	Modifications
LGEHNLDVLEGNEQFLNA[-499.331973]AK	2669.42	1_S_0.000000_N
GTFASLS[229.141559]ELHCDK	1923.00	1_S_0.000000_N, 11_S_57.021464, 13_S_229.162932
TLER[-499.281621]EACLLNANK	1990.115	1_S_0.000000_n, 7_S_57.021464, 13_S_229.162932

Output

TrunkSequence	TrunkDM	TrunkLabel	TrunkLabel_ppm	New_Theo_mh	New_DM	Static_modifications_position	match_number	Possible_option
LGEHNLDVLEGNEQFLN_-0.00015	-0.00015	Truncation; DM0	0.06	2170.09	-0.00015	1_L_229.162932_N	1	1
GTFASLSELHCDK_229.141	229.14	(+)TMT	9.93	2152.167	-0.02	1_G_229.162932_N,11_S_57.021464,13_S_229.162932	1	1
EACLLNANK_-0.006	-0.006	TrypticCut; DM0	4.14	1490.84012	-0.006	1_E_229.162932_N,3_S_57.021464,9_S_229.162932	1	1

PDMTableMaker, SiteSolver

PDMTableMakertMaker is an independent, but necessary, module for the proper operation of SiteSolver. PDMTableMaker contains the frequencies of the different sequence species (pdm) which allows SiteSolver to choose when there is a tie between two positions. SiteSolver will correct Δmass positions whenever appropriate.

PDMTableMaker

PDMTableMaker calculates the scan frequency which is used by SiteSolver but also some parameters that will allow to build the necessary relations files for the subsequent processing of the results by the statistical quantitation workflow, as it will be later explained: p, q, pdm, pd, d, Theo_mh, ScanFreq, a, m, n, l, qna, qdna, A, M, L, N, qdNA (Table 5).

PDMTableMaker input files:

- tsv file.
- .fasta file.
- Configuration file (.ini). Default configuration file can be modified by user:
  - PDMTableMaker parameters:
    - *Sequence_column_name:* Sequence that contains the Δmass colum name.
    - *DM_column_name:* Δmass column name.
    - *Theo_mh_column_name:* Theoretical mh column name.
    - *MasterProtein_column_name:* Master Protein column name.
    - *output_file_suffix:* Chosen suffix for output file.
  - PDMTableMaker conditions:
    - *number_of_conditions*: Number of conditions considered.
    - *Condition*$_i$: Column name of condition i.
    - *Value*$_i$: Chosen value for condition *i*.

    The program can use as many conditions as it is desired, considering that a match will be made between the condition$_i$ and the value$_i$.

PDMTableMaker delivers two output files:

- PDMTableMaker output (default suffix: “PDM”). An example can be seen in Table 5.
  - New columns:
    - *p:* peptide.
    - pdm: peptide form defined by peptide sequence, Δmass and position. Ex: ABCD[xxx]EFGHK.
    - *pd:* ABCEDEFGHK:XXX (includes a set of pdm elements, do not confound with ABCEDEFGHK_XXX, which is a pdm, not a pd)
    - *m:* position in peptide. Ct (C-terminal) =-1, Nt (Nterminal)=0.
    - *l:* position in peptide, from right to left. Ct=,0 Nt = -1.
    - *n:* position in protein. Ct =length peptide plus 1, Nt = 0.
    - *a:* modified aminoacid. Ct = U, Nt = U.
    - *q:* protein identifier.
    - *M:* it is the m corresponding to the pdm with highest original PSM frequency in a pdm table.
    - *L:* l which corresponds with the razor m. Razor l, property of a pd.
    - *N:* n which corresponds with the razor m. Nt=0. Razor n, property of a pd.
    - *A:* a which corresponds with the razor m. Razor a, property of a pd.
    - *qdna:* information of q,d,n,a (ex: HPT:yyy:300:M)
    - *qna:* information of q,n,a (ex: HPT: 300:M)
    - *qdNA:* razor qdna, property of a pd
    - *Theo_mh:* Theoretical mh
    - *ScanFreq:* Scan frequency

- A log file (default suffix: " PDM_logFile")

Table 5. *Example of PDMTableMaker Output*

p	q	pdm	pd	d	Theo_mh	ScanFreq	a	m	n	l	qdna	qna	A	M	L	N	qdNA
FRTEDGFSLK	P32261	FRTEDGFSLK_-8e-05	FRTEDGFSLK:-8e-05	-8e-05	1657.9313	2	U				P32261:-8e-05::U	P32261::U	U				P32261:-8e-05::U
AAPQLPMEELVSLSK	O89020	AAPQLPMEEL[-2.0718]VSLSK	AAPQLPMEELVSLSK:-2.0718	-2.071	2071.187	1.0	L	10	447	6	O89020:-2.0718:447L	O89020:447:L	L	10	6	447	O89020:-2.0718:447L

SiteSolver

SiteSolver is a module made with the aim of detecting if a modified peptide has its Δmass in an incorrect position. In such a case, Δmass location within the sequence is corrected in the “SiteSequence” column. If the module does not find any possible position, it passes the modified sequence without any modification to the output column. Foremost, it is ascertained whether the amino acid position, in which the Δmass is originally located, is prohibited. This is tested up using the frequency chosen in the PrimaryList file (PrimaryList_colum_name) and ScanFrequency of PDMTableMaker ouput file. From the

lists, the mass with the lowest relative error of place will be chosen.. If the amino acid is allowed, for that Δmass, it passes the modified sequence without any modification. If the amino acid is prohibited, SiteSolver will analyze the contiguous amino acids to that position. Provided that only one of them is prohibited the sequence will be corrected by assigning the Δmass to the amino acid that is allowed, according to PrimaryList or SecondayList file. If both are allowed, for that Δmass, the position with the highest frequency in the PDMTableMaker output file is chosen, but if the frequencies are equal, it will be taken into account the order of appearance in the PrimaryList or SecondayList file. If none of the amino acids are allowed, the contiguous, at each end, will be analyzed. The process is repeated until the number of positions to be analyzed, on each side, exceeds the maximum (X parameter) (Figures 1 D,E,F,G). To carry out these tests, the relative error must be calculated:

$$Relative\ error\ (ppm) = abs(\frac{(Theoretical\ mass + Experimental\ mass\ ) * 1000000}{Theoretical\ mass + Label\ mass})$$

SiteSolver needs the following input files:

- .tsv file.
- PDMTableMaker output file.
- PrimaryList.txt file (user).
- SecondaryList.txt file (unimod).
  Secondary and Primary lists will contain 3 columns:
  1) DM: first column must contain the Δmass.
  2) Residue: second column must contain the residue. In case of N-terminal and C-terminal, must appear, NT and CT respectively.
- Configuration file (.ini). Default configuration file can be modified by user:
  - SiteSolver parameters:
    - *Relative_Error_ppm:* relative error (ppm) allowed.
    - *Theo_mh_column_name:* theoretical mh column name.
    - *Sequence_column_name:* sequence with Δmass column name.
    - *cal_Dm_mh_column_name:* calibrated Δmass MH name.
    - *PeakNaming:* Parameter that indicates how peaks are named.
    - *PeakAssignation_column_name:* name of column that contains peak assignation.
    - *x:* parameter that indicates the extension (left and right) from the original DM site of the amino acid positions to be analyzed.
    - *MinScanFreq:* parameter that indicates the frequency threshold for PDMTableMaker output file.
    - *MaxNSite:* parameter that indicates the number of amino acids allowed when the PDMTableMaker output file is analyzed.
    - *PrimaryList_column_name:* column name from Primary List wanted to be used for SiteSolver.

- *SecondaryList_column_name:* column name from Secondary List wanted to be used for SiteSolver.
- *SiteCorrection_PrimaryList_label:* SiteCorrection when the SiteList selected is from the Primary list.
- *SiteCorrection_SecondaryList_label*: SiteCorrection when the SiteList selected is from the Secondary list.
- *SiteSequence_column_name:* column name of the output in which the sequence is annotated.
- *SiteCorrection_column_name:* column name of the output in which the correction site is annotated,
- *SiteOption_column_name:* column name of the output in which the different possible options are annotated.
- *SiteDM_column_name:* column name of the output in which the selected Δmass is annotated.
- *SiteDMError_ppm_column_name:* column name of the output where the error of the selected DM is annotated.
- *Output_file_suffix:* chosen suffix for output file.

SiteSolver delivers two output files:

- SiteSolver output (default suffix: "SS")
  - New output columns:
    - *SiteSequence:* sequence with the Δmass.
    - *SiteCorrection:* initial residue (position that has been displaced) and final residue.
    - *SiteOption:* if there is a tie, which are all the possible options.

- A log file (default suffix: " SS_logFile")

Sticker

Sticker is a module that for each experimental Δmass detects,by calculating the relative error the closest theoretical one from a DMID_list file . This file will contain some conditions and labels. When comparing PDMTabelMaker output file with DMID_list, all the conditions established in DMID_list, the labels and the error that produces the corresponding label selection, will be annotated in new columns. To do so, the relative error will be calculated:

$$Relative\ error\ (ppm) = abs(\frac{(Theoretical\ mass + Experimental\ mass\ ) * 1000000}{Theoretical\ mass + Label\ mass})$$

As input files Sticker requires:

- PDMTableMaker output file
- DMID_list.txt:
  - Column names:
    - d: Column with the Δmass.

        - Cond_: all columns with the conditions that must be met, should start with "Cond_a_". If "a" has to be met, a column named "Cond_a" must appear in he DMID_list.txt file. It should be noted that the columns denoting conditions must include after the prefix "Cond_" (condition), names of the columns present in PDMTableMaker, for instance "Cond_d", "Cond_p", "Cond_pdm". If the indicated parameter is not found in the PDMTableMaker file, the comparison will not be possible. The name of the columns that contains the label to be added will begin with the prefix "Lab_".

          The label for those rows that do not have any label will be specified at the first line of the DMID_list file. All columns will be empty except for the first column that will start with "Lab_" (Figure 1 G).

- Configuration file (.ini). Default configuration file can be modified by user:
    - Sticker parameters:
        - *Relative_Error_ppm:* relative error (ppm) allowed.
        - *Theo_mh_column_name*: theoretical mh column name.
        - *Sequence_column_name:* sequence with Δmass column name.
        - *Selected_DM_column_name:* selected Δmass column name.
        - *StickerLabel_User_output_column_name:* column name of the output where the chosen label is annotated.
        - *StickerLabel_ppm_User_output_column_name:* column name of the output where the calculated error in ppm for the selected label is annotated.
        - *output_file_suffix:* chosen suffix for output file.

It delivers two output files:

- Sticker output (default suffix: "Sticker"):
    - New columns:
        - *StickerLabel:* selected label from the user file. It will appear as many StikerLabel columns as "Lab_" columns appear in DMIDlist file.
        - *StickerLabel_ppm:* relative error calculated with the mass of the selected label from the user file. It will appear as many StikerLabel_ppm columns as "Lab_" columns appear in DMIDlist file.
- A log file (default suffix: "Sticker_logFile").

Workflow

CometPTM output files were used as input files for a pipeline composed by the modules developed by Andrea Laguillo and Rafael Barrero (ongoing PhD Thesis), as well as by the ones developed in this work. First the modules developed before this work: DMCalibrator, DMModeller, PeakInspector, PeakSelector, PeakAssignator, PeakFDRer were executed,

and then the modules developed in the present work FDRFilterer, DM0Solver, TrunkSolver, PeakAssignator again as it has been explained, PDMTableMaker, SiteSolver, PDMTableMaker and Sticker (Fig. 1H). Note, that when PDMTableMaker is executed the second time, it takes as input the SiteSolver corrected pdms. The configuration files containing the parameters used to execute all the pipeline can be found in Supplementary Tables 1-12 and Supplemenatry Lists 1-5.

Quantitative statistical analysis

The output files from SiteSolver and the PDMTable file from the second round of PDMTableMaker were used as input files for the quantitative statistical pipeline (Fig1H). To that end, for each scan from the SiteSolver output file the quantitative information from the TMT reporter intensities was extracted from the Quant_PSM.txt files previously generated by Proteome Discoverer (2.1, Thermo Scientific) during my previous Master Thesis [23], appended to the SiteSolver output file and the log2ratios (Xqps) as well as the scan weights (Vqps) were calculated [16,17,24-26]. Briefly, for each sample *i*, the values $x_{qps}=\log_2 A_i/C$ were calculated, where A_i is the intensity of the TMT reporter of the corresponding sample *i* in the MS/MS spectrum s coming from peptide *p* and protein *q*, and *C* is the average intensity of all the TMT reporters from the control samples. The SiteSolver output file was used to produce the *scan*-to-*pdm* relations tables. On the other hand, the PDMTable file was used to build the *pdm*-to-*protein*, the *pdm*-to-*protein_site_DM* (*qdna*) and the *qdna*-to-*protein_site* (*qna*) relations tables.

The quantitative information was integrated from the spectrum level to the *pdm* and then to the protein level on the basis of the WSPP model [16] using the Generic Integration Algorithm (GIA) [27]. Additionally, the algorithm was modified to include the integration of the normalized-by-protein *pdm* (*q_pdm*) quantitation to the protein_site_DM level (*q_pdm2qdna*), then to the protein_site level (*q_qdna2qna*), and finally, *qdna* and *qna* levels were integrated to the experiment Mean (*All*) (*q_qdna2ALL* and *q_qna2ALL*). All the integrations were performed as a part of an automated workflow (Figure 1 I), The software for statistical analysis of quantitative data can be downloaded from https://github.com/CristinaDevesaA/TFM

Computational services

Some PTMs identified were checked in the following public PTMs DB: CarbonylDB [28], GlyConnect [29], SwissPalm [30], PhosphoSitePlus [31], UniCarbKB [32], iPTMnet [33]. ΔM annotation has been performed according to a user curated DB (Ileana Beatriz González ongoing PhD Thesis and the Unimod DB [34]. All the developed tools can be downloaded from https://github.com/CristinaDevesaA/TFM

RESULTS

The following results were obtained by applying the developed tools to a high-throughput TMT proteomics experiment in order to characterize PTM in atherosclerotic mouse model plasma proteins.

FDRFilterer

As described in the Materials and Methods section, first the Δmasses were calibrated, peak-modelled and global, local and peak FDRs were calculated using the programs developed by Andrea Laguillo and Rafael Barrero (PhD Thesis in progress) (Figure 1H from Materials and Methods). The resulting text files contained all spectra coming from target and decoy proteins regardless of their FDR values. Thus, to be able to filter out non-reliable identifications in a fast and automatic way, we developed FDRFilterer. This program can quickly and easily handle several large files at the same time, producing a reliable dataset containing only target identifications with controlled FDR values at global, local and peak level. As a result, 217977 of 692215 peptide spectrum matches (PSM) passed the established filters.

DM0Solver and TrunkSolver

Comet-PTM calculates the experimental mass variation corresponding to a spectrum by comparing the theoretical mass, according to the peptide sequence, with the experimental mass detected by the mass spectrometer. The following formula is used (Figure 2 A):

$$\Delta\text{experimental mass} = \text{theoretical mass} - \text{experimental mass}$$

However, during careful data inspection, we realized that a big proportion of non-zero Δmasses actually were unmodified peptides (Figure 2B). Among those non-zero Δmasses were monoisotopic peak assignment errors, adducts with metals and loss of water and ammonia.

Additionally, many far negative Δmasses, in the overwhelming majority of the cases, had as a second candidate for the same scan identical sequence without a partially digested N- or C-terminal segment (Ileana Beatriz González Ongoing PhD Thesis). Therefore, many of those spectra with far negative Δmasses actually belonged to unmodified peptides that were mistakenly assigned by CometPTM to longer peptides. In part, this fact takes place due to the score we used to assign the PSM (Xcorr) which is proportional to the number of matched fragments [35] and thereby gives preference to larger peptides [36]. This Xcorr-associated effect is shown in Figure 2C were the Δmass decoy distribution is biased towards the negative Δmasses. An additional source of the artefactual far negative Δmasses were the unaccounted non-tryptic or semi-tryptic cuts since the search constrained the theoretical candidate sequences only to fully tryptic cuts. To resolve these unaccurate assignations we developed DM0Solver and TrunkSolver, with the common goal of detecting and recovering all those forms of unmodified peptides. Accordingly, DM0Solver

was employed to detect adduct formation, isotopomer errors, and water and ammonia losses/gains. TrunkSolver was set to find tryptic and non-tryptic cuts and their combinations with TMT losses/gains (a common artifact detected in this kind of isobaric labeling experiments).

We first executed DM0Solver, which for each scan calculated the composite mass of the theoretical $[M+H]^+$ plus each one of the Δmasses present in the user configuration file (Supplementary Table 4), that apart from Δmass =0, also included ^{13}C-errors, and some metal adducts. Afterwards, the program calculated the ppm mass difference between this composite mass and the calibrated experimental mass $[M+H]^+$. If the obtained ppm was below the user specified threshold (10 ppm in this case), the PSM was labelled by the corresponding Δmass from the configuration file. Furthermore, since the Δmass could not be ascribed to any specific position inside the peptide, the Δmass was appended at the end of the clean sequence in a new "DM0Sequence" output column (Materials &Methods Table 2). Otherwise, the modified peptide sequence was passed unchanged, and was used as input for TrunkSolver (Materials &Methods Figure 1H). Thus, to preserve DM0Solver results, TrunkSolver only considered peptide sequences from the "DM0Sequence" column with the Δmass in brackets, while the ones produced by DM0Solver were passed unchanged (Table 4 and Figure 1 H from Materials &Methods).

As it is schematized in Figure 1B from Materials and Methods, on the basis of the "DM0Sequence" and the corresponding protein sequence from the fasta file, TrunkSolver tested all possible tryptic and non-tryptic cuts and the combination with each one of the Δmasses set in the configuration file (Supplementary 5). For each one of these combinations, TrunkSolver chose the one that resulted in the lowest relative error as long as it was below the threshold established in the configuration file (10 ppm in the present work). The precursor tolerance used during the search was +-500 Da , the mean mass of an amino acid residue is around 100 Da, that is why the number of residues considered by the program ("x" parameter) were up to six positions away from the original Δmass site to the right or to the left. Thus, to reveal possible semi-tryptic or non-tryptic cuts, TrunkSolver tests were also extended beyond the peptide sequence N- or C-terminal and into the fasta protein sequence. If all the above-mentioned conditions were met, the new clean sequence with its new experimental Δmass appended at the end (the Δmass position becomes unknown), was annotated in the new "TrunkSequence" column. Moreover, if the removed/added segment of the sequence implicated a partial digestion, it was labelled as "tryptic cut" in the TrunkLabel column, while if it was a non-tryptic cut, it was labelled as "truncation". However, if TrunkSolver was unable to explain the Δmass, it passed the modified sequence and its original Δmass without any modification to the output columns. Therefore, when TrunkSolver changes the theoretical mass, it concomitantly changes the experimental Δmass. For this reason, and to re-establish the correct Δmass-to-peak assignment, it is necessary to run again *PeakAssignator* (Andrea Laguillo Barrero ongoing PhD Thesis) using the new columns produced by Trunkolver: TrunkSequence, New_DM and New_Theo_mh (Table 4 and Figure 1 H from Materials &Methods).

After using DM0Solver and TrunkSolver, 48.78% of the spectra (106326 of 217978) were recovered as unmodified peptides: 36.47% by DM0Solver and 12.31% by TrunkSolver (Figure 2 D and E). Of the percentage tagged by DM0Solver, as expected, most of the Δmasses (70.39%) were equal to zero. However, this tool also recovered an important additional 22.67% of the spectra related to ^{13}C errors (18.44%, 3.33% and 0.90% belonging to one, two and three ^{13}C, respectively) and a 1.19% to water loss. There was also a percentage of 5.75% pertaining to adducts, specifically 1.60% to sodium adducts, 2.97% and 1.18% gains and losses of ammonium adducts, respectively (Figure 2D).

On the other hand, attending to the percentage recovered by TrunkSolver, as expected, the vast majority (45%) were scans coming from peptides reassigned to the Δmass=0 peak due to tryptic cuts, and a minor proportion (9%) corresponded to non-tryptic cuts that were also reassigned to the Δmass=0 peak. In other matters**,** the 39.66% was related to TMT-gain (32.99% TMT without partial digestion, 4.33% TMT and tryptic cut and 2.34% TMT and non-tryptic cut). Moreover, a 3.63% was related to TMT-losses (1.94% TMT-loss without partial digestion, 1.68% TMT-loss and non-tryptic cut and 0.01% TMT-loss and tryptic cut.). 2.59% of cases were related to the double gain or loss of TMT (1.57% double gain of TMT, 0.49% non-tryptic cut and double gain of TMT, 0.46% double loss of TMT with tryptic cut, 0.06 % double loss of TMT and non-tryptic cut and 0.01% double loss of TMT) (Figure 2E).

The performance of DM0Solver and TrunkSolver is shown in the Figure 2F, where the Δmass peak frequency difference between PeakAssignator2 and PeakAssignator1 is presented. That is, after using DM0Solver and TrunkSolver (PeakAssignator2) compared to the original situation before using DM0Solver and TrunkSolver (PeakAssignator1), the intensity of Δmasses peaks in the negative zone is drastically reduced and we recover a lot of unmodified and extra TMT-labelled peptides, as well as peptides lacking one TMT moiety. The most prominent negative Δmasses decreasing were -357.259079Da and -156.102444 Da. Those Δmasses correspond to an extra lysine labelled with TMT and to an extra arginine, respectively, and were recognized as non-modified peptides by TrunkSolver. In addition, the frequency of +229.162932 Da Δmass increases owing to many peptides getting cleaned from their extra sequence segments due to the Xcorr effect, and this uncovers the TMT mass.

PDMTableMaker and SiteSolver

Comet-PTM automatically assigns the modification to the residue in the peptide sequence that produces the best score and that best explains the fragmentation data. Assigning modifications to specific residues is considered a much less reliable process than identifying peptides, partly because there is frequently insufficient information to determine the exact modified residue [37]. Consequently, one out of five of the sites are misallocated [2], leading to data dilution during modified peptide quantitation and loss of statistical power to detect changes. Thus, the goal of SiteSolver is to detect the wrongly

assigned modified residue, and to relocate them to the most probable allowed site (Figure 2G)

For each experimental Δmass SiteSolver selected the allowed sites first from a user-curated list of theoretical Δmasses and their corresponding modifiable residues (See Supplementary Lists, 1,2 , Figure 1H from Materials &Methods), using a user-defined ppm mass error (<10 ppm in this work) calculated as described in the Materials and Methods section. Note that at this stage, only experimental Peak Δmasses should be used (Figure 1H from Materials &Methods), which should be properly specified in the PDMTableMaker_Conditions (Supplementary Table 7, Figure 1G from Materials &Methods). If the peak Δmass was not found in this Primary list, it was checked in the Secondary list coming from Unimod (See Supplementary List 2 and Figure 1G from Materials &Methods). Thus, if for a given sequence the peak Δmass was situated on an allowed residue, it was kept there and the sequence was passed unchanged to a new "SiteSequence" column. However, if it was found on a forbidden residue, it was moved towards the closest allowed residue for that peak Δmass and in the "SiteSequence" column the sequence appeared with the Δmass in the new position (Figure 1D from Materials &Methods). Also, to allow the user to review this delicate process, the modified residue change was annotated in the "SiteCorrection" column together with the corresponding list from which the allowed site for that Δmass was selected. In addition, the calculated Δmass error (ppm) was reflected in the "SiteDMError_ppm" column. However, in many cases there are more than one allowed residue in the same sequence, and both could be true (Figure 2G). Thus, to keep both possibilities we introduced the parameter X, that limited the number of residues away from the original position that a Δmass could be moved to reach an allowed residue (Figure 2G). In such cases, when for a given peptide the Δmass is dispersed forming two distributions around the two possible allowed residues, thanks to the constrains given by the X parameter on the SiteSolver working range, the Δmass can be concentrated on the two specific sites for that peptide (Fig.2G). In addition, there are cases where in the same sequence the two allowed residues are at the same distance from the original prohibited Δmass site. To decide to which one of them to assign the given Δmass, we checked for that sequence, which site with the given Δmass had the highest number of spectra and that information was given by the PDMTableMaker program (Figure 1E from Materials &Methods). Finally, if the peak Δmass was found on a forbidden residue and there were no allowed residues in the peptide sequence fulfilling the X parameter working range, the pdm was passed to the "SiteSequence" column unchanged and the "SiteDMError_ppm" column was left blank.

To test the efficiency of SiteSolver we represented the percentage of relocations in terms of the number of residues apart from the original Δmass site (Figures 2H and 2I). Figure 2H was created setting the program parameter X equal to fourteen, so that the modification cannot be relocated more than fourteen residues away from the original site. As a result, and taking all peak Δmasses together (Figure 2H), the highest percentage of relocations takes place in the amino acid contiguous to the original site and so progressively until only the minority are repositioned further away. As a matter of fact, this is corroborated for

important modifications such as deamidation, phosphorylation and oxidation (Figure 2 I), for which corrections are made more frequently in the closest residues. The more SiteSolver moves away from the original site, the lower is the number of peptides that are repositioned, being the percentage difference between “x” equal 1 an “x” equal two less accentuated in phosphorylation. On the bases of these results, the analysis from there on were performed setting the parameter X equal to three. (Suppementary Table 8)

Sticker

The PTM analysis, after running DM0Solver, TrunkSolver and SiteSolver, results in a long list of Δmasses which are usually interpreted by databases such as Unimod. However, this way of analyzing has its drawbacks. For instance, Unimod neither contains frequent Δmass combinations, for example oxidation plus deamidation (15.994915 + 0.984016 = 16.978931 Da), nor Δmasses resulting from incompatible chemical reactions with the fixed modifications set during the search: for example, a pre-existing acetylation would hinder TMT labeling, resulting in a Δmass not found in Unimod (42.010565 - 229.162932 = -187.152367 Da). Moreover, some Δmasses with biological relevance, previously detected in the laboratory, are not in this database [2]. In parallel**,** it must be considered, that there are isobaric Δmasses with radically different chemical nature depending on the residue they are located on. For instance, 15.994915 Da can be oxidation or amino acid substitution (A->S), while 0.984016 Da can be deamidation or amino acid substitution (Asn->Asp or Gln->Glu). Finally, Unimod actions are only available for certain organisms.

Hence, for a given experimental Δmass, Sticker annotations are not only based on the relative mass error threshold not being exceeded (in the case of this work 10 ppms), but also on the residue being modified, and if it appears for that Δmass in any of the provided lists. In this way, from the lists based on artifacts, Unimod and the user curated database, Sticker labelled the 84,59% of the modifications. (Supplementary Tables 10-12 and Supplemnetary Lists 3-5)

Quantitation and statistical analysis workflow

We previously developed an algorithm based on an accurate and robust statistical model (WSPP) [27] that used an automated workflow for the simultaneous PTM, protein and systems biology quantitative statistical analysis [2,16,38]. The algorithm executed a peptide-to-protein integration to quantify protein values and then computed the standardized log2-ratio of the modified peptides with respect to these protein values (*Zpq*). That allowed the detection of modified peptides with a behavior deviating more than expected (FDRpq<5%) from the rest of the peptides from the same protein, independent of the changes in protein abundance. However, at that time there were no tools to classify PTM into categories, which hindered a system-wide analysis of PTMs. Here, in the output from PDMTableMaker, all *pdms* are presented in a handy way with plenty information in separate columns to easily classify *pdm* according to, for instance, the type of residue being modified, the Δmass, the clean peptide sequence, among other parameters (Table 5 from Materials & Methods). However, in this work we focused on the *pdm* classification

according to the protein site being modified by a given Δmass (*qdna*) or by groups of Δmasses (*qna*), since this kind of PTM grouping could automatically reveal the specific protein sites involved in a given biological model with a single PTM or a whole modifications pathway.

Biological results

The performance of the developed tools is presented by a comprehensive characterization of PTMs induced by high-fat diet in plasma from an atherosclerosis mouse model.

The comparative *pdm* profile revealed clear differences between atherosclerotic and healthy mice. The increased *pdms* were specifically concentrated in APOE, ALBU, APOB, APOA2, HPT, APOA1, D3Z5G7, A1AT5 and MUG1 (Figure 3A). APOE was the protein that accumulated most of the increases (50.78%), mainly due to Met and Pro monoxidation (60.18%), but also to dioxidation (16.66%) at unknown residues, replacement of 2 protons by calcium (13.31%) and Trp oxidation to dihydroxy-N-fomaylkynurenine (9.90%). albumin from atherosclerotic mice presented more DTT-resistant Cys oxidations in the form of sulfinic and sulfonic acids. Finally, APOA1, APOB and HPT had higher Trp oxidation to hydroxykynurenine or to dihydroxy-N-fomaylkynurenine, while APOA2 and MUG1 showed higher Phe and Met monooxidations.

At this point, some decreases have also been identified in ALBU, A1AT2 and APOA1, the three of them accumulating 59.74% of the decreases (Figure 3B). ALBU mainly with di/tri/ and tetraoxidations at unknown amino acids, and sulfonic and sulfinic Cys, among others; AIAT2 with Asn deamidation and di/trioxidations, and APOA1 with di/trioxidations, Trp oxidations to kynurenine, hydroxykynurenin and dihydroxy-N-fomaylkynurenine and Met oxidation, among others.

On the other hand, the analysis at specific modified protein sites (*qna*) confirmed the protagonist role of ApoE in plasma from atherosclerotic mice, although other proteins, such as ALBU, HPT, Q58EV2, FETUA, A1AG1, APOA4 and SPA3K were also involved (Figure 4A). APOE had higher modification of Met118, mostly in the form of oxidation to aspartic semialdehyde (-32.00Da) and monoxidation (15.99Da); of Pro94 in the form of oxidation to pyroglutamic acid (13.97 Da and its C13 isotope); and of Trp274 in the form of oxidation to hydroxyformylkynurenine (C13 isotopomer with 48.98Da) and dihydroxytryptophan (31.98Da).

Moreover, figure 4B shows decreases mainly in ALBU Tyr419, concentrating a group of three modifications, such as trioxidation (49.99Da), phosphorylation (2C13 isotopomer with 81.98Da) and at a much lower extent, TMT (229.16Da). In addition, Ala146 decrease is related to several Δmasses: -15.02 Da, 46.97 Da, 245.16 Da, 58.01 Da, 30.98 Da, 32.99 Da, 48.99 Da, 64.98 Da, -31.01Da, -16.02 Da, being the largest decrease in -16.02 Da and the smallest in -15.02 Da, both unknown.

DISCUSSION

The aim of the bioinformatics tools developed in this work was to address some of the remaining challenges in high-throughput identification and quantification of PTM using the open-search approach.

The first challenge was to reduce the misassignment as modified peptides. This was accomplished by developing DM0Solver and TrunkSolver. The first one recovers non-modified peptides from Δmass peaks produced by common MS-associated artefacts such as isotopic errors, in-source formation of labile adducts with metals, or water and ammonia losses/gains, among others. TrunkSolver, on the other hand, dealed with unaccounted semi- and non-tryptic cleavages, which were artifactually detected as false PTM. As a result, the Δmass histogram was greatly simplified, especially in the far-negative Δmass region and an important proportion of non-modified peptides was recovered (Figure 2F). TrunkSolver operation is similar to Crystal-C [39], a computational tool for refinement of open search results. Although the differences have to be studied in more depth, Crystal-C does not have a variable similar to the "x" parameter of TrunkSolver. As explained in Materials and Methods, "x" defines the number of allowed positions to the right and to the left from the original Δmass site, extending, in some cases, the peptide sequence beyond its termini and into the fasta protein sequence. This is an important factor for a better control of the program behavior. Due to the 500 Da constrains on the precursor mass during the CometPTM search it is pointless to allow TrunkSolver to inspect positions beyond that range established by the user, in the configuration file (X parameter, Supplementary Table 5), from the original Δmass site. Moreover, the lack of control of the "x" parameter not only increases the execution time but is also more error-prone.

The second challenge was to correct the localization of the modified residue to the most probable site when the information in the MSMS spectrum was not enough to accurately pinpoint the site. This was addressed by the combination of the programs PDMTableMaker and SiteSolver. Here is important to highlight that SiteSolver gives preference to the list of allowed residues in the user-curated Δmass file over the ones found in Unimod. In this way the user-accumulated experience in specific modifications prevails over the more general Unimod data base information. The scan frequency of each pdm calculated by PDMTableMaker is used by SiteSolver in cases where in the same sequence are two allowed residues at the same distance from the original prohibited Δmass site. Most Δmass relocations performed by SiteSolver took place in the residues closest to the original modified site. Besides, the Δmass dispersion along the peptide sequence was greatly reduced and got concentrated on the most probable site(s) (Figure 2G). These findings showed that SiteSolver was effective to resolve the site allocation problem (Figure 2H and 2I).

The third challenge was to enable quantitative ontology-based PTM analysis. This problem was approached by using PDMTableMaker and the new integration steps from our

quantitative workflow based on the WSPP [27] statistical model. PDMTableMaker allowed the integration of *pdm*s into upper level elements attending to the Δmass (*d*), the protein site being modified by a given Δmass (*qdna*), including the different peptide forms that contain the modification, or the modified residue (*qna*), including all the modifications that affect it. These *pdm* groupings were used as relation tables for new integrative steps in the workflow that concentrated the quantitative information, avoiding the "dilution" problem and increasing statistical significance, and revealed specific protein sites with a single PTM or whole modification pathways associated with atherosclerosis in the mouse model used in this work.

All the tools developed in this work were set into a modular pipeline schematized in Figure 1H from Materials and Methods. We should note that the tools are independent and flexible, obviating the need to follow a preestablished order. The user can decide to run all the modules in a given order or just part of them, depending on the objective to be achieved. The modules can also be run several times in the same workflow if needed. Also, by adjusting the parameters in the configuration files (Supplementary Tables 1-12) and lists (Supplementary Lists 1-5) the user has full control of the program's execution.

Taking advantage of the great versatility of the modular design, we decided to develop a workflow that would allow us to optimally analyze the PTM data. Thus, we decided to execute DM0Solver before TrunkSolver for the sake of time. TrunkSolver takes longer than DM0Solver, since it has to make all the combinations between the different cuts of sequences and the Δmasses established by the user (Figure1 B, Supplementary Table 5). Then, peptides assigned as unmodified by DM0Solver would not have to be reanalyzed by TrunkSolver. In this way, the workflow would start with DM0Solver followed by TrunkSolver and, as it has been previously explained, the latter must be accompanied by PeakAssignator to re-assign the newly generated Δmasses to the corresponding Δmass peaks.

On the other hand, we decided to run the combination of PDMTableMaker&SiteSolver after the combination of TrunkSolver&DM0Solver, and not before, also for the sake of time. First, any reassignments made by the latter two, as the Δmasses are extracted out of sequence, would not have to be examined by SiteSolver. Thus, once PDMTableMaker has been passed, SiteSolver can be executed, and to finish, with the aim of keeping SiteSolver corrections on the peak Δmasses together with orphan Δmasses in a final compact *pdm* table, PDMTableMaker has to be run once again.

Sticker can be used as a final module, with the advantage that it can be run on the PDMTable itself, on the SiteSolver output file, as well as on any file containing as a minimum requirement the 2 columns (Δmass, residue). Moreover, Sticker was also made flexible accepting not only the classical Unimod lists of Δmasses, but also user-curated lists from biological Δmasses not found in Unimod, like newly discovered PTM, or artefacts derived from a specific chemistry.

The tools developed in this work would contribute to understand the interplay between PTMs and protein function in the context of disease. To demonstrate it, we applied our tools to perform a comprehensive unbiased study of the changes of plasma modified peptidome in a atherosclerosis mouse model. This is a pathology underlying many other cardiovascular diseases which are the main cause of death worldwide [17]. The role of PTMs in this disease has never been studied from a truly hypothesis-free approach, which may be indispensable for advancing in atherosclerosis research.

Our results demonstrate that the development of atherosclerosis brought about by high-fat diet in LDLR KO mice is accompanied by altered PTM profiles, especially on ApoE and ALBU. Note that in this work we achieved a quite high sequence coverage for both proteins: 78% for ApoE and 90% for Albu, which allowed the construction of accurate PTM maps of both proteins. Thus, the induction of atherosclerosis brings about a significant increase in plasmatic ApoE oxidation pathways specifically on M118 and W274. Albu, on the other hand, is more irreversibly oxidized on its C591, while the rest of di-tri- and tetra oxidations on difficult to assign residues are reduced.

Apolipoprotein E is a structural component of all lipoproteins except for LDL (low density lipoprotein) and is a critical ligand for the hepatic clearance of plasma lipoproteins mediated by LDL and LDLR (low density lipoprotein receptor) [40]. ApoE has been found to provide extra atheroprotective properties since it plays a requisite role in remnant lipoprotein clearance by the liver, and although hepatic LDL receptors can clear both LDL and ApoE-containing lipoproteins, LDL receptor-related protein mediated clearance of remnants is dependent on ApoE [41]. Concerning ApoE M118, although one of the PTMs taking place on this residue is also a common MS-associated artefact (Met oxidation), it co-occurs with an extreme oxidation to aspartic semialdehyde on the same residue. Moreover, ApoE has 13 additional Met residues, from which we detected 7. One of those (M135) has a known sulfoxidation site differentiating C57BL/6 and BALB/cmouse strains [42], which we find decreased. However, from the 6 remaining Met residues, only oxidative modifications in M118 were found increased. Altogether, our results suggest that the oxidation pathway on M118 from ApoE is truly associated with atherosclerosis.

There are no functions reported for M118 in ApoE. In the primary ApoE sequence the closest known modified sites are quite far from Met118 (K105 ubiquitination and T134 phosphorylation), and the probability of interference among the three of them seems quite low. However, in the 3D structure M118 is situated in an internal position from a α-helical region facing the other 3 helixes composing ApoE structure (Figure 5). Moreover, these α-helical regions contain tandem repeats conserved among ApoA1, ApoA4 and ApoE, induce the helical structure of these proteins and afford them the ability to bind lipids [43]. In ApoE this region goes from residue 80 until 255 [43] and we find altered other residues falling in this region, such as the increasing P94, the already mentioned M135, the decreasing E87.

P94 oxidation to pyroglutamic acid, a ROS-induced extreme oxidation of Pro [44], is increased in atherosclerotic mice compared to the control ones. Moreover, out of 9 ApoE P residues, we detect 5, but only P94 is changing, reinforcing its biological nature.

Contrary to M118, M135 is situated on an external loop linking two α-helical regions (Figure 5), which could make this Met more oxidation-prone *in vivo*, compared to the rest of ApoE methionines. However, this M is less oxidized in atherosclerotic mice, suggesting there might be other mechanisms involved regulating M135 oxidation. Thus, we speculate that all alterations we observe in this conserved α-helical tandem repeat could affect ApoE assembly into its helical structure and change the protein ability to bind lipids.

On the other hand, we also find increased the oxidation pathway on ApoE W274 in the form of di and trioxidation. This seems a biologically significant finding since out of the 7 ApoE W residues, we detect 4 and W274 is the only one changing. Near W274 lies the residue E276, which is located on a region concentrating di- and trioxidations, most probably situated on nearby residues. Contrary to W274, the rest of the oxidations in this protein region are reduced in atherosclerotic mice. W274 and E276 are outside of the previously mentioned α-helical tandem repeats. However, they are found in a region responsible for the specific association with VLDL (very low density lipoportein). With this regard, W oxidations have already been related to atherosclerosis, being the kynurenine pathway associated with an increased probability of developing symptomatic unstable atherosclerotic disease [45,46]. Thus, we are tempted to speculate that alteration of this region my change the ApoE-VLDL union, and may alter the transport of lipids synthesized in the liver inside the tissues, as well as the lipase hydrolyzing activity, thus giving rise to hypertriglyceridemia and therefore atherosclerosis [47].

Moreover, oxidations located near ApoE L303 and P305 are decreased. These residues belong to an homooligomerization region and their modifications could alter ApoE 3D assembly. It has been studied that protein oligomerization generate non-functional proteins and form fiber deposits, leading to a large number of degenerative diseases such as Alzheimer's, Creutzfeld -Jacob, Parkin- son, Huntington, systemic amyloidosis, type II diabetes, amyotrophic lateral sclerosis and cerebral angiopathy, among others [48], but it has not been observed in ApoE.

In addition, plasmatic ApoE from atherosclerotic mice is more associated with calcium, than in control animals. This is an interesting finding, since calcium gives rise to calcification, contributing to the formation of atherosclerotic plaque [49,50], although our results are limited to take conclusions in this point.

Albumin, on the other hand, is the most abundant plasma protein with pleiotropic functions, ranging from regulation of plasma osmotic pressure, binding and transport of various endogenous or exogenous compounds, and finally to extracellular anti- oxidant defenses [51]. Lower serum albumin levels are associated with increased risk of cardiovascular mortality, coronary artery disease and stroke suggesting protective effects

of albumin against atherosclerosis [52]. This phenomenon could be related to the antioxidant properties of albumin and to its calcium transporter role [52,53]. Albumin is an helicoidal protein, consisting of three structurally similar domains I, II, and III, each of them formed by two subdomains (A and B) [54]. Domain II has been confirmed as the most rigid one [53].

Atherosclerotic animals show a striking decrease in the overall albumin oxidation (di/tri and tetraoxidations), which is concentrated mainly on domain I. Of note are C86 and C148, which are part of disulfide bridges. The bridge formed between C86 and C77 involves a region containing the reduced sites V78, A83, A84 and N85, while C86 is less sulfonated in diseased animals. As for C148, it seems to be in a region which is in general less oxidized in atherosclerotic animals. Contrasting with these two Cysteines is C591, located on domain III as part of a disulfide bond, and whose sulfonated and sulfinated forms are increased in the diseased animals compared to controls. The rest of the 31 Cys detected in albumin in different modified or non-modified forms, out of the 37 total Cys contained in the protein sequence, were not altered. Thus, all the observed PTM changes could be related with some albumin function acting as a redox defense mechanism [51] (Figure 6).

In conclusion, the physiological-mechanistic relevance of the obtained results cannot be determined yet since they are novel. There are still many obstacles to overcome in the biological field of the disease, as well as in the technological one regarding post-translational modifications. The detected PTMs could be markers that reflect the atherosclerosis, or effectors that are involved in the pathogenesis. Consequently, the need to continue our study is evident, not only in the technological part but also in the biological one, applying our tools to ongoing projects in atherosclerotic models such as the ApoE KO mice (David Sancho laboratory), minipigs (Jacob Fog Bentzon) and clinically-relevant projects in humans such as PESA (Progression of Early Subclinical Atherosclerosis) and AWHS (Aragon Workers Health Study).

CONCLUSIONS

The main conclusions reached in this study were the following:

1. DM0Solver and TrunkSolver produced an important simplification of the Δmass histogram, especially in the far negative Δmass region, almost doubling the number of PSMs assigned to unmodified peptides.
2. SiteSolver minimized the dispersion of modified residues, concentrating the quantitative information on the most probable site(s)
3. PDMTableMaker and Sticker on one side, and the new WSPP integrative algorithm on the other, enabled a quantitative ontology-based PTM analysis, detecting alterations of specific protein sites affected by single PTMs or even by whole PTM pathways.
4. The development of atherosclerosis brought about by high-fat diet in LDLR KO mice is accompanied by altered PTM profiles mainly on ApoE and ALBU, among other proteins. There is a significant increase in plasmatic ApoE oxidation pathways specifically on M118 and W274; Albu is more sulfonated and sulfinated on its disulfide bond-forming C591 from domain III, while domain I is in general less oxidized.

Figures

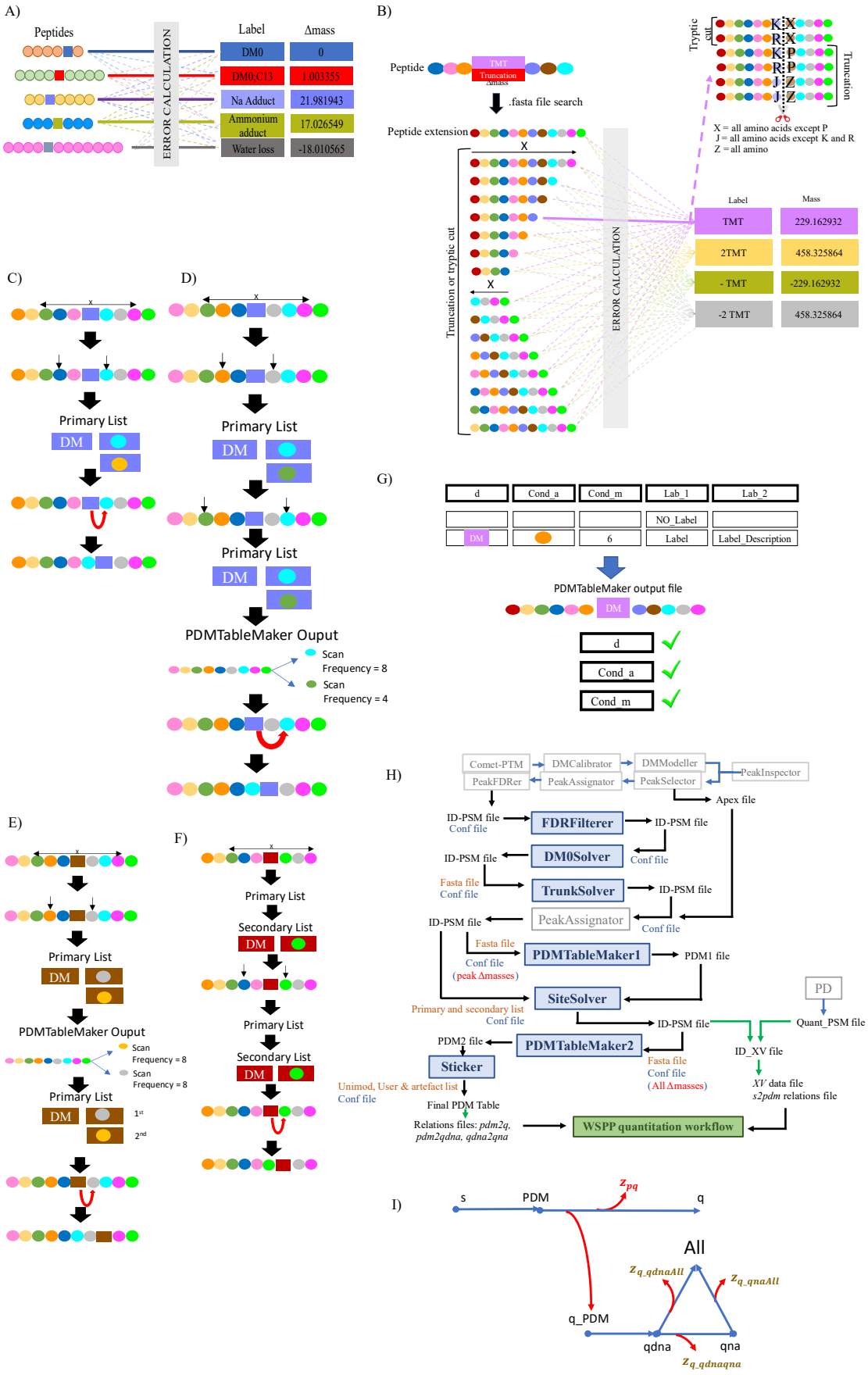

A)
Peptides
ERROR CALCULATION
Label
Δmass
DM0
0
DM0;C13
1.003355
Na Adduct
21.981943
Ammonium adduct
17.026549
Water loss
-18.010565
B)
Peptide
TMT
Truncation
Δmass
.fasta file search
Peptide extension
X
Truncation or tryptic cut
ERROR CALCULATION
Tryptic cut
Truncation
X = all amino acids except P
J = all amino acids except K and R
Z = all amino
Label
Mass
TMT
229.162932
2TMT
458.325864
- TMT
-229.162932
-2 TMT
458.325864
C)
Primary List
DM
D)
Primary List
DM
Primary List
DM
PDMTableMaker Ouput
Scan
Frequency = 8
Scan
Frequency = 4
E)
Primary List
DM
PDMTableMaker Ouput
Scan
Frequency = 8
Scan
Frequency = 8
Primary List
DM
1st
2nd
F)
Primary List
Secondary List
DM
Primary List
Secondary List
DM
G)
d
Cond_a
Cond_m
Lab_1
Lab_2
NO_Label
DM
6
Label
Label_Description
PDMTableMaker output file
DM
d
Cond_a
Cond_m
H)
Comet-PTM
PeakFDRer
DMCalibrator
PeakAssignator
DMModeller
PeakSelector
PeakInspector
Apex file
ID-PSM file
Conf file
FDRFilterer
ID-PSM file
DM0Solver
Conf file
Fasta file
Conf file
TrunkSolver
ID-PSM file
ID-PSM file
PeakAssignator
Conf file
Fasta file
Conf file
(peak Δmasses)
PDMTableMaker1
PDM1 file
SiteSolver
Primary and secondary list
Conf file
PD
ID-PSM file
Quant_PSM file
PDM2 file
PDMTableMaker2
ID_XV file
Sticker
Fasta file
Conf file
(All Δmasses)
XV data file
s2pdm relations file
Unimod, User & artefact list
Conf file
Final PDM Table
Relations files: pdm2q, pdm2qdna, qdna2qna
WSPP quantitation workflow
I)
s
PDM
q
All
q_PDM
qdna
qna

Figure 1. **DM0Solver, TrunkSolver, SiteSolver, Sticker operation and scheme of the integrative workflow used to quantify modified peptides.**

A) **DM0Solver Operation.** In the left part of the figure, a scheme of the peptides analyzed by DM0Solver is shown. Each peptide has a Δmass (square). DM0Solver, calculates the relative error with each of the masses of the labels introduced in DM0Solver configuration file (right part of the figure). DM0Solver saves the lowest error produced and if it is lower than the Relative_Error_ppm parameter, the error is noted at the output.

B) **TrunkSolver Operation**. At the top of the figure the peptide is represented with its Δmass. This peptide sequence is elongated from the modified residue onwards to the right and to the left following the protein sequence in the fasta file as many residues as it is set in the x parameter of the configuration file. In this case the Δmass corresponds to a cut plus a TMT. This cut can be a tryptic or non-tryptic cut (truncation) depending on the amino acids at the C-terminal of the cut. If the cut occurs between lysine or arginine and any other amino acid that is not proline, it is a tryptic cut. If the cut occurs between arginine or lysine and proline, it is a truncation, and if at the C-terminal end of the resulting peptide sequence there is no lysine or arginine, it will be a truncation. The left part of the figure schematically shows, the entire peptide and all the possible cuts analyzed by TrunkSolver. TrunkSolver, calculates the relative error with each of the masses of the labels introduced in the TrunkSolver configuration file (right part of the figure). TrunkSolver saves the lowest error produced, which, in this case, corresponds to a truncation plus a TMT. If the error is lower than the Relative_Error_ppm parameter, it is noted in the output file.

C) **Example of SiteSolver operation.** A PDM is analyzed to detect if its Δmass is at an incorrect position. SiteSolver will consider as many positions as it is set in the configuration file. In a first step, SiteSolver checks if the amino acid, to which the Δmass is originally assigned (pink circle), is forbidden or allowed by looking in the Primary list. As it does not appear (prohibited), it evaluates the contiguous amino acids (dark blue and light blue circles). SiteSolver concludes that one of the two is allowed (light blue circle) in the Primary List. Hence, the sequence is reassigned.

D) **SiteSolver operation instance.** This PDM analysis follows the same schema. In a first step, SiteSolver determines if the amino acid, to which the Δmass is originally assigned (dark blue circle), is forbidden or allowed, by looking in the Primary list. As it does not appear (prohibited) it evaluates the immediate amino acids (orange, gray circles). The amino acids are not allowed in the Primary list. Accordingly, SiteSolver examines to the following immediate amino acids (light blue and green). At this point a tie can be seen since both are allowed. In these cases, SiteSolver assigns the Δmass to the aminoacid with the highest scan frequency which are obtained from PDMTableMaker output.

E) **Example of SiteSolver operation.** After verifying that the Δmass was assigned, in its original form, to a forbidden amino acid for that Δmass (dark blue circle), Site Solver tests out the immediate amino acids (gray and orange). Considering that both are allowed and have the same scan frequency the order in which they appear in the Primary list is decisive. It is assigned to the amino acid that appears first (gray circle).

F) **Example of SiteSolver operation**. This example is similar to D), with the difference that the Δmass is not in the Primary list so the SecondaryList is observed. This list shows how the Δmass was assigned, in its original form, to a prohibited amino acid for that Δmass (prink circle). Site Solver evaluates the contiguous amino acids (blue and lemon green circle). The Δmass is reassigned to the only allowed amino acid (lemon green circle).

G) **Sticker operation Instance.** An example of a DMID_list.txt is outlined at the top of the figure. The first column contains the Δmass, followed by two terms "a" and "m", which are existing fields in the PDMTableMaker output file. The fourth and fifth columns are the labels that are assigned. The first row of DMID_list.txt always specifies which label must be annotated when the conditions are not met, in this case "No_Label". For each row of the PDMTableMaker output file, if DMID_list.txt terms are met it will be noted in the output file.

H) **Workflow scheme**. Dataset was searched with Comet-PTM followed by DMCalibrator, DMModeller , PeakInspector, PeakSelector, PeakAssignator, PeakFDRer (Andrea Laguillo ongoing PhD Thesis). PeakInspector (Rafael Barrero ongoing PhD Thesis) allows the inspection of the DM histogram to set PeakSelector parameters. PeakFDR's output was filtered using FDRFilterer. The filtered ID_PSM table was passed to DM0Solver. Then TrunkSolver was run with the fasta file along with the DM0Solver outout. To re-adjust the DM peaks due to the action of TrunkSolver, PeakAssignator was used again with the apex list from PeakSelector and TrunkSolver output files. With the PeakAssignator output file, together with the fasta file and considering only peaks, PDMTableMaker, gave rise to the PDMTable. SS took the ID_PSM PeakAssignator output file, together with the PDMTable and which served SiteSolver, together with the primary and the secondary lists, to correct the wrongly assigned modified sites in the ID_PSMTable. To save the peak Δmasses data corrected by SiteSolver together with the rest of orphan Δmasses in the same PDMTable, PDMTableMaker was run again on the ID_PSM file output from SiteSolver considering all Δmasses, and give rise to a Final PDMTable, which was labeled with Sticker making use of the artifacts, Unimod and user lists (Supplementary Lists 3,4,5 respectiveluy). The ID_XV file was obtained by merging the ID_PSM file output from SiteSolver with the Quant_PSM files output from Proteome Discoverer containing the intensities of the TMT tags. From the ID_XV file the XV_data and *scan2pdm* relations files for the WSPP quantitative workflow were obtained, while the rest of the relations files (*pdm*-to-*protein*, the *pdm*-to-*protein_site_DM* (*qdna*) and the *qdna*-to-*protein_site* (*qna*)) were obtained from the FinalPDMTable. For the correct functioning of the programs, the configuration files (Supplementary tables 1-12) and certain lists (Aupplementary lists 1-5) must be provided. The previously developed programs are lined in gray, while the programs developed in the present work are in blue. In addition, the files to be provided by the user are shown in orange, the configuration files (Conf file) are in blue and the input/ output files are in black.

I) **Scheme of the integrative workflow used to quantify modified peptides.** Each arrow represents a step performed with the generic integration algorithm [27]. Standardized log2-ratio values at the PDM level z_{pq} are obtained from the PDM-to-protein integration. The algorithm provides corrected PDM values by the corresponding protein value (x_{q_PDM}), which are integrated to a category composed by a DM modifying a certain protein site (qdna), getting $z_{q\_pdm2qdna}$. Subsequently, x_{q_qdna} is integrated to a category composed by all detected DM modifying a certain protein site (qna, $z_{q\_qdna2qna}$) or to ALL ($z_{q\_qdna2ALL}$). Finally, x_{q_qna} is integrated to ALL, obtaining $z_{q\_qna2ALL}$.

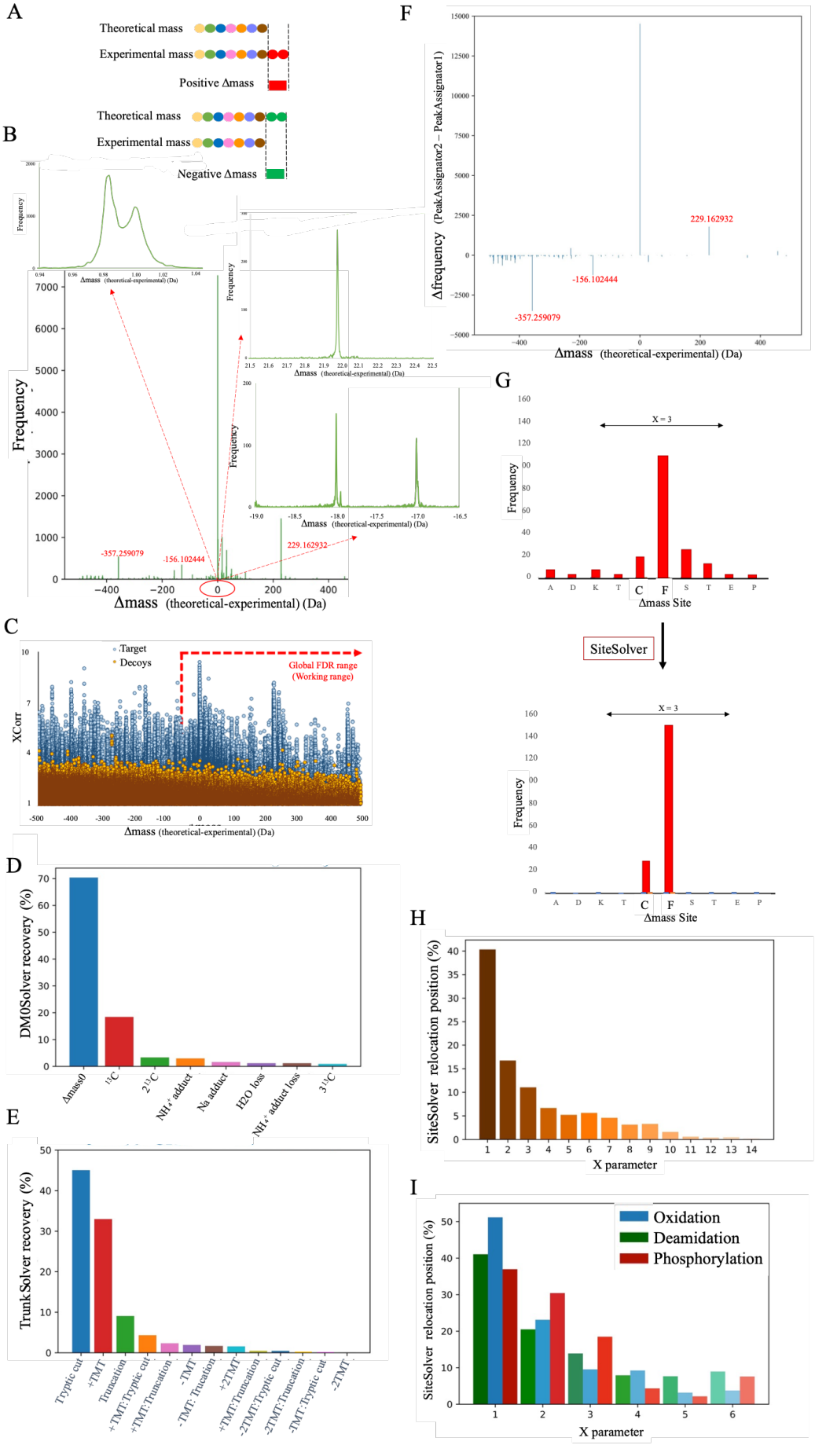

A
Theoretical mass
Experimental mass
Positive Δmass
B
Theoretical mass
Experimental mass
Negative Δmass
Frequency
Δmass (theoretical-experimental) (Da)
-357.259079
-156.102444
229.162932
C
Target
Decoys
Global FDR range
(Working range)
XCorr
Δmass (theoretical-experimental) (Da)
D
DM0Solver recovery (%)
Δmass0
13C
2 13C
NH4+ adduct
Na adduct
H2O loss
NH4+ adduct loss
3 13C
E
TrunkSolver recovery (%)
Tryptic cut
+TMT
Truncation
+TMT:Tryptic cut
+TMT:Truncation
-TMT
-TMT: Truncation
+2TMT
+TMT:Truncation
-2TMT:Tryptic cut
-2TMT:Truncation
-TMT:Tryptic cut
-2TMT
F
Δfrequency (PeakAssignator2 – PeakAssignator1)
229.162932
-156.102444
-357.259079
Δmass (theoretical-experimental) (Da)
G
X = 3
Frequency
A D K T C F S T E P
Δmass Site
SiteSolver
X = 3
Frequency
A D K T C F S T E P
Δmass Site
H
SiteSolver relocation position (%)
X parameter
I
SiteSolver relocation position (%)
Oxidation
Deamidation
Phosphorylation
X parameter

Figure 2. **DM0Solver, TrunkSolver and SiteSolver effectiveness.**

A) **Representation of Comet-PTM experimental Δmass calculation.** Δmass = theoretical mass - experimental mass. In the first example the Δmass is negative since the theoretical mass is greater than the experimental mass. In the second instance the Δmass is positive, since the experimental mass is higher than the theoretical mass.

B) **Typical CometPTM Δmass histogram after executing FDRFilterer.** The zoomed insets show some non-zero Δmass corresponding to non modified peptides.

C) **XCorr distribution for targets (blue) and decoys (yellow).in the whole Δmass range**. The graph shows an Xcorr-bias towards far-negative Δmasses for targets and decoys. This representation was made after executing the PeakFDRer.

D) **Percentage of each label assignation by DM0Solver.** Of the 36.47% of the total that DM0Solver recovers, the following percentages per label were allocated: 70.39% Δmass0; 18.44% ^{13}C; 3.33% $2^{13}C$ (two^{13}C);2.97% NH_4^+ adduct (ammonium adduct);1.60% Na adducts (sodium adducts); 1.19% H_2O loss (water loss); 1.18% NH_4^+ loss (ammonium adduct loss); 0.90% $3^{13}C$.

E) **Percentage of each label assignation by TrunkSolver.** Of the 12.31% of the total that TrunkSolver recovers, the following percentages per label were allocated: 45.06% tryptic cuts;32.99% +TMT(TMT gain);9.06% Truncation(non-tryptic cut); 4.33% +TMT:Tryptic cut (TMT gain and a tryptic cut); 2.34% +TMT:Truncation (TMTgain and a non-tryptic cut);1.94% -TMT (TMT loss);1.68% -TMT: Trucation (TMT loss and non-tryptic cut); 1.57% +2TMT (double gain of TMT); 0.49% +TMT:Truncation (non-tryptic cut and double gain of TMT); 0.46% -2TMT:Tryptic cut (double loss of TMT with tryptic cut);0.06 % -2TMT:Truncation (double loss of TMT and non-tryptic cut);0.01% -TMT:Tryptic cut (TMT loss and tryptic cut);0.01% -2TMT (double loss of TMT).

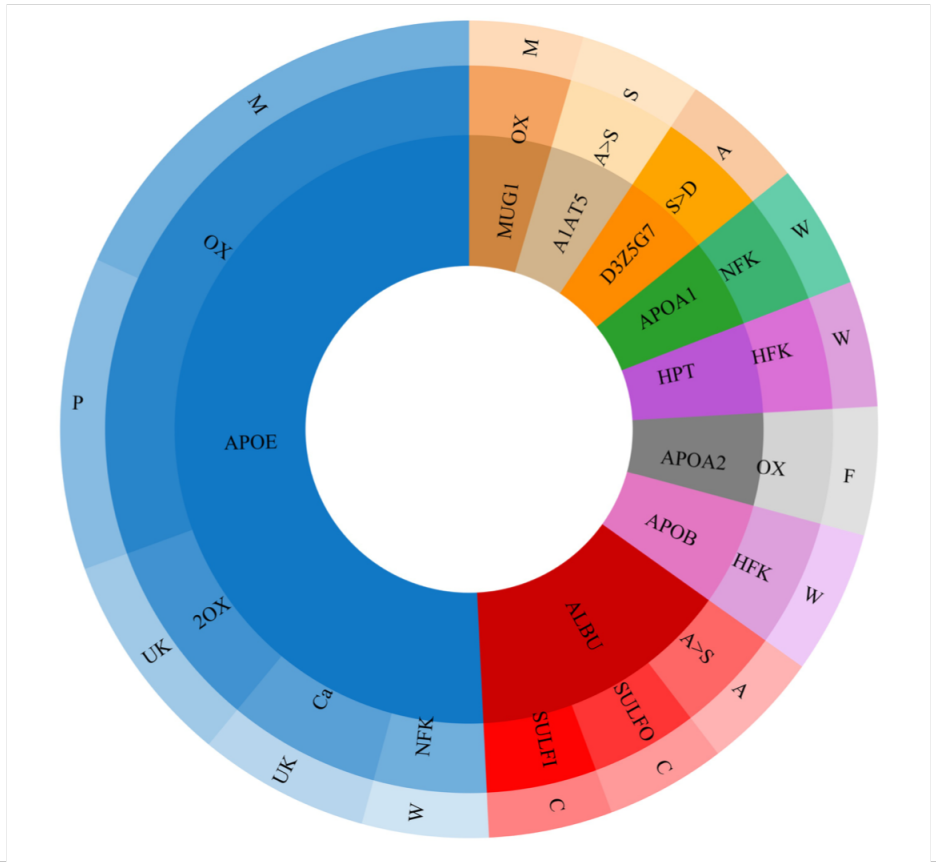
F) **Δmasses frequency variation before and after applying DM0Solver and TrunkSolver.** The difference between the frequency of the Δmasses assigned by PeakAssignator after and before passing DM0Solver and TrunkSolver is shown. The Δmasses with the highest frequency differences are highlighted, such as: -357.259079Da (extra lysine plus TMT), -156.102444 Da (extra arginine), 0Da, 229.162932 (TMT).

G) **SiteSolver efficiency.** The frequency of oxidation (15.993411 Da) distribution in the mouse ALBU peptide ADKTCFSTEP before and after passing SiteSolver is shown. The sequence contains only two allowed residues for this DM (C and F) in the user-curated list. Before SiteSolver the DM is distributed in two dispersed populations around C5 and F6, respectively. However, after passing SiteSolver with a working range of 3 in the parameter X, we managed to concentrate all the scans from this DM in this peptide on C5 and F6, demonstrating the reassignment was correct.

H) **Assignment correction percentage by SiteSolver for each position, setting the program parameter X equal to fourteen.** The x axis represents the values of the X parameter and the y axes the percentage. One is the closest position to the original Δmass site and fourteen the furthest. The percentages are reflected: X = 1:40.33%; X=2:16.75%; X=3:11.06%; X=4:6.67%; X=5:5.22%; X=6: 5.62%; X=7:4.60%; X=8:3.14%;X=9:3.30%,X=10:1.59%; X=11:0.59%; X=12: 0.40%; X=13:0.51%; X=14:0.22%.

I) **Assignment correction percentage by SiteSolver for each position, setting the program parameter X equal to six.** The x axis represents the values of the X parameter and the y axes the percentage. One is the closest position to the original Δmass site and six the furthest. The percentages are reflected for eah modification. Oxidation: X=1:51.16%; X=2:23.12%; X=3:9.54%, X=4:9.25%; X=5:3.18%; X=6:3.76%. Deamidation: X=1:41.07%; X=2:20.50%; X=3:13.9%; X=4:7.92%; X=5:7.66%; X=6:8.95%. Phosphorylation: X=1:36.96%; X=2:30.43%; X=3:18.48&; X=4:4.35%; X=5:2.17%; X=6:7.61%.

A


APOE
ALBU
MUG1
A1AT5
D3Z5G7
APOA1
HPT
APOA2
APOB
SULFI
SULFO
OX
2OX
UK
Ca
NFK
HFK
A->S
S->D
F
P
M
S
A
W
C

B

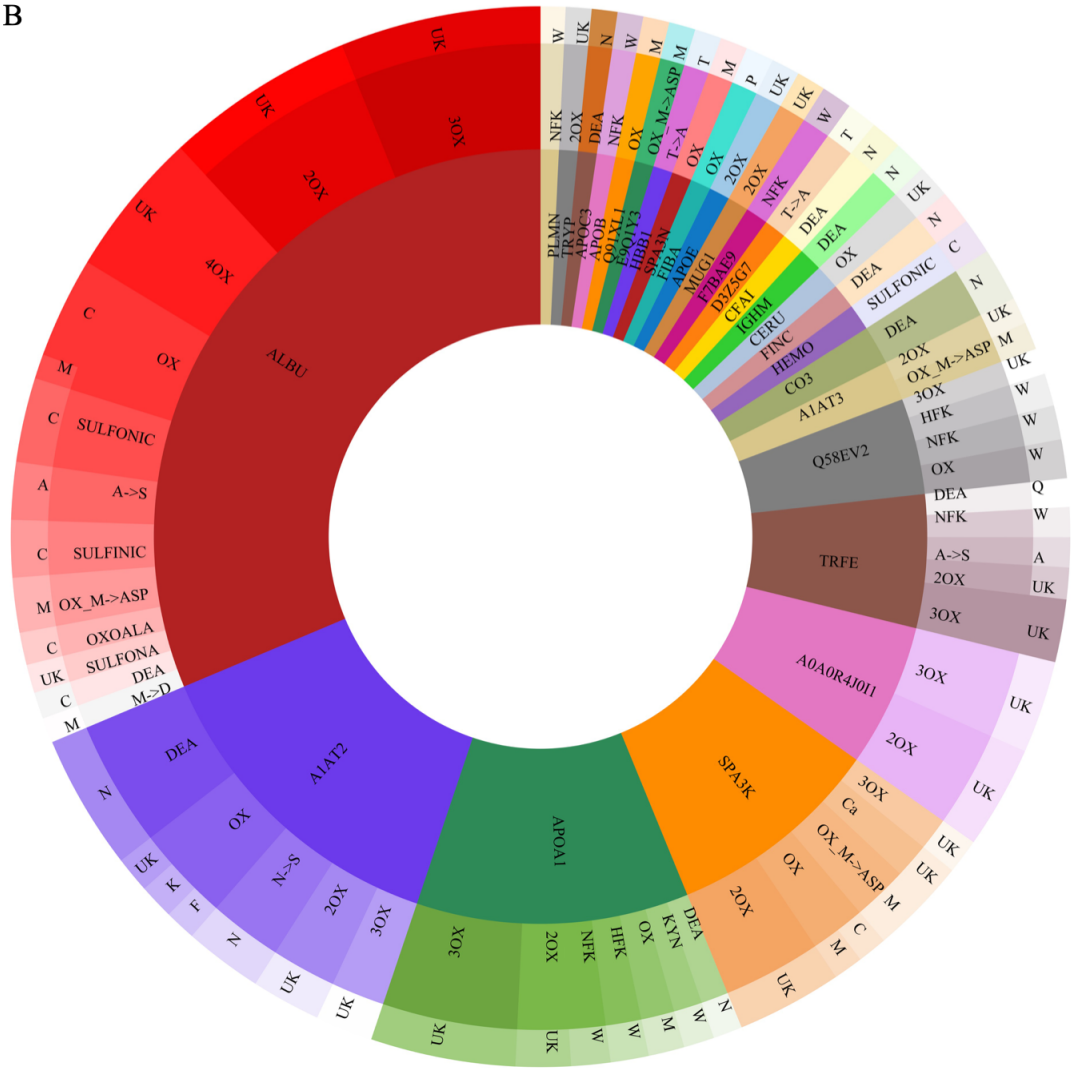
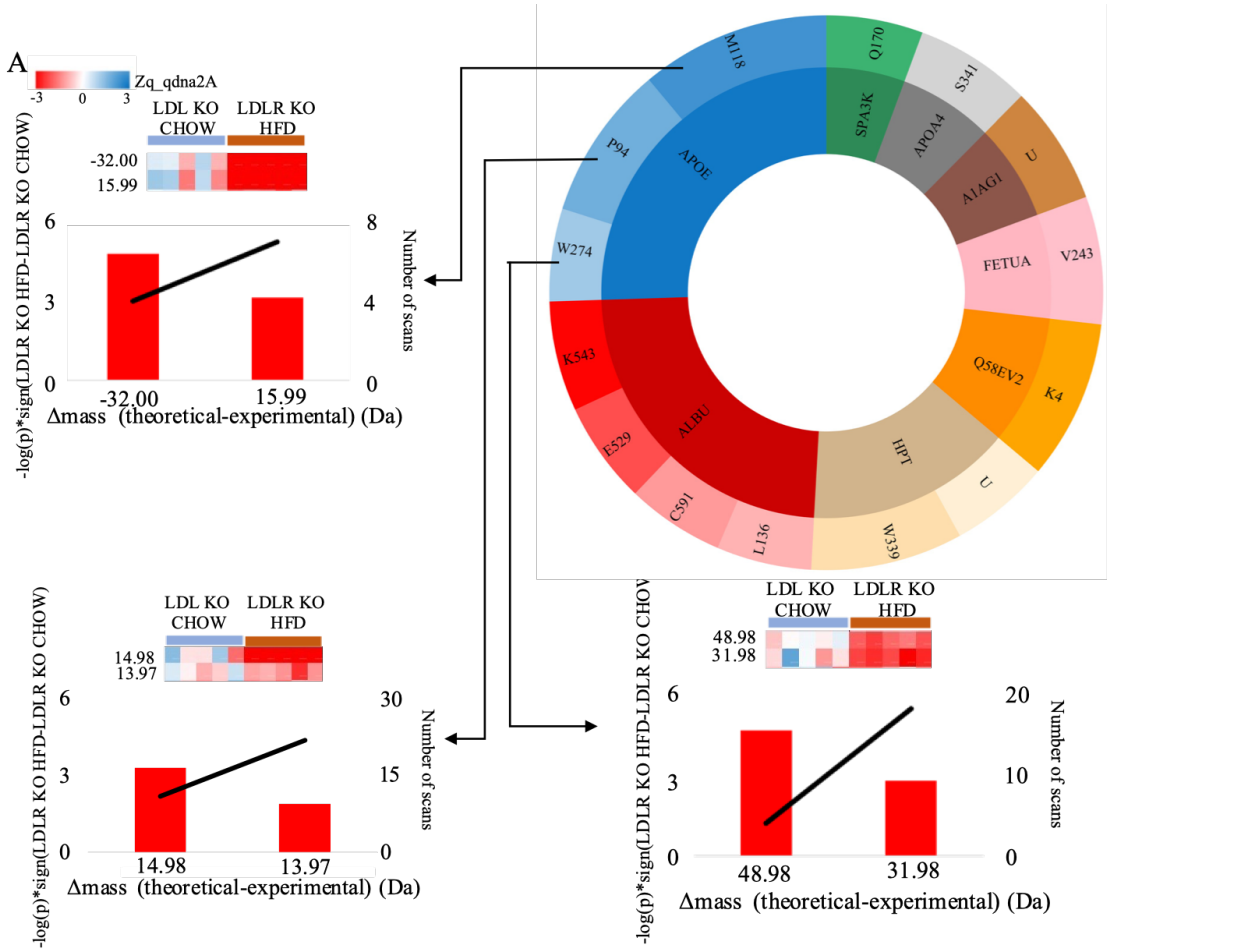

ALBU
A1AT2
APOA1
SPA3K
A0A0R4J0I1
TRFE
Q58EV2
A1AT3
CO3
HEMO
FINC
CERU
IGHM
CFAI
D3Z5G7
F7BAE9
APOE
AMUG1
FIBA
SPA3N
HBB1
Q9UXL1
Q9UIY3
APOC3
APOB
TRYP
PLMN
3OX
2OX
4OX
OX
SULFONIC
A->S
SULFINIC
OX_M->ASP
OXOALA
SULFONA
M->D
DEA
N->S
NFK
HFK
KYN
Ca
T->A
OX_M->ASP

Figure 3. **High-fat diet-associated PTM changes in plasma from LDLRKO mice.**

Nested pie charts show the *pdms* (with more than 5 quantified spectra) with statistically significant (n=5, ttest, *FDR*<=0.15) increases (A) and decreases (B). The graphs are proportional to the magnitude of the change between the control (LDLR KO CHOW) and the LDLR KO HFD samples, expressed in terms of -log(p-value)*sign($z_{pdm2q\_chow} - z_{pdm2q\_HFD}$). The adscription of the type of modification and the modified residue to the plasma proteins is depicted, and the magnitude of the *pdms* composing each one of these three graph layers was correspondingly summed. Ox, 2Ox, 3Ox and 4Ox: addition of one, two, three and four oxygens, respectively; Ca: replacement of 2 protons by calcium; HFK: hydroxykynurenine; NFK: dihydroxy-N-fomaylkynurenine; KYN: Kynurenine; SULFI: cysteine sulfinic acid; SULFO: cysteine sulfonic acid; A>S: alanine by serine amino acid substitution; S>D: serine for aspartic acid substitution; A>S: alanine for serine substitution; OX_M->ASP: methionine oxidation to aspartic semialdehyde; OXOALA: cysteine oxidation to oxoalanine; SULFONA: sulfonation; DEA: deamidation; M<D: methionine for aspartic acid substitution; N<S: asparagine for serine substitution; T<A: threonine alanine substitution; UK: unknown; ALBU: Serum albumin; APOA1: Apolipoprotein A-I; APOA2: Apolipoprotein A-II; APOB: Apolipoprotein B-100; APOC3: Apolipoprotein C3; APOE: Apolipoprotein E; HPT: Haptoglobin; D3Z5G7: Carboxylic ester hydrolase; A1AT2: Alpha-1-antitrypsin 1-2; A1AT3: Alpha-1-antitrypsin 1-3; A1AT5: Alpha-1-antitrypsin 1-5; MUG1: Murinoglobulin-1; SPA3K (Serine protease inhibitor A3K); A0A0R4J0I1: Serpina3k; TRFE: Serotransferrin; CO3: Complement component 3; HEMO: Hemopexin; FINC: Fibronectin; CERU: Ceruloplasmin; IGHM: Immunoglobulin heavy constant mu; CFAI: Complement factor I; FIBA: Fibrinogen alpha chain; SPA3N: Serine protease inhibitor A3N; HBB1: Hemoglobin subunit beta-1; Q91XL1: Leucine-rich HEV glycoprotein; PLMN: Plasminogen.


A
Zq_qdna2A
-3 0 3
LDL KO CHOW
LDLR KO HFD
-32.00
15.99
-log(p)*sign(LDLR KO HFD-LDLR KO CHOW)
Number of scans
Δmass (theoretical-experimental) (Da)
-32.00
15.99
14.98
13.97
48.98
31.98
M118
Q170
S341
P94
APOE
SPA3K
APOA4
U
A1AG1
W274
FETUA
V243
K543
ALBU
Q58EV2
K4
E529
HPT
U
C591
L136
W339

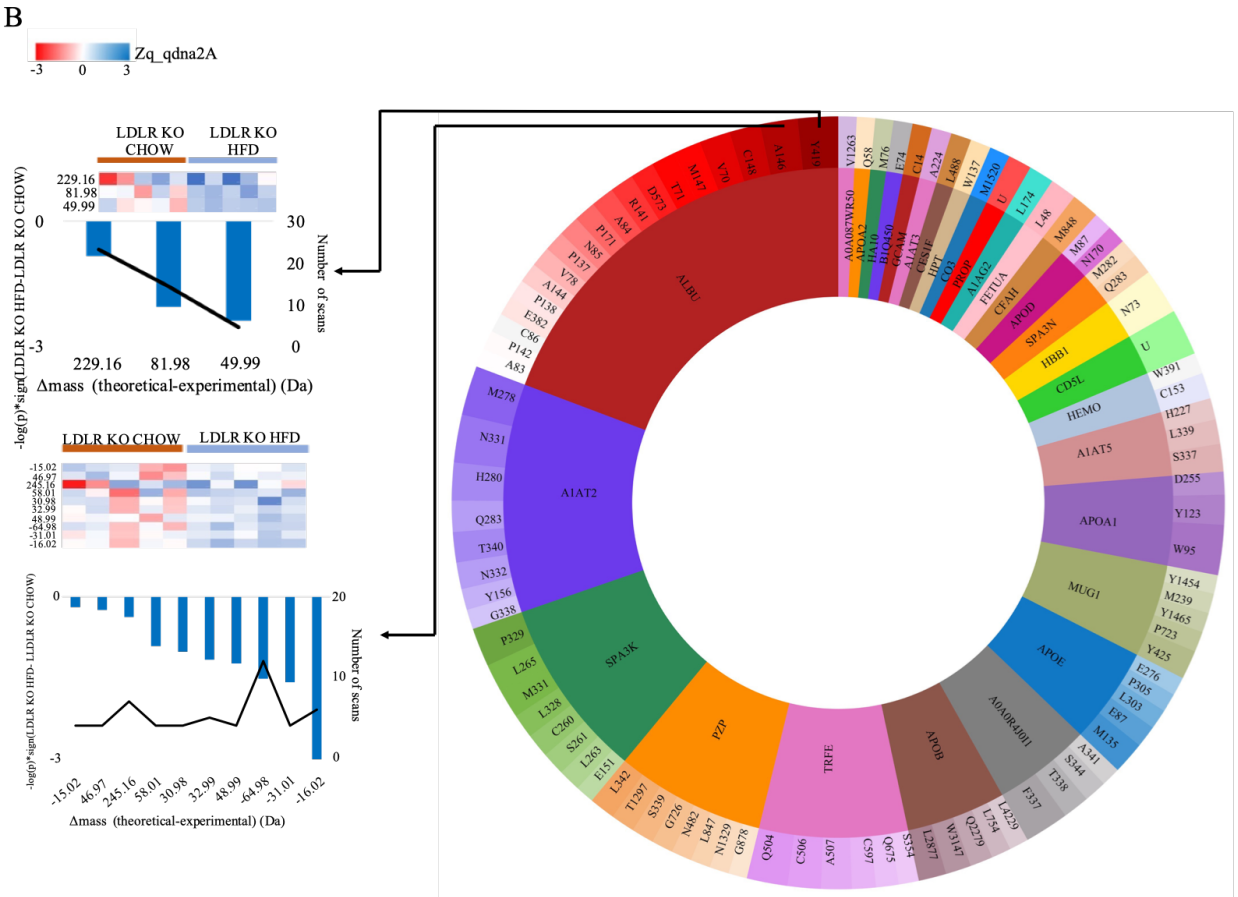

B
Zq_qdna2A
-3 0 3
LDLR KO CHOW
LDLR KO HFD
229.16
81.98
49.99
-log(p)*sign(LDLR KO HFD-LDLR KO CHOW)
Number of scans
Δmass (theoretical-experimental) (Da)
-15.02
46.97
245.16
58.01
30.98
32.99
48.99
-64.98
-31.01
-16.02
ALBU
A1AT2
SPA3K
PZP
TRFE
APOB
A0A0R4J011
APOE
MUG1
APOA1
A1AT5
HEMO
CD5L
HBB1
SPA3N
APOD
CFAH
FETUA
A1AG2
PROP
CO5
HPT
CES1F
A1AT3
GCAM
BTD450
APOA2
A0A087WR50

Figure 4. **High-fat diet-associated protein sites changes in plasma from LDL KO mice.**

Nested pie charts show the protein sites (*qna*) (with more than 4 quantified spectra) with statistically significant (n=5, ttest, *FDR*<=0.1) increases (A) and decreases (B). The graphs are proportional to the magnitude of the change between the control (LDLR KO CHOW) and the LDLR KO HFD samples, expressed in terms of -log(*p-value*)*sign($z_{q\_qna2A\_HFD} - z_{q\_qna2A\_CHOW}$). The adscription of the protein modified site to the plasma proteins is depicted and the magnitude of the *qnas* composing each protein was correspondingly summed. For the 2-3 most relevant protein sites a heatmap of the standardized protein-corrected protein_site_DM quantitation (z_{q_qdna2A}) in each one of the samples according to the color scale is presented. The DMses modifying each protein site are depicted. Additionally, the magnitude of the change between the control (LDLR KO CHOW) and the LDLR KO HFD samples, expressed in terms of -log(*p-value*)*sign($z_{q\_qdna2A\_HFD} - z_{q\_qdna2A\_CHOW}$) (bars) and the number of spectra quantified (Number of scans) are shown. APOE: Apolipoprotein E; ALBU: Serum albumin; HPT: Haptoglobin; Q58EV2: Apoa1; FETUA: Alpha-2-HS-glycoprotein; A1AG1: Alpha-1-acid glycoprotein 1; APOA4: Apolipoprotein A-IV; SPA3K: Serine protease inhibitor A3K; A1AT2: Alpha-1-antitrypsin 1-2; SPA3K:Serine protease inhibitor A3; PZP :Pregnancy zone; TRFE: Serotransferrin; APOB: Apolipoprotein B-100; A0A0R4J0I1: Serpina3k; APOE: Apolipoprotein E; MUG1: Murinoglobulin-1; APOA1: Apolipoprotein A; A1AT5: Alpha-1-antitrypsin 1-5; HEMO: Hemopexin; CD5L: CD5 antigen-like; HBB1: Hemoglobin subunit beta-1; SPA3N: Serine protease inhibitor A3N; APOD: Apolipoprotein D; CFAH: Complement factor H; FETUA: Alpha-2-HS-glycoprotein; A1AG2: Alpha-1-acid glycoprotein; PROP: Properdin; CO3: Complement component 3; HPT: Haptoglobin; CES1F: Carboxylic ester hydrolase; A1AT3: Alpha-1-antitrypsin 1-3; GCAM: Ig gamma-2A chain C region; B1Q450: Hemoglobin beta chain subunit; HA10: H-2 class I histocompatibility antigen, Q10 alpha chain; APOA2: Apolipoprotein A-II; A0A087WR50: Fibronectin.

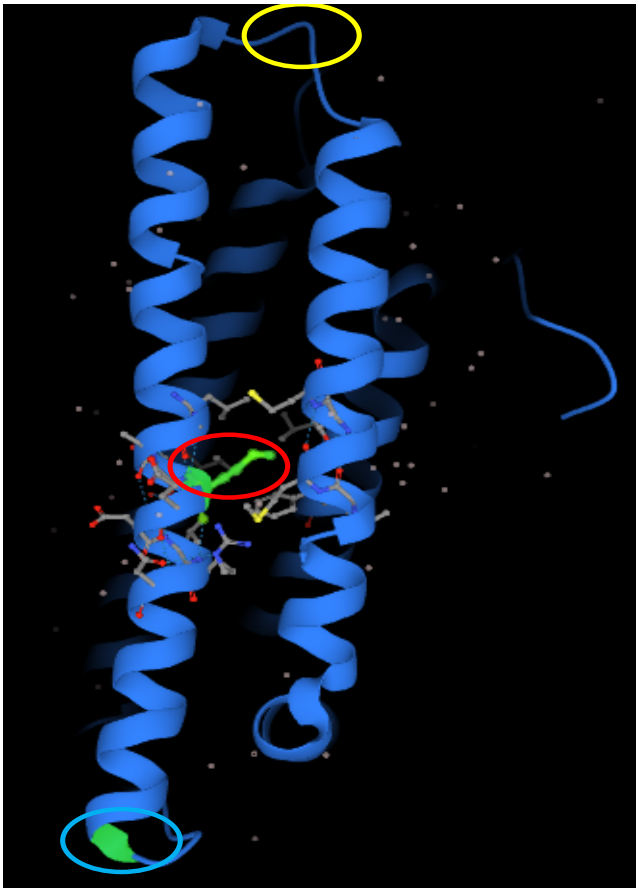


Figure 5. **Apolipoprotein E 3D structure.**

ApoE 3D structure taken from [55] P94 (circled in yellow), M118 (circled in red), M135 (circled in blue) are depicted in green.

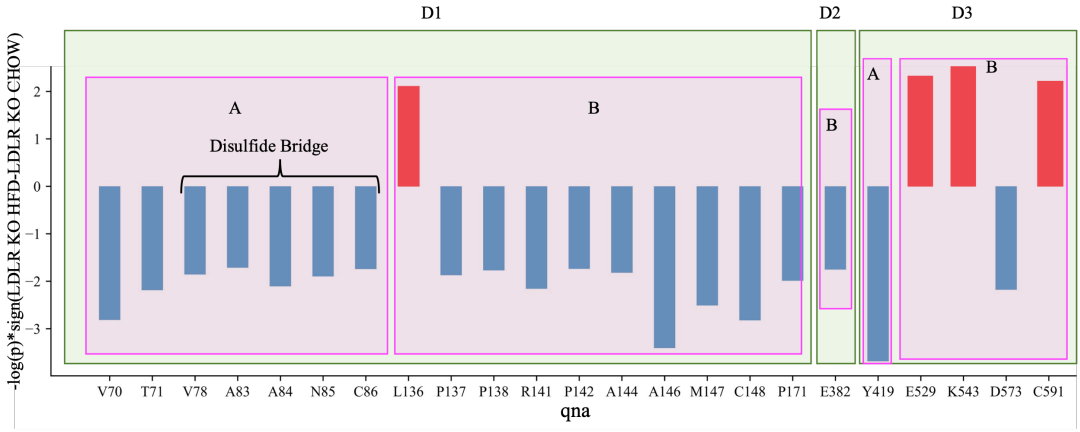


Figure 6. **Albumin qnas and domains.**

Albumin sites (qna) (with more than 4 quantified spectra) with statistically significant (n=5, ttest, FDR<=0.1) increases (red) and decreases (blue). The bars are proportional to the magnitude of the change between the control (LDLR KO CHOW) and the LDLR KO HFD samples, expressed in terms of -log(p-value)*sign($z_{q\_qna2A\_HFD} - z_{q\_qna2A\_CHOW}$). Green squares represent domains, being D1 domain one, D2 domain two and D3 domain three. Pink boxes represent the subdomains A or B. Disulfide bridge is also represented englobing V78, A83, A84, N85 and C86.

Bibliography

1. Mann, M., & Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nature biotechnology*, *21*(3), 255-261.
2. Bagwan, N., Bonzon-Kulichenko, E., Calvo, E., Lechuga-Vieco, A. V., Michalakopoulos, S., Trevisan-Herraz, M., ... & Enríquez, J. A. (2018). Comprehensive quantification of the modified proteome reveals oxidative heart damage in mitochondrial heteroplasmy. *Cell reports*, *23*(12), 3685-3697.
3. Ryan, B. J., Nissim, A., & Winyard, P. G. (2014). Oxidative post-translational modifications and their involvement in the pathogenesis of autoimmune diseases. *Redox biology*, *2*, 715-724.
4. Cubedo, J., Padro, T., Garcia-Moll, X., Pinto, X., Cinca, J., & Badimon, L. (2011). Proteomic signature of Apolipoprotein J in the early phase of new-onset myocardial infarction. *Journal of proteome research*, *10*(1), 211-220.
5. Spinelli, F. R., Pecani, A., Conti, F., Mancini, R., Alessandri, C., & Valesini, G. (2016). Post-translational modifications in rheumatoid arthritis and atherosclerosis: focus on citrullination and carbamylation. *Journal of International Medical Research*, *44*(1_suppl), 81-84.
6. Natarajan, P., Bis, J. C., Bielak, L. F., Cox, A. J., Dörr, M., Feitosa, M. F., ... & Jhun, M. A. (2016). Multiethnic exome-wide association study of subclinical atherosclerosis. *Circulation: Cardiovascular Genetics*, *9*(6), 511-520.
7. López-Díez, R., Shekhtman, A., Ramasamy, R., & Schmidt, A. M. (2016). Cellular mechanisms and consequences of glycation in atherosclerosis and obesity. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, *1862*(12), 2244-2252.
8. Jakubowski, H., Zhang, L., Bardeguez, A., & Aviv, A. (2000). Homocysteine thiolactone and protein homocysteinylation in human endothelial cells: implications for atherosclerosis. *Circulation research*, *87*(1), 45-51.
9. Delporte, C., Boudjeltia, K. Z., Noyon, C., Furtmüller, P. G., Nuyens, V., Slomianny, M. C., ... & Koyani, C. N. (2014). Impact of myeloperoxidase-LDL interactions on enzyme activity and subsequent posttranslational oxidative modifications of apoB-100. *Journal of lipid research*, *55*(4), 747-757.
10. Schurgers, L. J., Spronk, H. M. H., Skepper, J. N., Hackeng, T. M., Shanahan, C. M., Vermeer, C., ... & Proudfoot, D. (2007). Post-translational modifications regulate matrix Gla protein function: importance for inhibition of vascular smooth muscle cell calcification. *Journal of Thrombosis and Haemostasis*, *5*(12), 2503-2511.
11. Jiang, W., Agrawal, D. K., & Boosani, C. S. (2018). Cell-specific histone modifications in atherosclerosis. *Molecular Medicine Reports*, *18*(2), 1215-1224.
12. Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, *422*(6928), 198-207.
13. Pathan, M., Samuel, M., Keerthikumar, S., & Mathivanan, S. (2017). Unassigned MS/MS Spectra: Who Am I?. In *Proteome Bioinformatics* (pp. 67-74). Humana Press, New York, NY.
14. Griss, J., Perez-Riverol, Y., Lewis, S., Tabb, D. L., Dianes, J. A., Del-Toro, N., ... & Wang, R. (2016). Recognizing millions of consistently unidentified spectra across hundreds of shotgun proteomics datasets. *Nature methods*, *13*(8), 651-656.
15. Skinner, O. S., & Kelleher, N. L. (2015). Illuminating the dark matter of shotgun proteomics. *Nature biotechnology*, *33*(7), 717-718.
16. Trevisan-Herraz, M., Bagwan, N., García-Marqués, F., Rodriguez, J. M., Jorge, I., Ezkurdia, I., ... & Vázquez, J. (2019). SanXoT: a modular and versatile package for the quantitative analysis of high-throughput proteomics experiments. *Bioinformatics*, *35*(9), 1594-1596.
17. Li, B., Li, W., Li, X., & Zhou, H. (2017). Inflammation: a novel therapeutic target/direction in atherosclerosis. *Current pharmaceutical design*, *23*(8), 1216-1227.

18. Mushenkova, N. V., Summerhill, V. I., Silaeva, Y. Y., Deykin, A. V., & Orekhov, A. N. (2019). Modelling of atherosclerosis in genetically modified animals. *American journal of translational research*, *11*(8), 4614.
19. Chambless, L. E., Folsom, A. R., Davis, V., Sharrett, R., Heiss, G., Sorlie, P., ... & Evans, G. W. (2002). Risk factors for progression of common carotid atherosclerosis: the Atherosclerosis Risk in Communities Study, 1987–1998. *American journal of epidemiology*, *155*(1), 38-47.
20. Thompson, D., Morrice, N., Grant, L., Le Sommer, S., Lees, E. K., Mody, N., ... & Delibegovic, M. (2017). Pharmacological inhibition of protein tyrosine phosphatase 1B protects against atherosclerotic plaque formation in the LDLR−/− mouse model of atherosclerosis. *Clinical Science*, *131*(20), 2489-2501.
21. Westmuckett, A. D., & Moore, K. L. (2009). Lack of tyrosylprotein sulfotransferase activity in hematopoietic cells drastically attenuates atherosclerosis in Ldlr−/− mice. *Arteriosclerosis, thrombosis, and vascular biology*, *29*(11), 1730-1736.
22. Stehbens, W. E. (1995). The hypothetical epidemic of coronary heart disease and atherosclerosis. *Medical hypotheses*, *45*(5), 449-454.
23. Devesa, C. (2019). Comprehensive proteomics analysis of plasma from ApoE- and LDLR-deficient mice reveals atherosclerosis biomarkers with potential application in clinic. (Trabajo Fin de Máster). Centro Nacional de Investigaciones Cardiovasculares (CNIC). Universidad Complutense de Madrid (UCM), Madrid, Spain.
24. Wiśniewski, J. R., Zougman, A., Nagaraj, N., & Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nature methods*, *6*(5), 359.
25. Martínez-Acedo, P., Nu´n˜ez, E., Go´mez, F.J., Moreno, M., Ramos, E., Iz- quierdo-A´ lvarez, A., Miro´ -Casas, E., Mesa, R., Rodriguez, P., Martı´nez-Ruiz, A., et al. (2012). A novel strategy for global analysis of the dynamic thiol redox proteome. Mol. Cell. Proteomics *11*, 800–813.
26. Navarro, P., Trevisan-Herraz, M., Bonzon-Kulichenko, E., Núñez, E., Martínez-Acedo, P., Pérez-Hernández, D., ... & Hernáez, M. L. (2014). General statistical framework for quantitative proteomics by stable isotope labeling. *Journal of proteome research*, *13*(3), 1234-1247.
27. García-Marqués, F., Trevisan-Herraz, M., Martínez-Martínez, S., Camafeita, E., Jorge, I., Lopez, J. A., ... & Andrés, V. (2016). A novel systems-biology algorithm for the analysis of coordinated protein responses using quantitative proteomics. *Molecular & Cellular Proteomics*, *15*(5), 1740-1760.
28. R. S. P. Rao, N. Zhang, D. Xu, and I. M. Moller, "CarbonylDB: a curated data- resource of protein carbonylation sites.," *Bioinformatics*, vol. 34, no. 14, pp. 2518–2520, Jul. 2018.
29. D. Alocci, J. Mariethoz, A. Gastaldello, E. Gasteiger, N. G. Karlsson, D. Kolarich, N. H. Packer, and F. Lisacek, "GlyConnect: Glycoproteomics Goes Visual, Interactive, and Analytical.," *J. Proteome Res.*, vol. 18, no. 2, pp. 664–677, Feb. 2019.
30. M. Blanc, F. David, L. Abrami, D. Migliozzi, F. Armand, J. Burgi, and F. G. van der Goot, "SwissPalm: Protein Palmitoylation database.," *F1000Research*, vol. 4, p. 261, 2015.
31. P. V Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, and E. Skrzypek, "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations.," *Nucleic Acids Res.*, vol. 43, no. Database issue, pp. D512-20, Jan. 2015.
32. M. P. Campbell and N. H. Packer, "UniCarbKB: New database features for integrating glycan structure abundance, compositional glycoproteomics data, and disease associations.," *Biochim. Biophys. Acta*, vol. 1860, no. 8, pp. 1669– 1675, Aug. 2016.
33. K. E. Ross, H. Huang, J. Ren, C. N. Arighi, G. Li, C. O. Tudor, M. Lv, J.-Y. Lee, S.-C. Chen, K. Vijay-Shanker, and C. H. Wu, "iPTMnet: Integrative Bioinformatics for Studying PTM Networks.," *Methods Mol. Biol.*, vol. 1558, pp. 333–353, 2017.
34. D. M. Creasy and J. S. Cottrell, "Unimod: Protein modifications for mass spectrometry.," *Proteomics*, vol. 4, no. 6, pp. 1534–1536, Jun. 2004.
35. J.K. Eng, A.L. McCormack, J.R. Yates. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* 5 (1994), pp. 976-989. doi: 10.1016/1044-0305(94)80016-2
36. Nesvizhskii, A. I., Roos, F. F., Grossmann, J., Vogelzang, M., Eddes, J. S., Gruissem, W., ... & Aebersold, R. (2006). Dynamic spectrum quality assessment and iterative computational analysis of shotgun proteomic data: toward more efficient identification of post-translational modifications, sequence polymorphisms, and novel peptides. *Molecular & Cellular Proteomics*, *5*(4), 652-670.

37. Chalkley, R. J., Baker, P. R., Medzihradszky, K. F., Lynn, A. J., & Burlingame, A. L. (2008). In-depth analysis of tandem mass spectrometry data from disparate instrument types. *Molecular & Cellular Proteomics*, *7*(12), 2386-2398.
38. Bonzon-Kulichenko, E., Camafeita, E., López, J. A., Gómez-Serrano, M., Jorge, I., Calvo, E., ... & Vázquez, J. (2020). Improved integrative analysis of the thiol redox proteome using filter-aided sample preparation. *Journal of proteomics*, *214*, 103624.
39. Chang, H. Y., Kong, A. T., da Veiga Leprevost, F., Avtonomov, D. M., Haynes, S. E., & Nesvizhskii, A. I. (2020). Crystal-C: A Computational Tool for Refinement of Open Search Results. *Journal of proteome research*, *19*(6), 2511-2515.
40. Getz, G. S., & Reardon, C. A. (2009). Apoprotein E as a lipid transport and signaling protein in 15.The blood, liver, and artery wall. *Journal of lipid research*, *50*(Supplement), S156-S161.
41. Pendse, A. A., Arbones-Mainar, J. M., Johnson, L. A., Altenburg, M. K., & Maeda, N. (2009). Apolipoprotein E knock-out and knock-in mice: atherosclerosis, metabolic syndrome, and beyond. *Journal of lipid research*, *50*(Supplement), S178-S182.
42. Puppione, D.L., et al., *Mass spectral analysis of the apolipoproteins on mouse high density lipoproteins. Detection of post-translational modifications.* BiochimBiophys Acta, 2006. **1764**(8): p. 1363-71.
43. Emamzadeh, F. N. (2017). Role of Apolipoproteins and α-Synuclein in Parkinson's Disease. *Journal of Molecular Neuroscience*, *62*(3), 344-355.
44. Berlett, B. S., & Stadtman, E. R. (1997). Protein oxidation in aging, disease, and oxidative stress. *Journal of Biological Chemistry*, *272*(33), 20313-20316.
45. Baumgartner, R., Berg, M., Matic, L., Polyzos, K. P., Forteza, M. J., Hjorth, S. A., ... & Ketelhuth, D. F. J. (2020). Evidence that a deviation in the kynurenine pathway aggravates atherosclerotic disease in humans. *Journal of Internal Medicine*.
46. Baumgartner, R., Forteza, M. J., & Ketelhuth, D. F. (2019). The interplay between cytokines and the Kynurenine pathway in inflammation and atherosclerosis. *Cytokine*, *122*, 154148.
47. Takahashi, S., Sakai, J., Fujino, T., Hattori, H., Zenimaru, Y., Suzuki, J., ... & Yamamoto, T. T. (2004). The very low-density lipoprotein (VLDL) receptor: characterization and functions as a peripheral lipoprotein receptor. *Journal of atherosclerosis and thrombosis*, *11*(4), 200-208.
48. Ali, M. H., & Imperiali, B. (2005). Protein oligomerization: how and why. *Bioorganic & medicinal chemistry*, *13*(17), 5013-5020.
49. Demer, L. L., Watson, K. E., & Boström, K. (1994). Mechanism of calcification in atherosclerosis. *Trends in cardiovascular medicine*, *4*(1), 45-49.
50. Kalampogias, A., Siasos, G., Oikonomou, E., Tsalamandris, S., Mourouzis, K., Tsigkou, V., ... & Tousoulis, D. (2016). Basic mechanisms in atherosclerosis: the role of calcium. *Medicinal Chemistry*, *12*(2), 103-113.
51. Sitar, M. E., Aydin, S., & Cakatay, U. (2013). Human serum albumin and its relation with oxidative stress. *Clin Lab*, *59*(9-10), 945-52.
52. Lapenna, D., Ciofani, G., Ucchino, S., Pierdomenico, S. D., Cuccurullo, C., Giamberardino, M. A., & Cuccurullo, F. (2010). Serum albumin and biomolecular oxidative damage of human atherosclerotic plaques. *Clinical biochemistry*, *43*(18), 1458-1460.
53. Beddhu, S., Kaysen, G. A., Yan, G., Sarnak, M., Agodoa, L., Ornt, D., ... & HEMO Study Group. (2002). Association of serum albumin and atherosclerosis in chronic hemodialysis patients. *American journal of kidney diseases*, *40*(4), 721-727.
54. Guizado, T. C. (2014). Analysis of the structure and dynamics of human serum albumin. *Journal of molecular modeling*, *20*(10), 1-13.
55. Hatters, D.M., C.A. Peters-Libeu, and K.H. Weisgraber, *Engineering conformational destabilization into mouse apolipoprotein E. A model for a unique property of human apolipoprotein E4.* J Biol Chem, 2005. 280(28): p. 26477-82.

Supplementary material

Supplementary configuration files tables:

Supplementary Table1. Mass modification configuration file

[Aminoacids]	
A	71.037114
B	115,026944
R	156.101111
N	114.042927
D	115.026943
C	103.009185
E	129.042593
Q	128.058578
G	57.021464
H	137.058912
I	113.084064
L	113.084064
K	128.094963
M	131.040485
F	147.068414
P	97.052764
S	87.032028
T	101.047679
U	150.953630
W	186.079313
Y	163.063329
V	99.068414
O	132.089878
Z	129.042594
[Fix_modifications]	
Nt	229.162932;TMT-labeled N-term
C	57.021464;C-IAM
K	229.162932;K-TMT
[Masses]	
m_proton	1.007276
m_hydrogen	1.007825
m_oxygen	15.994915

Supplementary Table2. DMCalibrator, PeakModeller and PeakFDRer configuration files.

[DMcalibrator]	
scorecolumn	xcorr
mzcolumn	exp_mz
zcolumn	charge
seqcolumn	plain_peptide
seqdmcolumn	delta_peptide
proteincolumn	protein
decoyprefix	DECOY
score_min	3
ppm_max	31
calseqcolumn	Cal_Sequence
decimal_places	6
[PeakModeller]	
bins	0.002
slope_points	7
smooth_points	7
[PeakSelector]	
slope	1000
frequency	0
apex_points	4
[PeakAssignator]	
ppm_max	10
peak_label	PEAK
orphan_label	ORPHAN
caldeltamh_column	cal_dm_mh
theomh_column	Theo_mh
closestpeak_column	ClosestPeak
peak_column	PeakAssignation
deltamass_column	deltaMass
ppm_column	ppmError
assignseqcolumn	Assign_Sequence
[PeakFDRer]	
score_column	xcorr
recom_data	0
[Logging]	
create_log	1
create_ini	0

Supplementary Table 3. FDRFilterer configuration file.

[Parameters]	
GlobalThres	0.05
PeakThres	0.01
LocalThres	0.01
GlobalFDR_colum_name	GlobalFDR
PeakFDR_colum_name	PeakFDR
LocalFDR_colum_name	LocalFDR
Label_colum_name	Label
decoys_naming	DECOY

Supplementary Table 4. DM0Solver configuration file.

[DM0Solver_Parameters]	
Relative_Error_ppm	10
Exp_mh_colum_name	cal_exp_mh
Theo_mh_colum_name	theo_mh
Sequence_colum_name	Cal_Sequence
[DM0Solver_DM0List]	
DM0	0
DM0;C13	1.003355
DM0;2C13	2.00671
DM0;3C13	3.010065
Na_adduct	21.981943
Ammonium_adduct	17.026549
Ammonium_adduct_loss	-17.026549
H20_loss	-18.010565

Supplementary Table 5. TrunkSolver configuration file.

[TrunkSolver_Parameters]	
Relative_Error	10
Exp_mh_column_name	cal_exp_mh
Theo_mh_column_name	theo_mh
Sequence_column_name	DM0Sequence
Calibrated_Delta_MH_column_name	cal_dm_mh
MasterProtein_column_name	protein
decnum	6
static_modifications_column_name	modifications
New_Deltamass_output_column_name	New_DM
New_Theo_mh_output_column_name	New_Theo_mh
x	6
TrunkSequence_output_column_name	TrunkSequence
TrunkDM_output_column_name	TrunkDM
TrunkLabel_output_column_name	TrunkLabel
TrunkLabel_ppm_output_column_name	TrunkLabel_ppm
Static_modifications_position_output_column_name	Static_modifications_position
Matchnumber_output_column_name	Match_number
Possible_option_output_column_name	Possible_option
output_file_suffix	_TS_C
[TrunkSolver_CombList]	
DM0	0
(+)TMT	229.162932
(+)2TMT	458.325864
(-)TMT	-229.162932
(-)2TMT	-458.325864

Supplementary Table 6. PeakAssignator second execution configuration file.

[PeakAssignator]	
ppm_max	10
peak_label	PEAK
orphan_label	ORPHAN
caldeltamh_column	TrunkDM
theomh_column	New_Theo_mh
closestpeak_column	New_ClosestPeak
peak_column	New_PeakAssignation
deltamass_column	New_Assigned_deltaMass
ppm_column	New_ppmError
assignseqcolumn	New_Assign_Sequence

Supplementary Table 7. PDMTableMaker first execution configuration file.

[PDMTableMaker_Parameters]	
Sequence_column_name	New_Assigned_Sequence
DM_column_name	New_Assigned_deltaMass
Theo_mh_column_name	New_Theo_mh
Outfile_suffix	_PDMTable
MasterProtein_column_name	protein
[PDMTableMaker_Conditions]	
number_of_conditions	1
Condition1	New_PeakAssignation
Value1	PEAK
MasterProtein_column_name	protein

Supplementary Table 8. SiteSolver configuration file.

[SiteSolver_Parameters]	
Relative_Error_ppm	10
Theo_mh_column_name	New_Theo_mh
Sequence_column_name	New_Assigned_Sequence
cal_Dm_mh_column_name	New_Assigned_deltaMass
PeakNaming	PEAK
PeakAssignation_column_name	New_PeakAssignation
x	3
MinScanFreq	2
MaxNSite	3
PrimaryList_column_name	Freq
SecondaryList_column_name	Freq
SiteCorrection_PrimaryList_label	PL_
SiteCorrection_SecondaryList_label	SL_
SiteSequence_column_name	SiteSequence
SiteCorrection_column_name	SiteCorrection
SiteOption_column_name	SiteOption
SiteDM_column_name	SiteDM
SiteDMError_ppm_column_name	SiteDMError_ppm
Output_file_suffix	_SS

Supplementary Table 9. PDMTableMaker second execution configuration file.

[PDMTableMaker_Parameters]	
Sequence_column_name	SiteSequence
DM_column_name	New_Assigned_deltaMass
Theo_mh_column_name	New_Theo_mh
Outfile_suffix	_PDMTable
MasterProtein_column_name	protein
[PDMTableMaker_Conditions]	
number_of_conditions	0
Condition1	New_PeakAssignation
Value1	PEAK
MasterProtein_column_name	protein

Supplementary Table 10. Sticker configuration file using Artefact file (Supplementary List X)

[Sticker_Parameters]	
Relative_Error_ppm = 10	10
Theo_mh_column_name	Theo_mh
Sequence_column_name	pdm
Selected_DM_column_name	d
StickerLabel_User_output_column_name	Artefact_StickerLabel
StickerLabel_ppm_User_output_column_name	Artefact_StickerLabel_ppm
output_file_suffix	_Artefact

Supplementary Table 11. Sticker configuration file using Unimod file (Supplementary List X)

[Sticker_Parameters]	
Relative_Error_ppm = 10	10
Theo_mh_column_name	Theo_mh
Sequence_column_name	pdm
Selected_DM_column_name	d
StickerLabel_User_output_column_name	Unimod_StickerLabel
StickerLabel_ppm_User_output_column_name	Unimod_StickerLabel_ppm
output_file_suffix	_Unimod

Supplementary Table 12. Sticker configuration file using User file (Supplementary List X)

[Sticker_Parameters]	
Relative_Error_ppm = 10	10
Theo_mh_column_name	Theo_mh
Sequence_column_name	pdm
Selected_DM_column_name	d
StickerLabel_User_output_column_name	Artefact_StickerLabel
StickerLabel_ppm_User_output_column_name	Artefact_StickerLabel_ppm
output_file_suffix	_Artefact

Supplementary lists:

Supplementary List 1. Primary List for SiteSolver.

DM	Residue	Freq
-458.325864	NT	1
-442.330949	K	1
-442.330949	K	1
-422.325864	NT	1
-421.341848	K	1
-406.330949	K	1
-404.315299	K	1
-403.331283	K	1
-388.320384	K	1
-376.312384	K	1
-376.312384	K	1
-360.330488	NT	1
-360.330488	NT	1
-359.333453	K	1
-359.333453	NT	1
-350.304734	K	1
-344.276762	K	1
-344.276762	NT	1
-320.172118	K	1

-300.236431	C	1
-286.220781	K	1
-270.225867	K	1
-260.205131	K	1
-258.189481	K	1
-256.210217	K	1
-254.248711	K	1
-246.189481	K	1
-244.173831	NT	1
-243.214968	K	1
-242.230952	K	1
-230.194566	K	1
-230.194566	NT	1
-229.162932	K	1
-229.162932	NT	1
-228.282014	K	1
-228.282014	K	1
-226.21741	NT	1
-220.198983	K	1
-217.162932	K	1
-214.199652	NT	1
-213.168017	K	1
-210.222496	NT	1
-203.18058	K	1
-203.18058	NT	1
-201.168017	K	1
-201.168017	K	1
-197.173102	NT	1
-197.173102	NT	1
-194.194566	K	1
-187.185665	K	1
-187.185665	NT	1
-187.185665	K	1
-187.152367	M	1
-186.154031	K	1
-186.154031	K	1
-175.152367	NT	1
-174.168351	K	1
-172.138381	K	1
-172.138381	K	1
-171.178582	NT	1
-170.159116	K	1

-170.159116	K	1
-170.159116	M	1
-167.147282	NT	1
-159.157452	K	1
-157.133802	K	1
-157.133802	K	1
-156.143466	NT	1
-156.143466	K	1
-156.143466	M	1
-147.149452	NT	1
-147.149452	NT	1
-143.162537	K	1
-143.126153	K	1
-141.138887	K	1
-141.138887	M	1
-129.138887	NT	1
-129.138887	K	1
-121.141803	NT	1
-115.120005	K	1
-114.141007	K	1
-114.141007	K	1
-113.143972	NT	1
-113.143972	K	1
-112.138097	NT	1
-103.009185	K	1
-98.146092	C	1
-98.146092	K	1
-97.14905703	NT	1
-97.14905703	K	1
-95.14234703	NT	1
-94.13899203	NT	1
-91.009186	NT	1
-90.025169	C	1
-88.993535	C	1
-82.151177	C	1
-82.151177	K	1
-75.014270	NT	1
-72.998620	C	1
-72.11742303	C	1
-72.11742303	K	1
-67.110108	NT	1
-62.977885	K	1

-60.962235	C	1
-59.037114	C	1
-58.982970	C	1
-56.12250803	C	1
-56.12250803	K	1
-46.946584	NT	1
-45.987721	C	1
-45.003705	C	1
-33.987721	C	1
-31.972071	C	1
-31.935685	C	1
-30.988055	C	1
-28.990164	C	1
-25.031634	C	1
-22.971737	C	1
-22.067486	C	1
-12.995249	C	1
-12.962235	C	1
-11.033743	C	1
-9.036719	C	1
-3.929537	C	1
-2.015650	C	1
-1.031634	C	1
-1.012295	NA	1
-0.028279	NA	1
3.032680	NA	1
4.978931	C	1
15.994915	W	1
15.994915	A	1
15.994915	F	1
15.994915	M	1
15.994915	C	1
15.994915	P	1
15.994915	W	1
26.048664	Y	1
31.989829	C	1
31.989829	M	1
31.989829	P	1
31.989829	W	1
31.989829	Y	1
36.000000	F	1
47.984745	K	1

47.984745	W	1
49.955943	Y	1
49.955943	Y	1
54.010565	W	1
57.024551	K	1
57.024551	M	1
57.024551	K	1
57.024551	T	1
57.024551	H	1
62.01565	S	1
61.982636	K	1
71.032095	C	1
71.032095	K	1
72.02912997	NT	1
72.02912997	M	1
72.02912997	Y	1
72.02912997	W	1
73.016379	H	1
73.019466	C	1
89.014381	M	1
91.975442	M	1
100.024045	C	1
100.024045	S	1
152.980587	T	1
154.065171	C	1
168.080821	K	1
182.096471	K	1
197.227247	K	1
209.018035	C	1
230.146948	C	1
230.146948	K	1
244.173831	NT	1
245.157847	Y	1
245.157847	K	1
246.189481	NT	1
246.189481	K	1
261.152762	NT	1
261.152762	K	1
261.152762	NT	1
261.152762	K	1
286.184396	NT	1
286.184396	K	1

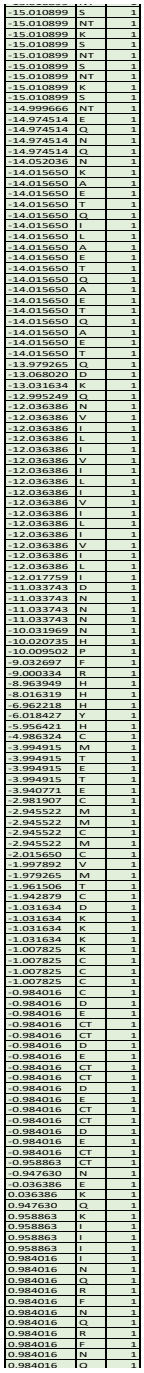
297.084987	NT	1
297.084987	Y	1
298.069003	Y	1
458.325864	Y	1
458.325864	T	1

Supplementary List 2. Secondary list for SiteSolver

DM	AA	Freq
-131.040485	R	1
-129.057849	M	1
-115.042199	W	1
-106.041865	W	1
-99.079647	Y	1
-99.047285	R	1
-94.041865	W	1
-94.041865	V	1
-94.041865	V	1
-94.041865	V	1
-92.026215	V	1
-89.029920	Y	1
-89.026549	M	1
-87.010899	W	1
-85.063997	W	1
-85.031634	R	1
-83.070128	W	1
-80.037448	W	1
-76.031300	H	1
-76.031300	Y	1
-76.031300	F	1
-76.031300	Y	1
-74.019021	F	1
-72.995249	M	1
-72.036386	W	1
-72.021129	W	1
-71.073499	E	1
-71.037114	K	1
-69.069083	Q	1
-66.021798	R	1
-66.010565	H	1
-63.994915	Y	1
-62.015650	Y	1
-60.054144	Y	1
-60.036386	Y	1
-60.003371	F	1
-59.048347	M	1
-58.020735	R	1
-58.005479	W	1
-58.005479	D	1
-58.005479	E	1
-58.005479	G	1
-58.005479	D	1
-58.005479	E	1
-58.005479	G	1
-58.005479	D	1
-58.005479	E	1
-57.057849	G	1
-57.032697	K	1
-57.021464	R	1
-57.021464	N	1
-57.021464	Q	1
-57.021464	N	1
-56.062600	Q	1
-56.062600	L	1
-56.062600	I	1
-56.062600	L	1
-55.053433	I	1
-55.038828	R	1
-53.091927	W	1
-50.026883	R	1
-50.015650	H	1
-49.979265	F	1
-49.020401	Y	1
-49.020401	Y	1
-49.020401	W	1
-49.020401	Y	1
-48.036386	W	1
-48.003371	Y	1
-48.003371	M	1
-48.003371	M	1
-48.003371	M	1
-48.000000	M	1
-46.020735	F	1
-46.005479	F	1
-46.005479	CT	1
-46.005479	CT	1
-46.005479	CT	1
-45.987721	CT	1
-44.059229	C	1
-44.026215	F	1
-44.008456	T	1
-43.989829	M	1
-43.053433	D	1
-43.017047	R	1
-43.005814	R	1
-42.058184	N	1
-42.046950	R	1
-42.021798	V	1
-42.010565	R	1
-41.074168	E	1
-41.062935	R	1
-41.026549	K	1
-40.031300	Q	1
-40.006148	P	1
-39.010899	H	1
-37.990498	W	1
-36.011233	H	1
-35.004751	H	1
-34.968366	Y	1
-34.049727	Y	1
-34.020735	H	1
-33.987721	Y	1
-33.987721	M	1
-33.987721	C	1
-33.987721	C	1
-33.987721	M	1
-33.987721	C	1
-33.987721	C	1
-33.987721	M	1
-33.987721	C	1
-33.984350	C	1
-33.025486	F	1
-33.003705	F	1
-33.003705	C	1
-33.003705	C	1
-33.003705	C	1
-32.041471	C	1
-32.022844	F	1
-32.008456	Y	1
-31.989829	M	1
-31.972071	E	1
-31.972071	M	1
-31.972071	C	1
-31.972071	M	1
-31.042199	C	1
-31.005814	K	1

-30.010565	Q	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-30.010565	E	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-30.010565	E	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-30.010565	E	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-30.010565	E	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-30.010565	E	1
-30.010565	S	1
-30.010565	T	1
-30.010565	P	1
-30.010565	E	1
-30.010565	P	1
-29.992806	E	1
-29.992806	M	1
-29.992806	M	1
-29.992806	M	1
-29.978202	M	1
-29.974179	W	1
-29.026549	E	1
-28.990164	K	1
-28.042534	Q	1
-28.031300	R	1
-28.031300	V	1
-28.031300	M	1
-28.031300	V	1
-28.006148	M	1
-27.994915	R	1
-27.994915	D	1
-27.994915	E	1
-27.994915	P	1
-27.994915	P	1
-27.994915	D	1
-27.994915	E	1
-27.994915	P	1
-27.994915	P	1
-27.994915	D	1
-27.994915	E	1
-27.994915	P	1
-27.994915	P	1
-27.994915	D	1
-27.994915	E	1
-27.994915	P	1
-27.058518	P	1
-27.047285	R	1
-27.010899	K	1
-27.010899	N	1
-27.010899	Q	1
-27.010899	N	1
-26.052036	Q	1
-26.033409	I	1
-26.015650	E	1
-26.004417	P	1
-25.085779	Y	1
-25.060626	K	1
-25.049393	R	1
-23.974848	Q	1
-23.015984	H	1
-23.015984	W	1
-23.015984	H	1
-23.015984	H	1
-23.015984	W	1
-23.015984	H	1
-23.015984	H	1
-23.015984	W	1
-23.015984	H	1
-22.031969	H	1
-22.031969	H	1
-22.031969	H	1
-22.031969	H	1
-21.987721	H	1
-20.026215	M	1
-20.026215	C	1
-20.026215	C	1
-20.026215	C	1
-19.042199	C	1
-19.009836	R	1
-18.973451	F	1
-18.025821	F	1
-18.010565	F	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1

-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-18.010565	E	1
-18.010565	N	1
-18.010565	Q	1
-18.010565	S	1
-18.010565	T	1
-18.010565	Y	1
-18.010565	D	1
-18.010565	C	1
-18.010565	E	1
-17.992806	E	1
-17.992806	C	1
-17.992806	C	1
-17.992806	C	1
-17.974179	C	1
-17.956421	D	1
-17.026549	M	1
-17.026549	T	1
-17.026549	S	1
-17.026549	C	1
-17.026549	N	1
-17.026549	Q	1
-17.026549	T	1
-17.026549	S	1
-17.026549	C	1
-17.026549	N	1
-17.026549	Q	1
-17.026549	T	1
-17.026549	S	1
-17.026549	C	1
-17.026549	N	1
-17.026549	Q	1
-17.026549	T	1
-17.026549	S	1
-17.026549	C	1
-17.026549	N	1
-17.026549	Q	1
-17.026549	T	1
-17.026549	S	1
-17.026549	C	1
-17.026549	N	1
-16.997557	Q	1
-16.990164	M	1
-16.031300	N	1
-16.027929	I	1
-16.013542	F	1
-15.994915	M	1
-15.994915	S	1
-15.994915	Y	1
-15.994915	D	1
-15.994915	T	1
-15.994915	S	1
-15.994915	Y	1
-15.994915	D	1
-15.994915	T	1
-15.994915	S	1
-15.994915	Y	1
-15.994915	D	1
-15.994915	T	1
-15.994915	S	1
-15.994915	Y	1
-15.994915	D	1
-15.977156	T	1
-15.958529	C	1
-15.958529	D	1
-15.958529	E	1
-15.958529	D	1
-15.010899	E	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	K	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	K	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	K	1
-15.010899	S	1
-15.010899	NT	1



-15.010899	S	1
-15.010899	NT	1
-15.010899	K	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	S	1
-15.010899	NT	1
-15.010899	K	1
-15.010899	S	1
-14.999666	NT	1
-14.974514	E	1
-14.974514	Q	1
-14.974514	N	1
-14.974514	Q	1
-14.052036	N	1
-14.015650	K	1
-14.015650	A	1
-14.015650	E	1
-14.015650	T	1
-14.015650	Q	1
-14.015650	I	1
-14.015650	L	1
-14.015650	A	1
-14.015650	E	1
-14.015650	T	1
-14.015650	Q	1
-14.015650	A	1
-14.015650	E	1
-14.015650	T	1
-14.015650	Q	1
-14.015650	A	1
-14.015650	E	1
-14.015650	T	1
-13.979265	Q	1
-13.068020	D	1
-13.031634	K	1
-12.995249	Q	1
-12.036386	N	1
-12.036386	V	1
-12.036386	I	1
-12.036386	L	1
-12.036386	I	1
-12.036386	V	1
-12.036386	I	1
-12.036386	L	1
-12.036386	I	1
-12.036386	V	1
-12.036386	I	1
-12.036386	L	1
-12.036386	I	1
-12.036386	V	1
-12.036386	I	1
-12.017759	I	1
-11.033743	D	1
-11.033743	N	1
-11.033743	N	1
-11.033743	N	1
-10.031969	N	1
-10.020735	H	1
-10.009502	P	1
-9.032697	F	1
-9.000334	R	1
-8.963949	H	1
-8.016319	H	1
-6.962218	H	1
-6.018427	Y	1
-5.956421	H	1
-4.986324	C	1
-3.994915	M	1
-3.994915	T	1
-3.994915	E	1
-3.994915	T	1
-3.940771	E	1
-2.981907	C	1
-2.945522	M	1
-2.945522	M	1
-2.945522	C	1
-2.945522	M	1
-2.015650	C	1
-1.997892	V	1
-1.979265	M	1
-1.961506	T	1
-1.942879	C	1
-1.031634	D	1
-1.031634	K	1
-1.031634	K	1
-1.031634	K	1
-1.007825	K	1
-1.007825	C	1
-1.007825	C	1
-1.007825	C	1
-0.984016	C	1
-0.984016	D	1
-0.984016	E	1
-0.984016	CT	1
-0.984016	CT	1
-0.984016	D	1
0.984016	E	1
-0.984016	CT	1
-0.984016	CT	1
-0.984016	D	1
-0.984016	E	1
-0.984016	CT	1
-0.984016	CT	1
-0.984016	D	1
-0.984016	E	1
-0.984016	CT	1
-0.958863	CT	1
-0.947630	N	1
-0.036386	E	1
0.036386	K	1
0.947630	Q	1
0.958863	K	1
0.958863	I	1
0.958863	I	1
0.958863	I	1
0.984016	I	1
0.984016	N	1
0.984016	Q	1
0.984016	R	1
0.984016	F	1
0.984016	N	1
0.984016	Q	1
0.984016	R	1
0.984016	F	1
0.984016	N	1
0.984016	Q	1

Supplementary List 3. Artefact file for Sticker.

d	Cond_a	Lab_ArtefactIdentifier	Lab_Description	Cond_m
		NoArt		
-229.16358		(-TMT)	No marcado	
-91.01015	C	C_CM-IAM-NL	Artefactos en Cys: C_DHA	
-320.172758	C	C_CM-IAM-NL	Artefactos en Cys: C_DHA-TMT	
0.98554	C	C_CM-IAM-NL	Artefactos en Cys: C_CM/DEA	
155.179435	C	C_CM-IAM-NL	Artefactos en Cys: C_UK	
120.022358	C	C_CM-IAM-NL	Artefactos en Cys: C_DTT	

1.986792	C	C_CM-IAM-NL	Artefactos en Cys: C_2DEA	
-148.032887	C	C_CM-IAM-NL	Artefactos en Cys: C_DHA+C_-IAM	
-246.189985	Q	Cyclation_Nterm-C-Q	C y Q en N-term ciclados y no marcados	1
-246.189985	C	Cyclation_Nterm-C-Q	C y Q en N-term ciclados y no marcados	1
0.98554	Q	DEA_2DEA	Deamidaciones en general	
0.98554	N	DEA_2DEA	Deamidaciones en general	
1.986792	N	DEA_2DEA	Deamidaciones en general	
1.986792		QDEA_2DEA	Deamidaciones en general	
-357.259079		DigPar (-K+TMT)	Dihestiv≥n parcial	
15.993412	M	M_ox,C_ox	Oxidaciv≥n en M,Oxidaciv≥n en C	
15.993412	C	M_ox,C_ox	Oxidaciv≥n en M,Oxidaciv≥n en C	
16.980965	M	M_OX+DEA	En M : OX+DEA	
16.980965	N	M_OX+DEA	En N: OX+DEA	
32.991275	M	M_2OX+DEA/M_2OX+2DEA	En M: 2OX+DEA	
32.991275	N	M_2OX+DEA/M_2OX+2DEA	En N: 2OX+DEA	
33.95246547	M	M_2OX+DEA/M_2OX+2DEA	En M : 2OX+2DEA	
33.95246547	N	M_2OX+DEA/M_2OX+2DEA	En N: 2OX+2DEA	
57.021005	M	M_IAM_CM	Artefactos por IAM y urea en M: M_IAM	
57.021005	H	H_IAM_CM	Artefactos por IAM y urea en H: H_IAM	
57.021005	K	K_IAM_CM	Artefactos por IAM y urea en K: K_IAM	
58.006072	M	M_IAM_CM	Artefactos por IAM y urea en M: M_CM	
58.006072	H	H_IAM_CM	Artefactos por IAM y urea en H: H_CM	
58.006072	K	K_IAM_CM	Artefactos por IAM y urea en K: K_CM	
100.015262	S	T_100.01,S_100.01	Artefactos por hidorxilamina	
100.015262	T	T_100.01,S_100.01	Artefactos por hidorxilamina	
229.162594		TMT	Sobre-marcaje	
230.150072	SN	TMT+DEA	Sobre-marcaje+DEA (sin C)	
230.150072	T	TMT+DEA	Sobre-marcaje+DEA (sin C)	
230.150072	Y	TMT+DEA	Sobre-marcaje+DEA (sin C)	
230.150072	N	TMT+DEA	Sobre-marcaje+DEA (sin C)	
-228.177409		(-TMT+DEA)	No marcado+DEA	
-129.14713	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-129.14713		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-114.137012	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-114.137012		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-147.157817	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	

-147.157817		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-113.152907	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-113.152907		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-203.183662	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-203.183662		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-128.146038	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-128.146038		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-187.153235	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-187.153235		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-157.142248	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-157.142248		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-95.150944	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-95.150944		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-141.078696	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-141.078696		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-172.142429	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-172.142429		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-186.157635	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-186.157635		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1
-171.159939	K	TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	
-171.159939		TMT-fragments	Artefactos asociados al TMT de masa teV≥rica desconocida.	1

Supplementary List 4. Unimod file for Sticker.

d	Lab_Unimod_Label	Lab_Unimod_Description
	NoLab	
495.19519	HexNAc1dHex2	HexNAc1dHex2
494.174789	HexNAc(1)NeuAc(1)	HexNAc NeuAc
486.158471	Hex3	Hex3
486.11556	HexNAc(2)Sulf(1)	HexNAc(2) Sulf
484.228162	EQIGG	Sumo mutant Smt3-WT tail following trypsin digestion
470.266839	Withaferin	Modification of cystein by withaferin
470.163556	dHex(1)Hex(2)	Hex2dHex1
469.716159	triiodo	triiodo

469.228496	NQIGG	SUMOylation by Giardia lamblia
469.143155	Hex(1)NeuGc(1)	Hex NeuGc
456.104615	FMN3	S-(4a-FMN)
456.069261	EGCG1	(-)-epigallocatechin-3-gallate
454.210387	betaFNA	beta-Funaltrexamine
454.088965	FMN	flavin mononucleotide
453.212452	Puromycin	Puromycin
453.14824	Hex(1)NeuAc(1)	Hex NeuAc ---OR--- HexNAc Kdn
452.034807	IASD	Iodoacetamide derivative of stilbene (reaction product with thiol)
445.098527	Hex(1)HexNAc(1)Phos(1)	Hex HexNAc Phos
445.089011	Hex(1)HexNAc(1)Sulf(1)	Hex HexNAc Sulf
440.152991	Hex(1)Pent(2)Me(1)	Hex:1 Pent:2 Me:1
438.094051	FMN2	O3-(riboflavin phosphoryl)
437.201774	bisANS-sulfonates	BisANS with loss of both sulfonates
426.137341	Hex(1)Pent(2)	Hex Pent(2)
420.051719	Unknown:420	Unidentified modification of 420.0506 found in open search
418.137616	dipyrrole	dipyrrolylmethanemethyl
411.259403	LRGG+dimethyl	LeudimethylArgGlyGly
406.158745	HexNAc2	HexNAc2
404.071978	PhosphoHex(2)	H1O3P1Hex2
404.062462	Hex(2)Sulf(1)	Hex(2) O(3) S
397.243753	LRGG+methyl	LeumethylArgGlyGly
387.127779	-Glu-Glu-Glu-	triglutamyl
386.110369	C-Asn-deriv	(3-aminopropyl)(L-aspartyl-1-amino)phosphoryl-5-adenosine
383.228103	LeuArgGlyGly	Ubiquitination
380.147118	cytopiloyne+H2O	nucleophilic addition to cytopiloyne+H2O
377.689944	tri-Iodination	tri-Iodination
372.142033	4AcAllylGal	2,3,4,6-tetra-O-Acetyl-1-allyl-alpha-D-galactopyranoside modification of cysteine
368.344302	C-cholesterol	cholesterol ester
365.132196	Hex(1)HexNAc(1)	Hex1HexNAc1
364.076278	Dap-DSP	Diaminopimelic acid-DSP monolinked
362.136553	cytopiloyne	nucleophilic addtion to cytopiloyne
349.137281	HexNAc1dHex1	HexNAc1dHex1
348.193674	LG-Hlactam-K	Levuglandinyl - lysine hydroxylactam adduct
345.047435	p-guanosine	phospho-guanosine
343.149184	QTGG	SUMOylation leaving GlnThrGlyGly
343.031785	cGMP	S-guanylation Fe-cluster hydrogenase diiron subcluster
340.167459	BADGE	Bisphenol A diglycidyl ether derivative
340.100562	glucosylgalactosyl	glucosylgalactosyl hydroxylysine

340.085794	p-pantetheine	Phosphopantetheine
338.084912	Hex(1)HexA(1)	Hex HexA
335.121631	Pent(1)HexNAc(1)	Pent HexNAc
332.19876	LG-lactam-K	Levuglandinyl - lysine lactam adduct
332.19876	Andro-H2O	andrographolide with the loss of H2O
330.136176	ClGG	Ubiquitin D (FAT10) leaving after chymotrypsin digestion Cys-Ile-Gly-Gly
329.05252	p-adenosine	AMP
327.240959	NA-OA-NO2	Nitroalkylation by Nitro Oleic Acid
325.225309	NA-LNO2	Nitroalkylation by Nitro Linoleic Acid
324.105647	Gal-Glu	Lactosylation
324.035867	RNPXL	Simulate peptide-RNA conjugates
322.020217	IodoU-AMP	(Iodo)-uracil MP
320.100836	ZQG	carbobenzoxy-L-glutaminyl-glycine
316.203845	LG-pyrrole	Levuglandinyl-lysine pyrrole adduct
316.138088	IBTP	Thio Ether Formation - BTP Adduct
314.188195	LG-anhydrolactam	Levuglandinyl-lysine anhydrolactam adduct
310.047738	AFB1_Dialdehyde	adduction of aflatoxin B1 Dialdehyde to lysine
309.205242	Ahx2+Hsl	C-terminal homoserine lactone and two aminohexanoic acids
308.148455	UgiJoullieProGlyProGly	Side reaction of PGPG with Side chain of aspartic or glutamic acid
308.110732	dHex(1)Hex(1)	Hex1dHex1
307.090331	NeuGc	N-glycoyl neuraminic acid
306.171876	LG-Hlactam-R	Levuglandinyl - arginine hydroxylactam adduct
306.095082	Unknown:306	Unidentified modification of 306.0952 found in open search
306.025302	p-uridine	uridine phosphodiester
305.068156	Glutathione	glutathione disulfide
305.041287	PhosphoCytidine	Cytidine monophosphate
301.986514	Unknown:302	Unidentified modification of 301.9864 found in open search
298.19328	LG-anhyropyrrole	Levuglandinyl-lysine anhyropyrrole adduct
294.183109	LTP1-lipid	cis-14-hydroxy-10,13-dioxo-7-heptadecenoic ester
294.095082	Hex(1)Pent(1)	Hex Pent
291.095417	NeuAc	N-acetyl neuraminic acid
290.176961	LG-lactam-R	Levuglandinyl - arginine lactam adduct
287.055563	EGCG2	(-)-dehydroepigallocatechin
283.045704	p-GlcNAc	N-acetylglucosamine-1-phosphoryl
283.036187	s-GlcNAc	O3S1HexNAc1
282.052824	Arg2HPG	bis(hydroxphenylglyoxal) arginine
276.055146	CresylSaligeninPhosphate	Cresyl-Saligenin-phosphorylation
275.100502	MurNAc	N-Acetylmuramic acid

272.250401	Geranyl-geranyl	Geranyl-geranyl
268.039202	NO_SMX_SMCT	Nitroso Sulfamethoxazole semimercaptal thiol adduct
267.158292	DimethylamineGMBS	Modified GMBS X linker
267.031377	NO_SMX_SIMD	Nitroso Sulfamethoxazole Sulfinamide thiol adduct
266.203451	retinal	retinal
266.13068	EHD-diphenylpentanone	2-ethyl-3-hydroxy-1,3-diphenylpentan-1-one
266.057909	Arg2PG	Adduct of phenylglyoxal with Arg
265.146664	MM-diphenylpentanone	3-methyl-5-(methylamino)-1,3-diphenylpentan-1-one
264.187801	didehydroretinyl	3,4-didehydroretinylidene
264.084518	Pent(2)	Pent(2)
263.237491	Tween80	Tween 80 synthetic polymer terminus
263.131014	BMP-piperidinol	1-methyl-3-benzoyl-4-hydroxy-4-phenylpiperidine
258.085186	-Glu-Glu-	diglutamyl
258.014069	Phosphogluconoylation	Phosphogluconoylation
252.044287	NO_SMX_SEMD	Nitroso Sulfamethoxazole Sulphenamide thiol adduct
251.793296	di-Iodination	di-Iodination
249.981018	Unknown:250	Unidentified modification of 249.981 found in open search
248.19876	Unknown:248	Unidentified modification of 248.1986 found in open search
243.085521	pupylation	addition of GGE
242.101505	GGQ	SUMOylation leaving GlyGlyGln
242.019154	p-Man	phosphoglycosyl-D-mannose-1-phosphoryl
240.11503	lapachenole	lapachenole photochemically added to cysteine
238.229666	Palmitoylation	Palmitoylation
236.214016	palmitoleyl	palmitoleyl
234.16198	BHTOH	Michael addition of t-butyl hydroxylated BHT (BHTOH) to C, H or K
234.073953	Unknown:234	Unidentified modification of 234.0742 found in open search
233.051049	dansyl	5-dimethylaminonaphthalene-1-sulfonyl
232.064354	propyl-NAG-thiazoline	propyl-1,2-dideoxy-2\'-methyl-alpha-D-glucopyranoso-[2,1-d]-Delta2\'-thiazoline
231.02966	PyridoxalPhosphateH2	PLP bound to lysine reduced by sodium borohydride (NaBH4) to create amine linkage
229.162932	TMT6plex	Sixplex Tandem Mass Tag
229.014009	Pyridoxal-phos	Pyridoxal phosphate
228.111007	Bacillosamine	2,4-diacetamido-2,4,6-trideoxyglucopyranose
225.090212	Dipyridyl	Cys modified with dipy ligand
220.182715	hydroxyfarnesyl	hydroxyfarnesyl
220.058303	Kdo	Glycosylation with KDO

218.167065	BHT	Michael addition of BHT quinone methide to Cysteine and Lysine
216.099774	Unknown:216	Unidentified modification of 216.1002 found in open search
214.971084	SPITC	4-sulfophenyl isothiocyanate
212.00859	phosphoRibosyl	phosphate-ribosylation
210.986535	TNBS	tri nitro benzene
210.198366	Myristoylation	Myristoylation
210.16198	Unknown:210	Unidentified modification of 210.1616 found in open search
210.00205	CarboxymethylDTT	Carboxymethylated DTT modification of cysteine
209.018035	CarbamidomethylDTT	Carbamidomethylated DTT modification of cysteine
208.182715	myristoleylation	(cis-delta 5)-tetradecaenoyl
206.167065	myristoyl-4H	(cis,cis-delta 5, delta 8)-tetradecadienoyl
204.187801	Farnesylation	Farnesylation
203.950987	Haloxon	O-Dichloroethylphosphate
203.079373	HexNAc	N-Acetylhexosamine
201.970617	Hg	Mercury Mercaptan
198.981352	DNPS	2,4-Dinitrobenzenesulfenyl
197.04531	glycerylPE	glycerylphosphorylethanolamine
192.063388	Hep	Heptose
190.074228	Bromobimane	Monobromobimane derivative
189.188947	EQATd5	EAPTA d5
188.156501	Triton	Triton synthetic polymer terminus
188.032956	Lipoyl	Lipoyl
185.189198	spermine	spermine adduct
184.157563	EQAT	EAPTA d0
184.07961	MTSL	Cys modification by (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanesulfonate (MTSL)
183.983029	3sulfo	derivatization by N-term modification using 3-Sulfobenzoic succinimidyl ester
183.035399	AEBS	Aminoethylbenzenesulfonylation
178.047738	Galactosyl	Gluconoylation
176.744957	Unknown:177	Unidentified modification of 176.7462 found in open search
176.032088	N-glucuronyl	hexuronic acid
175.042199	NDA	naphthalene-2,3-dicarboxaldehyde
175.030314	NEMsulfurWater	N-ethylmaleimideSulfurWater
174.025169	Thiadiazole	Thiadiazolydation of Cys
173.092617	Ub-VME	Ubiquitin vinylmethylester
172.992127	Cys->SecNEM	N-ethylmaleimide on selenocysteines
172.05243	Menadione-HQ	Menadione hydroquinone derivative
170.036779	Menadione-Q	Menadione quinone derivative
170.013281	Cresylphosphate	o-toluyl-phosphorylation

169.048741	His+O(2)	Photo-induced histidine adduct
168.0245	phenylsulfonylethyl	reaction with phenyl vinyl sulfone
167.982375	3-phosphoglyceryl	3-phosphoglyceryl
166.001457	DNCB_hapten	Chemical reaction with 2,4-dinitro-1-chloro benzene (DNCB)
165.164326	Tween20	Tween 20 synthetic polymer terminus
164.060231	O-Diisopropylphosphate	O-Diisopropylphosphorylation
163.04557	PEITC	Phenethyl isothiocyanate
162.125595	Unknown:162	Unidentified modification of 162.1258 found in open search
162.080967	O-pinacolylmethylphosphonate	O-pinacolylmethylphosphonylation
162.052824	Hex	Hexose
161.068808	HexN	Hexosamine
160.037173	glucosone	Condensation product of glucosone
159.932662	pyrophospho	pyrophosphorylation of Ser/Thr
159.068414	serotonylation	5-glutamyl serotonin
159.035399	#¿NOMBRE?	Addition of N-formyl met
159.035399	thioacylPA	membrane protein extraction
158.13068	redHNE	reduced 4-Hydroxynonenal
157.019749	NEMsulfur	N-ethylmaleimideSulfur
156.11503	HNE	4-hydroxynonenal (HNE)
156.101111	Arg	Addition of arginine due to transpeptidation
155.821022	Dibromo	Dibromo
154.135765	decanoyl	lipid
154.09938	4-ONE	4-Oxononenal (ONE)
154.074228	UgiJoullieProGly	Side reaction of PG with Side chain of aspartic or glutamic acid
154.026609	Delta:H(6)C(7)O(4)	methylglyoxal-derived tetrahydropyrimidine
154.00311	glycerophospho	glycerophospho
152.006087	Diethylphosphothione	O-diethylphosphothione
151.996571	DTT	DTT adduct of cysteine
150.041585	cGMP+RMP-loss	S-guanylation-2
149.02992	BITC	Benzyl isothiocyanate
148.037173	glycosyl	glycosyl-L-hydroxyproline
146.057909	Fuc	Fucose
146.036779	hydroxycinnamyl	hydroxycinnamyl
145.019749	CAMthiopropanoyl	3-(carbamidomethylthio)propanoyl
144.042259	3-deoxyglucosone	Condensation product of 3-deoxyglucosone
144.042259	2-dimethylsuccinyl	S-(2-dimethylsuccinyl) cysteine
143.058243	NEMhyd	Nethylmaleimidehydrolysis
142.110613	Diphthamide	Diphthamide
141.042593	Oxidation+NEM	N-ethylmaleimide on cysteine sulfenic acid

138.104465	HNE-Delta:H(2)O	Dehydrated 4-hydroxynonenal
137.16403	TMAB-d9	d9-4-trimethyllammoniumbutyryl-
136.088815	4-ONE+Delta:H(-2)O(-1)	Dehydrated 4-Oxononenal Michael adduct
136.028931	Diethylphosphate	O-Diethylphosphorylation
136.001656	BEMAD_ST	Beta elimination of modified S or T followed by Michael addition of DTT
135.983029	CAF	sulfonation of N-terminus
135.044916	O-Et-N-diMePhospho	O-ethyl, N-dimethyl phosphate
135.032028	2-nitrobenzyl	Tyrosine caged with 2-nitrobenzyl (ONB)
134.048013	PyMIC	3-methyl-2-pyridyl isocyanate
133.052764	Iodoacetanilide	iodoacetanilide derivative
133.019749	HCysteinyl	S-homocysteinylation
132.068748	benzylguanidine	modification of the lysine side chain from NH2 to guanidine with a H removed in favor of a benzyl group
132.057515	Propiophenone	Propiophenone
132.042259	Pentose	Pentose
132.021129	Arg1HPG	Hydroxyphenylglyoxal arginine
132.021129	Difuran	Chemical modification of the diiodinated sites of thyroglobulin by Suzuki reaction
130.026609	2-monomethylsuccinyl	S-(2-monomethylsuccinyl) cysteine
129.057849	Gly->Trp	Gly->Trp substitution
129.042593	-Glu-	monoglutamyl
129.042593	NMMhyd	Nmethylmaleimidehydrolysis
128.131349	spermidine	spermidine adduct
128.107539	TMAB-d0	4-trimethyllammoniumbutyryl-
128.094963	Lys	Addition of lysine due to transpeptidation
127.063329	SMA	N-Succinimidyl-2-morpholine acetate
126.104465	octanoyl	octanoyl
125.896648	Iodination	Iodination
125.047679	NEM	N-ethylmaleimide on cysteines
123.974787	Dimethylphosphothione	O-dimethylphosphothione
123.00853	GPI-anchor	glycosylphosphatidylinositol
122.073165	Delta:H(10)C(8)O(1)	crotonaldehyde-derived dimethyl-FDP-lysine
122.036779	Carboxyethylpyrrole	Carboxyethylpyrrole
122.013281	O-Isopropylphosphate	O-Isopropylphosphorylation
121.035005	PET	phosphorylation to pyridyl thiol
120.034017	O-Isopropylmethylphosphonate	O-Isopropylmethylphosphonylation
120.0245	BEMAD_C	Beta elimination of alkylated Cys followed by Michael addition of DTT
120.0245	ethylsulfonylethyl	reaction with ethyl vinyl sulfone
120.021129	PS_Hapten	reaction with phenyl salicylate (PS)
119.037114	pyridylacetyl	pyridylacetyl

119.037114	PIC	phenyl isocyanate
119.004099	Cysteinyl	Cysteinylation
118.065674	DHP	Dehydropyrrolizidine alkaloid (dehydroretronecine) on cysteines
117.024835	HCysThiolactone	N-Homocysteine thiolactone
116.997965	Phosphopropargyl	phospho-propargylamine
116.010959	2-succinyl	S-(2-succinyl) cysteine
115.042199	Ala->Trp	Ala->Trp substitution
114.042927	GlyGly	ubiquitinylation residue
114.042927	Dicarbamidomethyl	Double Carbamidomethylation
114.031694	Gluratylation	Glutarylation
113.084064	Acetylhypusine	Acetylhypusine
112.05243	Acrolein112	Acrolein addition +112
111.068414	DMPO	DMPO spin-trap nitrone adduct
111.032028	NMM	Nmethylmaleimide
109.959137	monomethylphosphothione	O-methylphosphothione
109.052764	me-pyrroline	4-methyl-delta-1-pyrroline-5-carboxyl
109.048119	dNIC	deuterated Nicotinic Acid
108.975121	Methamidophos-S	S-methyl amino phosphinate
108.021129	HydroxymethylOP	2-ammonio-6-[4-(hydroxymethyl)-3-oxidopyridinium-1-yl]- hexanoate
107.997631	O-Dimethylphosphate	O-Dimethylphosphorylation
107.997631	Ethylphosphate	O-Ethylphosphorylation
107.979873	MeMePhosphorothioate	S-Methyl Methyl phosphorothioate
107.077339	AHA-Alkyne	Azidohomoalanine (AHA) bound to propargylglycine-NH2 (alkyne)
107.013615	N-dimethylphosphate	N-dimethylphosphate
106.041865	Gly->Tyr	Gly->Tyr substitution
106.041865	Saligenin	o-toluene
106.00885	methylsulfonylethyl	reaction with methyl vinyl sulfone
105.897267	Cation:Ag	Replacement of proton by silver
105.057849	S-pyridylethyl	S-pyridylethylation
105.021464	NIC	Nicotinic Acid
104.965913	Cys->CamSec	Sec Iodoacetamide derivative
104.071154	Tris	tris adduct causes 104 Da addition at asparagine-succinimide intermediate
103.960719	DimethylArsino	Reaction with dimethylarsinous (AsIII) acid
101.084064	HN2_mustard	Modification by hydroxylated mechloroethamine (HN-2)
100.016044	Methylmalonylation	Methylmalonylation on Serine
99.079647	Gly->Arg	Gly->Arg substitution
99.068414	NIPCAM	N-isopropylcarboxamidomethyl
99.047285	Ser->Trp	Ser->Trp substitution
98.073165	MesitylOxide	Acetone chemical artifact
97.089149	Acetyldeoxyhypusine	Acetyldeoxyhypusine

97.016378	maleimide	maleimide
96.057515	Delta:H(8)C(6)O(1)	Reduced acrolein addition +96
95.943487	Thio-phospho	Thiophosphorylation
94.967714	sulfo+amino	aminotyrosine with sulfation
94.041865	Acrolein94	Acrolein addition +94
93.981981	O-Methylphosphate	O-Methylphosphorylation
92.997965	Methamidophos-O	O-methyl amino phosphinate
92.026215	Ala->Tyr	Ala->Tyr substitution
90.04695	Gly->Phe	Gly->Phe substitution
89.026549	Pro->Trp	Pro->Trp substitution
87.068414	hypusine	hypusine
87.050655	ser_thr_DAET	phosphorylation to amine thiol
87.032028	glycidamide	glycidamide adduct
87.010899	Val->Trp	Val->Trp substitution
86.036779	HMVK86	Michael addition of hydroxymethylvinyl ketone to cysteine
86.036779	GEE	transamidation of glycine ethyl ester to glutamine
86.036779	hydroxyisobutyryl	2-hydroxyisobutyrylation
86.000394	Malonyl	Malonylation
85.063997	Ala->Arg	Ala->Arg substitution
85.052764	NEIAA-d0	N-ethyl iodoacetamide-d0
85.031634	Thr->Trp	Thr->Trp substitution
83.070128	Cys->Trp	Cys->Trp substitution
80.985078	Arg->Npo	Arginine replacement by Nitropyrimidyl ornithine
80.037448	Gly->His	Gly->His substitution
80.026215	Delta:H(4)C(5)O(1)	methylglyoxal-derived argpyrimidine
79.966331	Phospho	Phosphorylation
79.956815	Sulfation	O-Sulfonation
79.91652	selenyl	selenyl
78.04695	pyrrole	2,5-dimethypyrrole
77.987066	Methylphosphonate	Methylphosphonylation
77.910511	bromo	bromination
76.0313	Acrolein76	Acrolein addition +76
76.0313	Ser->Tyr	Ser->Tyr substitution
76.0313	Ala->Phe	Ala->Phe substitution
75.998285	DeStreak	Cysteine mercaptoethanol
75.980527	EDT	EDT
74.019021	Gly->Met	Gly->Met substitution
72.995249	Xle->Trp	Leu/Ile->Trp substitution
72.036386	Asn->Trp	Asn->Trp substitution
72.021129	carboxyethyl	carboxyethyl
72.021129	Gly->Glu	Gly->Glu substitution

72.021129	Dihydroxyimidazolidine	Dihydroxy methylglyoxal adduct
72.021129	Ethoxyformyl	Ethoxyformylation
71.073499	Deoxyhypusine	Deoxyhypusine
71.073499	Gly->Lys	Gly->Lys substitution
71.073499	Dimethylaminoethyl	Cys alkylation by dimethylaminoethyl halide
71.05237	Asp->Trp	Asp->Trp substitution
71.037114	Propionamide	Acrylamide adduct
71.037114	Gly->Gln	Gly->Gln substitution
71.013304	Delta:H(3)C(3)O(2)	methylglyoxal-derived carboxyethyllysine
70.041865	Croton	Crotonaldehyde
70.041865	Butyryl	Butyryl
70.005479	pyruv-iminyl	N-pyruvic acid 2-iminyl
69.069083	Ser->Arg	Ser->Arg substitution
68.0626	Piperidination	Piperidination
68.026215	Crotonyl	Crotonylation
67.922055	dichlorination	Dichlorination
66.021798	Ala->His	Ala->His substitution
66.010565	Pro->Tyr	Pro->Tyr substitution
66.010565	Furan	Chemical modification of the iodinated sites of thyroglobulin by Suzuki reaction
63.994915	Val->Tyr	Val->Tyr substitution
63.979659	Delta:O(4)	Tryptophan oxidation to dihydroxy-N-formaylkynurenine
63.9619	SulfurDioxide	SulfurDioxide
62.01565	MDA62	MDA adduct +62
62.01565	Thr->Tyr	Thr->Tyr substitution
61.921774	Cation:Cu	Replacement of proton by copper
61.913495	Cation:Zn[II]	Replacement of 2 protons by zinc
60.054144	Cys->Tyr	Cys->Tyr substitution
60.036386	Ser->Phe	Ser->Phe substitution
60.003371	MercaptoEthanol	2-OH-ethyl thio-Ser
60.003371	Ala->Met	Ala->Met substitution
59.04969	trimethyl-OH	5-hydroxy-N6,N6,N6-trimethyl
59.048347	Pro->Arg	Pro->Arg substitution
59.019355	AEC-MAEC	aminoethylcysteine
58.041865	Delta:H(6)C(3)O(1)	Reduced acrolein addition +58
58.029289	Carbofuran	2,3-dihydro-2,2-dimethyl-7-benzofuranol N-methyl carbamate
58.020735	Gln->Trp	Gln->Trp substitution
58.005479	Carboxymethyl	Iodoacetic acid derivative
58.005479	Ala->Glu	Ala->Glu substitution
58.005479	Gly->Asp	Gly->Asp substitution

57.98435	Lys->Trp	Lys->Trp substitution
57.057849	Ala->Lys	Ala->Lys substitution
57.03672	Glu->Trp	Glu->Trp substitution
57.032697	Val->Arg	Val->Arg substitution
57.021464	Carbamidomethyl	Iodoacetamide derivative
57.021464	Ala->Gln	Ala->Gln substitution
57.021464	Gly->Asn	Gly->Asn substitution
57.021464	Gly	Addition of Glycine
56.0626	Diethylation	Diethylation, analogous to Dimethylation
56.0626	Gly->Xle	Gly->Leu/Ile substitution
56.026215	Acrolein56	Acrolein addition +56
55.919696	Cation:Ni[II]	Replacement of 2 protons by nickel
55.053433	Thr->Arg	Thr->Arg substitution
55.038828	Met->Trp	Met->Trp substitution
54.010565	MDA54	MDA adduct +54
54.010565	MG-H1	Methylglyoxal-derived hydroimidazolone
53.971735	trifluoro	trifluoroleucine replacement of leucine
53.919289	Cation:Fe[II]	Replacement of 2 protons by iron
53.091927	Cys->Arg	Cys->Arg substitution
52.911464	Cation:Fe[III]	Replacement of 3 protons by iron
50.026883	Ser->His	Ser->His substitution
50.01565	Pro->Phe	Pro->Phe substitution
49.979265	Xle->Tyr	Leu/Ile->Tyr substitution
49.020401	Asn->Tyr	Asn->Tyr substitution
49.020401	His->Trp	His->Trp substitution
48.036386	Asp->Tyr	Asp->Tyr substitution
48	Val->Phe	Val->Phe substitution
47.984744	Cysteic_acid	cysteine oxidation to cysteic acid
47.944449	SeMet	Selenium replaces sulfur
46.020735	Thr->Phe	Thr->Phe substitution
45.987721	b-methylthiol	Beta-methylthiolation
45.987721	Gly->Cys	Gly->Cys substitution
44.985078	Nitro	Oxidation to nitro
44.059229	Cys->Phe	Cys->Phe substitution
44.026215	EtOH	Ethanolation
44.026215	Gly->Thr	Gly->Thr substitution
44.008456	S-Eth	S-Ethylcystine from Serine
44.008456	Ser->Met	Ser->Met substitution
43.989829	carboxyl	Carboxylation
43.989829	Ala->Asp	Ala->Asp substitution
43.953444	Delta:H(-4)O(3)	Tryptophan oxidation to hydroxy-bis-tryptophandione

43.042199	Ethanolamine	Carboxyl modification with ethanolamine
43.017047	Xle->Arg	Leu/Ile->Arg substitution
43.005814	Carbamyl	Carbamylation
43.005814	Ala->Asn	Ala->Asn substitution
42.058184	Asn->Arg	Asn->Arg substitution
42.04695	tri-Methylation	tri-Methylation
42.04695	Gly->Val	Gly->Val substitution
42.04695	Ala->Xle	Ala->Leu/Ile substitution
42.04695	Propyl	Propyl
42.021798	Guanidination	Guanidination
42.021798	amidino	amidino
42.010565	Acetyl	Acetylation
42.010565	Ser->Glu	Ser->Glu substitution
41.074168	Asp->Arg	Asp->Arg substitution
41.062935	Ser->Lys	Ser->Lys substitution
41.026549	amidine	amidination of lysines or N-terminal amines with methyl acetimidate
41.026549	Ser->Gln	Ser->Gln substitution
41.001397	AzidoF	Azidophenylalanine
40.0313	Propionald+40	Propionaldehyde +40
40.0313	Gly->Pro	Gly->Pro substitution
40.006148	Pro->His	Pro->His substitution
39.994915	Pyro-cmC	S-carbamoylmethylcysteine cyclization (N-terminus)
39.994915	G-H1	Glyoxal-derived hydroimiadazolone
39.010899	Phe->Trp	Phe->Trp substitution
38.01565	Acrolein38	Acrolein addition +38
37.990498	Val->His	Val->His substitution
37.955882	Cation:K	Replacement of proton by potassium
37.946941	Cation:Ca[II]	Replacement of 2 protons by calcium
37.031634	Propargylamine	propargylamine
36.011233	Thr->His	Thr->His substitution
35.004751	Gln->Tyr	Gln->Tyr substitution
34.968366	Lys->Tyr	Lys->Tyr substitution
34.049727	Cys->His	Cys->His substitution
34.020735	Glu->Tyr	Glu->Tyr substitution
33.987721	Pro->Met	Pro->Met substitution
33.98435	Xle->Phe	Leu/Ile->Phe substitution
33.969094	Homocysteic_acid	methionine oxidation to homocysteic acid
33.961028	Chlorination	Chlorination of tyrosine residues
33.951335	Leu->MetOx	Leu->Met substitution and sulfoxidation
33.025486	Asn->Phe	Asn->Phe substitution

32.056407	CHD2	DiMethyl-CHD2
32.041471	Asp->Phe	Asp->Phe substitution
32.022844	Met->Tyr	Met->Tyr substitution
31.989829	dihydroxy	dihydroxy
31.989829	Pro->Glu	Pro->Glu substitution
31.972071	persulfide	persulfide
31.972071	Val->Met	Val->Met substitution
31.972071	Ala->Cys	Ala->Cys substitution
31.935685	Lys->CamCys	Lys->Cys substitution and carbamidomethylation
31.042199	Pro->Lys	Pro->Lys substitution
31.005814	Pro->Gln	Pro->Gln substitution
30.010565	hydroxymethyl	hydroxymethyl
30.010565	Ala->Thr	Ala->Thr substitution
30.010565	Gly->Ser	Gly->Ser substitution
30.010565	methylol	formaldehyde induced modifications
29.992806	Thr->Met	Thr->Met substitution
29.978202	Arg->Trp	Arg->Trp substitution
29.974179	quinone	quinone
29.974179	Val->Glu	Val->Glu substitution
29.039125	Delta:H(5)C(2)	Dimethylation of proline residue
29.026549	Val->Lys	Val->Lys substitution
29.015316	Ethyl+Deamidated	deamidation followed by esterification with ethanol
28.990164	SNO	nitrosylation
28.990164	Val->Gln	Val->Gln substitution
28.042534	Gln->Arg	Gln->Arg substitution
28.0313	di-Methylation	di-Methylation
28.0313	Acetald+28	Acetaldehyde +28
28.0313	Ethyl	Ethylation
28.0313	Ala->Val	Ala->Val substitution
28.0313	Cys->Met	Cys->Met substitution
28.006148	Lys->Arg	Lys->Arg substitution
27.994915	Formyl	Formylation
27.994915	Ser->Asp	Ser->Asp substitution
27.994915	Thr->Glu	Thr->Glu substitution
27.958529	Delta:H(-4)O(2)	Tryptophan oxidation to beta-unsaturated-2,4-bis-tryptophandione
27.058518	Glu->Arg	Glu->Arg substitution
27.047285	Thr->Lys	Thr->Lys substitution
27.047285	ethylamino	ethyl amino
27.010899	Ser->Asn	Ser->Asn substitution
27.010899	Thr->Gln	Thr->Gln substitution
26.052036	Ser->Xle	Ser->Leu/Ile substitution

26.033409	Cys->Glu	Cys->Glu substitution
26.01565	Acetald+26	Acetaldehyde +26
26.01565	Ala->Pro	Ala->Pro substitution
26.004417	His->Tyr	His->Tyr substitution
25.085779	Cys->Lys	Cys->Lys substitution
25.060626	Met->Arg	Met->Arg substitution
25.049393	Cys->Gln	Cys->Gln substitution
24.995249	cyano	cyano
23.974848	Xle->His	Leu/Ile->His substitution
23.958063	Cation:Al[III]	Replacement of 3 protons by aluminium
23.015984	Asn->His	Asn->His substitution
23.015984	Tyr->Trp	Tyr->Trp substitution
22.031969	Asp->His	Asp->His substitution
21.98435	glyoxalAGE	glyoxal-derived AGE
21.981943	Sodiated	Sodium adduct
21.969392	Cation:Mg[II]	Replacement of 2 protons by magnesium
19.989829	hydroxykynurenin	tryptophan oxidation to hydroxykynurenin
19.042199	His->Arg	His->Arg substitution
19.009836	Gln->Phe	Gln->Phe substitution
18.973451	Lys->Phe	Lys->Phe substitution
18.940436	Lys->MetOx	Lys->Met substitution and sulfoxidation
18.025821	Glu->Phe	Glu->Phe substitution
18.010565	Pro->HAVA	Proline oxidation to 5-hydroxy-2-aminovaleric acid
17.990578	Fphe	fluorination
17.974179	Pro->Asp	Pro->Asp substitution
17.956421	Xle->Met	Leu/Ile->Met substitution
17.03448	D3-Me-ester	deuterated methyl ester
17.026549	Ammonium	replacement of proton with ammonium ion
16.997557	Asn->Met	Asn->Met substitution
16.990164	Pro->Asn	Pro->Asn substitution
16.0313	Pro->Xle	Pro->Leu/Ile substitution
16.028204	DeMet	Deuterium Methylation of Lysine
16.027929	Met->Phe	Met->Phe substitution
16.013542	Asp->Met	Asp->Met substitution
15.994915	Hydroxylation	Oxidation or Hydroxylation
15.994915	Ala->Ser	Ala->Ser substitution
15.994915	Phe->Tyr	Phe->Tyr substitution
15.977156	thiocarboxy	thiocarboxylic acid
15.977156	Ser->Cys	Ser->Cys substitution
15.958529	Val->Asp	Val->Asp substitution

15.958529	Xle->Glu	Leu/Ile->Glu substitution
15.010899	aminotyrosine	Tyrosine oxidation to 2-aminotyrosine
15.010899	Xle->Lys	Leu/Ile->Lys substitution
15.010899	Hydroxamic_acid	ADP-ribosylation followed by conversion to hydroxamic acid via hydroxylamine
14.999666	Methyl+Deamidated	Deamidation followed by a methylation
14.999666	Asn->Glu	Asn->Glu substitution
14.974514	Xle->Gln	Leu/Ile->Gln substitution
14.974514	Val->Asn	Val->Asn substitution
14.96328	aminoadipic	alpha-amino adipic acid
14.052036	Asn->Lys	Asn->Lys substitution
14.01565	Methyl	Methylation
14.01565	Asp->Glu	Asp->Glu substitution
14.01565	Gly->Ala	Gly->Ala substitution
14.01565	Ser->Thr	Ser->Thr substitution
14.01565	Val->Xle	Val->Leu/Ile substitution
14.01565	Asn->Gln	Asn->Gln substitution
13.979265	Pyroglutamic	proline oxidation to pyroglutamic acid
13.979265	oxolactone	Tryptophan oxidation to oxolactone
13.979265	Thr->Asp	Thr->Asp substitution
13.979265	Carbonyl	aldehyde and ketone modifications
13.06802	Asp->Lys	Asp->Lys substitution
13.031634	MethylamineST	Michael addition with methylamine
13.031634	Asp->Gln	Asp->Gln substitution
12.995249	Thr->Asn	Thr->Asn substitution
12.995249	Nitrene	Loss of O2; nitro photochemical decomposition
12.962234	Phe->CamCys	Phe->Cys substitution and carbamidomethylation
12.036386	Thr->Xle	Thr->Leu/Ile substitution
12.036386	Ser->Val	Ser->Val substitution
12.017759	Cys->Asp	Cys->Asp substitution
12	Thiazolidine	formaldehyde adduct
11.070128	Cys->ethylaminoAla	Carbamidomethylated Cys that undergoes beta-elimination and Michael addition of ethylamine
11.033743	Cys->Asn	Cys->Asn substitution
10.07488	Cys->Xle	Cys->Leu/Ile substitution
10.020735	Ser->Pro	Ser->Pro substitution
10.009502	His->Phe	His->Phe substitution
9.032697	Phe->Arg	Phe->Arg substitution
9.000334	Gln->His	Gln->His substitution
8.963949	Lys->His	Lys->His substitution
8.016319	Glu->His	Glu->His substitution

6.962218	Arg->Tyr	Arg->Tyr substitution
6.018427	Met->His	Met->His substitution
6.008178	Cation:Li	Replacement of proton by lithium
5.956421	Pro->Cys	Pro->Cys substitution
4.97893	Formylasparagine	In Bachi as Formylaspargine (typo?)
3.994915	kynurenin	tryptophan oxidation to kynurenin
3.994915	Pro->Thr	Pro->Thr substitution
3.940771	Val->Cys	Val->Cys substitution
2.981907	Gln->Met	Gln->Met substitution
2.945522	Lys->Met	Lys->Met substitution
2.01565	Pro->Val	Pro->Val substitution
1.997892	Glu->Met	Glu->Met substitution
1.979265	Val->Thr	Val->Thr substitution
1.961506	Thr->Cys	Thr->Cys substitution
1.942879	Xle->Asp	Leu/Ile->Asp substitution
0.984016	Deamidation	Deamidation
0.984016	Asn->Asp	Asn->Asp substitution
0.984016	Gln->Glu	Gln->Glu substitution
0.958863	Xle->Asn	Leu/Ile->Asn substitution
0.94763	Lys->Glu	Lys->Glu substitution
0.036386	Gln->Lys	Gln->Lys substitution
-0.036386	Lys->Gln	Lys->Gln substitution
-0.94763	Glu->Lys	Glu->Lys substitution
-0.958863	Asn->Xle	Asn->Leu/Ile substitution
-0.984016	Amide	Amidation
-0.984016	Asp->Asn	Asp->Asn substitution
-0.984016	Glu->Gln	Glu->Gln substitution
-1.007825	Cystine	Half of a disulfide bridge
-1.031634	Lysaminoadipicsealde	Lysine oxidation to aminoadipic semialdehyde
-1.942879	Asp->Xle	Asp->Leu/Ile substitution
-1.961506	Cys->Thr	Cys->Thr substitution
-1.979265	Thr->Val	Thr->Val substitution
-1.997892	Met->Glu	Met->Glu substitution
-2.01565	didehydro	2-amino-3-oxo-butanoic_acid
-2.01565	Val->Pro	Val->Pro substitution
-2.945522	Met->Lys	Met->Lys substitution
-2.945522	Cys->methylaminoAla	carbamidomethylated Cys that undergoes beta-elimination and Michael addition of methylamine
-2.981907	Met->Gln	Met->Gln substitution
-3.940771	Cys->Val	Cys->Val substitution
-3.994915	Thr->Pro	Thr->Pro substitution
-3.994915	Glu->pyro-Glu+Methyl	Pyro-Glu from E + Methylation

-4.986324	Met->Aha	Methionine replacement by azido homoalanine
-5.956421	Cys->Pro	Cys->Pro substitution
-6.018427	His->Met	His->Met substitution
-6.962218	Tyr->Arg	Tyr->Arg substitution
-8.016319	His->Glu	His->Glu substitution
-8.963949	His->Lys	His->Lys substitution
-9.000334	His->Gln	His->Gln substitution
-9.032697	Arg->Phe	Arg->Phe substitution
-10.009502	Phe->His	Phe->His substitution
-10.020735	Pro->Ser	Pro->Ser substitution
-10.031969	Aspartylurea	Aspartylurea
-10.07488	Xle->Cys	Leu/Ile->Cys substitution
-11.033743	Asn->Cys	Asn->Cys substitution
-12.017759	Asp->Cys	Asp->Cys substitution
-12.036386	Xle->Thr	Leu/Ile->Thr substitution
-12.036386	Val->Ser	Val->Ser substitution
-12.995249	Asn->Thr	Asn->Thr substitution
-13.031634	Gln->Asp	Gln->Asp substitution
-13.06802	Lys->Asp	Lys->Asp substitution
-13.979265	Asp->Thr	Asp->Thr substitution
-14.01565	Ala->Gly	Ala->Gly substitution
-14.01565	Glu->Asp	Glu->Asp substitution
-14.01565	Xle->Val	Leu/Ile->Val substitution
-14.01565	Thr->Ser	Thr->Ser substitution
-14.01565	Gln->Asn	Gln->Asn substitution
-14.052036	Lys->Asn	Lys->Asn substitution
-14.974514	Gln->Xle	Gln->Leu/Ile substitution
-14.974514	Asn->Val	Asn->Val substitution
-14.999666	Glu->Asn	Glu->Asn substitution
-15.010899	lactic	lactic acid from N-term Ser
-15.010899	Lys->Xle	Lys->Leu/Ile substitution
-15.010899	ISD_z+2_ion	ISD (z+2)-series
-15.958529	Asp->Val	Asp->Val substitution
-15.958529	Glu->Xle	Glu->Leu/Ile substitution
-15.977156	Cys->Ser	Cys->Ser substitution
-15.994915	semialdehyde	reduction
-15.994915	Ser->Ala	Ser->Ala substitution
-15.994915	Tyr->Phe	Tyr->Phe substitution
-16.013542	Met->Asp	Met->Asp substitution
-16.027929	Phe->Met	Phe->Met substitution
-16.0313	Xle->Pro	Leu/Ile->Pro substitution
-16.990164	Asn->Pro	Asn->Pro substitution

-16.997557	Met->Asn	Met->Asn substitution
-17.026549	Pyro-glu	Pyro-glu from Q
-17.026549	N-oxobutanoic	Loss of ammonia
-17.956421	Met->Xle	Met->Leu/Ile substitution
-17.974179	Asp->Pro	Asp->Pro substitution
-17.992806	oxoalanine	oxoalanine
-18.010565	Phospho+PL	Dehydration
-18.010565	Pyro_glu	Pyro-glu from E
-18.025821	Phe->Glu	Phe->Glu substitution
-18.973451	Phe->Lys	Phe->Lys substitution
-19.009836	Phe->Gln	Phe->Gln substitution
-19.042199	Arg->His	Arg->His substitution
-20.026215	azole	Formation of five membered aromatic heterocycle
-21.987721	Met->Hpg	methionine replacement by homopropargylglycine
-22.031969	His2Asp	His->Asp substitution
-23.015984	His2Asn	His->Asn substitution
-23.015984	Trp->Tyr	Trp->Tyr substitution
-23.974848	His->Xle	His->Leu/Ile substitution
-25.049393	Gln->Cys	Gln->Cys substitution
-25.060626	Arg->Met	Arg->Met substitution
-25.085779	Lys->Cys	Lys->Cys substitution
-26.004417	Tyr->His	Tyr->His substitution
-26.01565	Pro->Ala	Pro->Ala substitution
-26.033409	Glu->Cys	Glu->Cys substitution
-26.052036	Xle->Ser	Leu/Ile->Ser substitution
-27.010899	Asn->Ser	Asn->Ser substitution
-27.010899	Gln->Thr	Gln->Thr substitution
-27.047285	Lys->Thr	Lys->Thr substitution
-27.058518	Arg->Glu	Arg->Glu substitution
-27.994915	OxPro	Pyrrolidone from Proline
-27.994915	Asp->Ser	Asp->Ser substitution
-27.994915	Glu->Thr	Glu->Thr substitution
-28.006148	Arg->Lys	Arg->Lys substitution
-28.0313	Val->Ala	Val->Ala substitution
-28.0313	Met->Cys	Met->Cys substitution
-28.042534	Arg->Gln	Arg->Gln substitution
-28.990164	Gln->Val	Gln->Val substitution
-29.026549	Lys->Val	Lys->Val substitution
-29.974179	Glu->Val	Glu->Val substitution
-29.978202	Trp->Arg	Trp->Arg substitution
-29.992806	HSe	Homoserine

-29.992806	Met->Thr	Met->Thr substitution
-30.010565	Pyrrolidinone	Proline oxidation to pyrrolidinone
-30.010565	Ser->Gly	Ser->Gly substitution
-30.010565	Thr->Ala	Thr->Ala substitution
-30.010565	Decarboxylation	Decarboxylation
-31.005814	Gln->Pro	Gln->Pro substitution
-31.042199	Lys->Pro	Lys->Pro substitution
-31.972071	Met->Val	Met->Val substitution
-31.972071	Cys->Ala	Cys->Ala substitution
-31.989829	Glu->Pro	Glu->Pro substitution
-32.008456	Met->AspSA	Methionine oxidation to aspartic semialdehyde
-32.022844	Tyr->Met	Tyr->Met substitution
-32.041471	Phe->Asp	Phe->Asp substitution
-33.003705	pyruvicC	pyruvic acid from N-term cys
-33.025486	Phe->Asn	Phe->Asn substitution
-33.98435	Phe->Xle	Phe->Leu/Ile substitution
-33.987721	DehydroalaC	Dehydroalanine (from Cysteine)
-33.987721	Met->Pro	Met->Pro substitution
-34.020735	Tyr->Glu	Tyr->Glu substitution
-34.049727	His->Cys	His->Cys substitution
-34.968366	Tyr->Lys	Tyr->Lys substitution
-35.004751	Tyr->Gln	Tyr->Gln substitution
-36.011233	His->Thr	His->Thr substitution
-37.990498	His->Val	His->Val substitution
-39.010899	Trp->Phe	Trp->Phe substitution
-40.006148	His->Pro	His->Pro substitution
-40.0313	Pro->Gly	Pro->Gly substitution
-41.026549	Gln->Ser	Gln->Ser substitution
-41.062935	Lys->Ser	Lys->Ser substitution
-41.074168	Arg->Asp	Arg->Asp substitution
-42.010565	Glu->Ser	Glu->Ser substitution
-42.021798	Arg2Orn	Ornithine from Arginine
-42.04695	Xle->Ala	Leu/Ile->Ala substitution
-42.04695	Val->Gly	Val->Gly substitution
-42.058184	Arg->Asn	Arg->Asn substitution
-43.005814	Asn->Ala	Asn->Ala substitution
-43.017047	Arg->Xle	Arg->Leu/Ile substitution
-43.053433	Argglutamicsealde	Arginine oxidation to glutamic semialdehyde
-43.989829	Asp->Ala	Asp->Ala substitution
-44.008456	Met->Ser	Met->Ser substitution
-44.026215	Thr->Gly	Thr->Gly substitution

-44.059229	Phe->Cys	Phe->Cys substitution
-45.987721	Cys->Gly	Cys->Gly substitution
-46.005479	a-type_Ion	ISD a-series (C-Term)
-46.020735	Phe->Thr	Phe->Thr substitution
-48	Phe->Val	Phe->Val substitution
-48.003371	Hse_lact	Homoserine lactone
-48.003371	Dethiomethyl	Prompt loss of side chain from oxidised Met
-48.036386	Tyr->Asp	Tyr->Asp substitution
-49.020401	Tyr->Asn	Tyr->Asn substitution
-49.020401	Trp->His	Trp->His substitution
-49.979265	Tyr->Xle	Tyr->Leu/Ile substitution
-50.01565	Phe->Pro	Phe->Pro substitution
-50.026883	His->Ser	His->Ser substitution
-53.091927	Arg->Cys	Arg->Cys substitution
-55.038828	Trp->Met	Trp->Met substitution
-55.053433	Arg->Thr	Arg->Thr substitution
-56.0626	Xle->Gly	Leu/Ile->Gly substitution
-57.021464	Asn->Gly	Asn->Gly substitution
-57.021464	Gln->Ala	Gln->Ala substitution
-57.032697	Arg->Val	Arg->Val substitution
-57.03672	Trp->Glu	Trp->Glu substitution
-57.057849	Lys->Ala	Lys->Ala substitution
-57.98435	Trp->Lys	Trp->Lys substitution
-58.005479	Asp->Gly	Asp->Gly substitution
-58.005479	Glu->Ala	Glu->Ala substitution
-58.005479	Gly-loss+Amide	Enzymatic glycine removal leaving an amidated C-terminus
-58.020735	Trp->Gln	Trp->Gln substitution
-59.048347	Arg->Pro	Arg->Pro substitution
-60.003371	Met->Ala	Met->Ala substitution
-60.036386	Phe->Ser	Phe->Ser substitution
-60.054144	Tyr->Cys	Tyr->Cys substitution
-62.01565	Tyr->Thr	Tyr->Thr substitution
-63.994915	Tyr->Val	Tyr->Val substitution
-66.010565	Tyr->Pro	Tyr->Pro substitution
-66.021798	His->Ala	His->Ala substitution
-69.069083	Arg->Ser	Arg->Ser substitution
-71.037114	Gln->Gly	Gln->Gly substitution
-71.05237	Trp->Asp	Trp->Asp substitution
-71.073499	Lys->Gly	Lys->Gly substitution
-72.021129	Glu->Gly	Glu->Gly substitution
-72.036386	Trp->Asn	Trp->Asn substitution

-72.995249	Trp->Xle	Trp->Leu/Ile substitution
-74.019021	Met->Gly	Met->Gly substitution
-76.0313	Tyr->Ser	Tyr->Ser substitution
-76.0313	Phe->Ala	Phe->Ala substitution
-80.037448	His->Gly	His->Gly substitution
-83.070128	Trp->Cys	Trp->Cys substitution
-85.031634	Trp->Thr	Trp->Thr substitution
-85.063997	Arg->Ala	Arg->Ala substitution
-87.010899	Trp->Val	Trp->Val substitution
-89.026549	Trp->Pro	Trp->Pro substitution
-89.02992	Met-loss+Acetyl	Removal of initiator methionine from protein N-terminus, then acetylation of the new N-terminus
-90.04695	Phe->Gly	Phe->Gly substitution
-92.026215	Tyr->Ala	Tyr->Ala substitution
-94.041865	DehydroalaY	Dehydroalanine (from Tyrosine)
-99.047285	Trp->Ser	Trp->Ser substitution
-99.079647	Arg->Gly	Arg->Gly substitution
-106.041865	Tyr->Gly	Tyr->Gly substitution
-115.042199	Trp->Ala	Trp->Ala substitution
-129.057849	Trp->Gly	Trp->Gly substitution
-131.040485	Met-loss	Removal of initiator methionine from protein N-terminus
-156.101111	Arg-loss	Loss of arginine due to transpeptidation

Supplementary List 5. User file for Sticker.

d	Lab_User_label	Cond_a
	NoLab	
-458.325864	TMT-derived(-TMT(2))	N-term,K
-442.330949	TMT-derived(-TMT(2)+O)	N-term,K
-422.325864	K-1-amino-3-iminopropene-K crosslink+2TMT-loss	K
-421.341848	K-1-amino-3-iminopropene-K crosslink+2TMT-loss+DEA	K
-406.330949	K-1-amino-3-iminopropene-K crosslink+2TMT-loss+OX	K
-404.315299	N-propenal+2-TMT-loss	K
-403.331283	N-propenal+2-TMT-loss+DEA	K
-388.320384	N-propenal+2-TMT-loss+OX	K
-376.312384	TMT-derived	N-term,K
-360.330488	TMT-derived	C or Q on N-term, K
-359.333453	TMT-derived	C or Q on N-term, K
-350.304734	2N-propenal +2TMT-loss	K
-344.276762	TMT-derived	N-term,K

-320.172118	Cys-> Dha -TMT	C
-246.189481	TMT-derived	C or Q on N-term
-230.194566	TMT-derived	C or Q on N-term
-229.162932	TMT-derived	N-term,K
-228.282014	TMT-derived	N-term,K
-217.162932	TMT-derived ('-TMT+C)	N-term
-214.199652	Lys->AminoadipicAcid	K
-213.168017	TMT-derived (O-TMT)	N-term
-203.18058	TMT-derived	N-term,K
-201.168017	Formyl	K
-201.168017	TMT-derived (formyl-TMT)	N-term
-197.173102	TMT-derived (O(2)-TMT)	N-term,K
-187.185665	TMT-derived	N-term,K,M
-187.152367	Acetyl	K
-186.154031	TMT-derived	N-term,K
-175.152367	N-propenal+1TMT-loss	K
-174.168351	N-propenal+1TMT-loss+DEA	K
-172.138381	TMT-derived	N-term,K
-170.159116	TMT-derived	N-term,K,M
-159.157452	N-propenal+1TMT-loss+OX	K
-157.133802	TMT-derived	N-term,K
-156.143466	TMT-derived	N-term,K,M
-147.149452	TMT-derived	N-term,K
-143.162537	N-propenal+1TMT-loss+2OX	K
-143.126153	2-hydroxyisobutyrylation	K
-141.138887	TMT-derived	N-term M
-129.138887	TMT-derived	N-term,K
-121.141803	HydroxymethylOP	K
-115.120005	GG	K
-114.141007	TMT-derived	N-term,K
-113.143972	TMT-derived	N-term,K
-112.138097	HCysThiolactone	K
-103.009185	Cys->Gly-IAM	C
-98.146092	TMT-derived	N-term,K
-97.14905703	TMT-derived	N-term,K
-95.14234703	TMT-derived	N-term
-94.13899203	TMT-derived	N-term
-91.009186	Cys-> Dha	C
-90.025169	Cys->PyruvicAcid-IAM	C
-88.993535	Cys->Ala-IAM	C
-82.151177	TMT-derived	N-term,K
-75.014270	Cys->Oxoalanine-IAM	C

-72.998620	Cys->Ser-IAM	C
-72.11742303	TMT-derived	N-term,K
-67.110108	Hex	K
-62.977885	Cys->Pro-IAM	C
-60.962235	Cys->Val-IAM	C
-59.037114	S-S->Thioaldehyde	C
-58.982970	Cys->Thr-IAM	C
-56.12250803	TMT-derived	N-term,K
-46.946584	Cys->Xle-IAM	C
-45.987721	Cys->Asn-IAM	C
-45.003705	C->D-IAM	C
-33.987721	C_reduced (loss of oxidized, alkylated cysteine)	C
-31.972071	Cys->Gln-IAM	C
-31.935685	Cys->Lys-IAM	C
-30.988055	Cys->Glu-IAM	C
-28.990164	Cys->Met-IAM	C
-25.031634	C_sulfinic	C
-22.971737	Cys->His-IAM	C
-22.067486	2DTT-sensitive C ox	C
-12.995249	Cys->MetOx-IAM	C
-12.962235	Cys->Phe-IAM	C
-11.033743	DTT-sensitive C ox	C
-9.036719	C_sulfonic	C
-3.929537	Cys->Arg-IAM	C
-2.015650	S-S->Thioaldehyde; water-loss +ox	C
-1.031634	ammonia-loss+ox	NA
-1.012295	water-loss +ox+C13	NA
-0.028279	ammonia-loss+ox+C13	NA
3.032680	Cys->Tyr-IAM	C
4.978931	KYN+DEA	W
26.048664	Cys->Trp-IAM	C
36.000000	K-1-amino-3-iminopropene-K crosslink	K
49.955943	Cl+O	W,Y
54.010565	N-propenal	K
57.024551	TMT-derived	S,H,T,K,M
61.982636	MMTS+O	C
71.032095	TMT-derived	N-term,K
72.02912997	TMT-derived	S,H,W,Y,M
73.016379	IAM+O	C
73.019466	TMT-derived	M
89.014381	TMT-derived	M
91.975442	MMTS(2)	C

100.024045	TMT-derived	S,T
152.980587	C_disulfur (CarbamidomethylDTT)+DEA	C
154.065171	LRGG	K
168.080821	K_LRGG+methyl	K
182.096471	K_LRGG+dimethyl	K
197.227247	Cys->Glu+TMT-IAM	C
209.018035	DTT+IAM	C
230.146948	TMT-derived	N-term,K
244.173831	NH+TMT	Y
245.157847	TMT-derived	N-term,K
246.189481	TMT-derived	N-term,K
261.152762	TMT-derived	N-term,K
261.152762	TMT-derived	N-term,K
286.184396	TMT-derived	N-term,K
297.084987	(+)TMT+CL(2)	Y
297.084987	TMT-derived	Y
298.069003	(+)TMT+CL(2)+DEA	Y
458.325864	TMT-derived	S,T