Glossary of proteomics terms and abbreviations

The IUPAC maintains a glossary of mass spectrometry terms here.

Term	Definition
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis. An electrophoretic technique for separating proteins according to two physicochemical properties - isoelectric point and protein mass.
AP-MS	Affinity purification-mass spectrometry. The coupling of affinity purification to MS analysis is used to identify and quantify protein-protein interactions.
APEX	Engineered ascorbate peroxidase. A genetically targetable tag employed for protein labelling. APEX is used to biotinylate proteins nearby a protein of interest. Subsequent affinity purification-mass spectrometry is carried out to analyse protein-protein interactions and map subcellular structure.
AQUA	A strategy for absolute quantification of proteins. Synthetic peptides incorporating a stable isotope labelled amino acid are used as internal standards for targeted proteomics.
Base peak chromatogram	A chromatogram in which the signal intensity of the most intense peak (base peak) in a series of mass spectra is plotted against retention time.
BCA assay	Bicinchoninic acid assay. A colourimetric assay for quantifying the total concentration of protein in a solution.
Biological process	A gene ontology category containing terms that describe series of molecular events with a defined beginning and end. For example, metabolic processes.
Bottom-up proteomics	One of the two main strategies for protein analysis within proteomics (contrast with Top-down proteomics). Bottom-up approaches involve the proteolytic digestion of protein samples prior to MS analysis. Peptide ions are fragmented and analysed by tandem mass spectrometry. Peptides are typically identified by comparing the MS2 spectra to theoretical MS2 spectra generated by *in silico* digestion of a proteome.
C18	Octadecylsilane. Commonly used as the non-polar stationary phase for RPLC.
Cellular compartment	A gene ontology category containing terms that describe gene product location. For example, localisation to subcellular structures or macromolecular complexes.

Term	Definition
CID	Collision-induced dissociation. A MS fragmentation technique in which precursor ions are accelerated in an electric field and allowed to collide with molecules of an inert gas.
Coselection	The isolation and fragmentation of ions of similar m/z to a selected precursor ion due to their coelution during MS/MS. Coselection negatively impacts upon protein quantification accuracy in MS2-based methods.
CRAPome	Contaminant repository for affinity purification. A database of common mass spectrometry contaminants.
Data-dependent acquisition (DDA)	A tandem mass spectrometry data collection mode in which precursor ions are selected from survey scans for MS/MS analysis based on predefined rules, such as peak intensity.
Data-independent acquisition (DIA)	A tandem mass spectrometry data collection mode in which all precursor ions in a defined m/z range are selected for MS/MS analysis. For example, SWATH-MS.
Deamidation	A chemical reaction in which an amide group is removed by hydrolysis. During proteomics sample preparation, glutamine and asparagine amino acids can be deamidated to glutamate and aspartate, respectively.
DIGE	Difference gel electrophoresis. An electrophoretic technique for determining differences in protein expression between samples. Protein samples are labelled with spectrally-resolvable fluorescent dyes, combined and separated by 2D-PAGE. Proteins from each sample are visualized using the different fluorophores and relative expression of protein spots determined. Protein spots may be excised and analysed by MS for protein identification.
Dimethyl labelling	A stable isotope labelling method for quantitative proteomics which uses a dimethyl group to label peptide primary amines. Quantification is carried out on survey scans.
Discovery proteomics	Discovery or shotgun proteomics strategies are designed to sample and analyse a large number of proteins in a complex mixture using bottom-up proteomics methods. Contrast with Targeted proteomics.
DTT	Dithiothreitol. A redox reagent commonly used to reduce protein disulphide bonds.

Term	Definition
Dynamic exclusion	A mass spectrometry feature that allows the mass spectrometer to stop selecting a precursor ion for fragmentation once it has been selected a specified number of times. This promotes the analysis of less abundant ions.
emPAI	Exponentially modified protein abundance index. A label-free quantitative proteomics method. Protein coverage is determined using MS2 spectra for peptide identification and used to estimate protein abundance.
ERLIC	Electrostatic repulsion hydrophilic interaction chromatography. A type of hydrophilic interaction chromatography (HILIC) on which ion-exchange chromatography is superimposed by using an ionic column surface chemistry. ERLIC is employed for phosphopeptide enrichment.
ESI	Electrospray ionization. A 'soft' ionization technique for mass spectrometry that produces gaseous ions by applying an electric field to a liquid analyte eluting from a narrow capillary tube.
ETD	Electron-transfer dissociation. A MS fragmentation technique in which electron transfer from radical anions to multiply-charged precursor cations induces fragmentation of the cations.
Expression proteomics	The large-scale analysis of the relative or absolute abundances of proteins. For example, to identify proteins that are differentially expressed in response to a drug.
FASP	Filter-aided sample preparation. A method for preparing peptide samples for proteomics analysis. Samples are solubilized in sodium dodecyl sulphate, which is subsequently exchanged for urea using an ultrafiltration device.
FDR	False discovery rate. In proteomics, FDR is a statistical metric for the rate of false positives in peptide spectral matches (PSMs) accepted for protein identification. This can be estimated by adding decoy sequences to a database for sequence database searching.
GeLC-MS/MS	A strategy for MS analysis of complex protein mixtures. Proteins are separated by SDS-PAGE, digested in-gel and analysed by liquid chromatography-tandem mass spectrometry.
Gene ontology (GO)	Hierarchical, defined terms for gene product properties. GO terms are categorised into three domains - cellular component, biological process and molecular function.
GOrilla	Gene ontology enrichment analysis and visualization tool. A tool for identifying enriched GO terms.
HCD	Higher-energy collisional dissociation. A CID technique in a hybrid linear ion trap-orbitrap mass spectrometer in which fragmentation occurs externally to the ion trap in a dedicated collision cell. Consequently, HCD fragmentation does not suffer from loss of low m/z fragment ions.

Term	Definition
Heat map	A diagram used to display data values represented as colours.
Hierarchical cluster analysis	A statistical technique for arranging objects into groups based on similarity. Objects are either sequentially split (divisive) or combined (agglomerative) into groups as the hierarchy progresses along a dendrogram. This method may be used to visualize protein expression data.
HILIC	Hydrophilic interaction chromatography. A chromatographic technique for separating mixtures using a hydrophilic stationary phase and a hydrophobic mobile phase. Solutes are eluted from the column in order of increasing polarity. HILIC can be used for the enrichment of post-translationally modified peptides, such as phosphorylated and glycosylated peptides.
HPLC	High-performance liquid chromatography. The chromatographic separation of compounds in solution at high pressure, in narrow columns packed with small sorbent particles. This provides higher resolution than traditional low-pressure liquid chromatography.
Human protein atlas (HPA)	An antibody-based proteomics resource containing expression and localisation profiles of human proteins in different tissues, cancer types and cell lines.
HyperLOPIT	Hyperplexed LOPIT. A proteomics method for the high-throughput characterization of protein subcellular localisation. HyperLOPIT is an extension of the LOPIT method, with improved multiplexing, MS data acquisition (SPS-MS^3) and multivariate data analysis.
ICAT	Isotope-coded affinity tag. Thio-reactive isotopic tagging reagents for MS-based protein quantification of cysteine-containing peptides.
IMAC	Immobilized metal affinity chromatography. A technique for enriching for phosphopeptides or histidine-containing peptides/proteins. This is based on the coordination of phosphate groups or histidine residues to immobilized metal ions. The choice of metal ion and chelating ligand depends upon the application. Fe^{3+}, Ga^{3+} or Ti^{4+} are commonly used for phosphopeptide enrichment.
In-gel digestion	The digestion of protein samples within a polyacrylamide matrix, following protein separation by polyacrylamide gel electrophoresis. This sample preparation method permits efficient solubilization of proteins using a sodium dodecyl sulphate-containing buffer, since low molecular weight components are removed during gel electrophoresis.

Term	Definition
In-solution digestion	A sample preparation method in which proteins are solubilized and proteolytically digested in a solution (contrast with In-gel digestion). In-solution digestion methods are typically detergent-free and rely on high concentrations of chaotropes, such as urea, to denature and solubilize proteins.
IntAct	A database of molecular interactions containing data with direct experimental evidence.
Iodoacetamide (IAA)	An alkylating agent commonly used to covalently bind the reduced thiol of cysteine residues and prevent disulphide bond formation.
iPAC	Interactomes by parallel affinity capture. An affinity purification-mass spectrometry method which uses dual tags for the affinity purification of a protein of interest in two independent experiments. Genuine binding partners are identified by their consistent identification in both affinity purification experiments.
Isobaric tagging	A quantitative proteomics technique in which peptides are chemically labelled with tags of the same nominal mass. Fragmentation during tandem mass spectrometry releases reporter ions of a $m{:}z$ characteristic to each tag. By differentially labelling samples with isobaric tagging reagents, one can quantify the relative abundances of proteins between samples. Isobaric tags include TMT and iTRAQ.
Isoelectric focusing	An electrophoretic technique for separating molecules by their isoelectric point (pI). For example, proteins can be separated by isoelectric point in 2D-PAGE. This involves the application of an electric field to samples in a fixed pH gradient.
Isotopologues	Molecules with the same mass number, but different isotope compositions.
iTRAQ	Isobaric tag for relative and absolute quantitation. Amine-reactive isobaric tagging reagents for protein quantification by tandem mass spectrometry.
KEGG	Kyoto encyclopedia of genes and genomes. A database of molecular functions (of genes and proteins) integrated into networks, for example signalling pathways.
Label-free quantification	Any MS method for protein quantification that does not rely on stable isotope labelling. Label-free methods include summing precursor ion signal intensity and spectral counting.
LC-MS/MS	Liquid chromatography tandem mass spectrometry. The coupling of a liquid chromatography system to a tandem mass spectrometer is used to reduce sample complexity before MS analysis.
LCM	Laser capture microdissection. A method for isolating specific regions/cells from tissue sections viewed under microscopy.

Term	Definition
LOPIT	Localisation of organelle proteins by isotope tagging. A proteomics method for the high-throughput characterization of protein subcellular localisation. Proteins are partially resolved according to subcellular location using density gradient centrifugation. Fractions are proteolytically digested and differentially labelled with isobaric tags. Relative abundances of proteins between fractions are determined using reporter ion profiles acquired via tandem mass spectrometry. Subcellular localisation is calculated using classification algorithms that assign proteins based on the profiles of organelle marker proteins.
Lys-C	Protease that hydrolyses peptide bonds at the C-terminal side of lysine residues. Used for sample preparation in bottom-up methods. Unlike trypsin, Lys-C cleaves Lys-Pro bonds.
$m{:}z$	Mass-to-charge ratio. The quantity formed by dividing the mass number of an ion by its charge number.
MALDI	Matrix-assisted laser desorption/ionization. A 'soft' ionization technique for mass spectrometry that produces gaseous ions by pulsed laser irradiation of an analyte embedded in an organic matrix.
Mascot	A search engine that uses mass spectrometry data to identify proteins from peptide sequence databases.
Mass defect	The difference between the exact mass of an atom or molecule and its nominal mass (the sum of the masses of its components).
Mass spectrum	A plot of relative abundance of ions (y axis) against their m/z (x axis).
Middle-down proteomics	A strategy for protein analysis within proteomics that is a hybrid of top-down and bottom-up approaches. Middle-down methods employ restricted proteolysis to generate larger peptides than bottom-up methods.
MOAC	Metal oxide affinity chromatography. A chromatographic technique for enriching phosphopeptides based on the affinity of phosphate groups for metal oxides. For example, TiO_2.
Molecular function	A gene ontology category containing terms that describe the molecular activities of gene products. For example, catalytic or binding activities.
MRM	Multiple reaction monitoring. The application of SRM to multiple fragment ions produced from one or more precursor ions.
MS	Mass spectrometry. A technique for analysing ions based on their mass-to-charge ratio. Mass spectrometers consist of three core components - an ion source, one or more mass analysers and a detector. The ion source generates gaseous ions from an analyte, the mass analyser/s separate ions in space or time and the detector system records the relative abundances of ions resolved by m/z.

Term	Definition
MS/MS	Tandem mass spectrometry. The acquisition and analysis of mass spectra collected via more than one stage of mass analysis. Precursor ions are fragmented prior to the second stage of MS analysis.
MS^n	Multistage mass spectrometry, where n is the number of stages of mass analysis. For example, MS/MS is also known as MS^2.
MS^E	A data-independent acquisition (DIA) method for tandem mass spectrometry in which the mass spectrometer alternates between acquiring mass spectra from low energy collisions and high energy collisions to obtain accurate masses of precursor ions and fragment ions, respectively.
MS1	Mass spectrometry survey scan. In data-dependent MS/MS methods, mass spectra are collected on precursor ions over a broad range of m/z values. These spectra are known as full scans or survey scans.
MS2	Mass spectrometry fragment scan. A mass spectrum acquired on fragment ions during tandem mass spectrometry.
NeuCode SILAC	Neutron encoding SILAC. A quantitative proteomics method based on SILAC. NeuCode SILAC exploits the mass defects of stable isotopes to achieve higher multiplexing than SILAC. Peptides with the same number of extra neutrons incorporated into different atoms have subtly different masses due to differences in nuclear binding energies. For example, swapping a ^{12}C atom for a ^{13}C atom along with a ^{15}N atom for a ^{14}N atom produces labels differing by 6 mDa. This method requires high-resolution mass spectrometers to resolve peaks of similar m/z.
NHS	N-hydroxysuccinimide. NHS esters are commonly used to modify peptides for quantitative proteomics. For example, tandem mass tagging employs reagents containing amine-reactive NHS esters to label peptides with isobaric tags for protein quantification.
PCA	Principal components analysis. A multivariate data analysis technique used to summarize variation in a dataset and identify underlying patterns. A group of correlated variables are transformed into linearly uncorrelated variables, termed principal components, which summarize the variation in the dataset. Condensing complex data into principal components allows one to identify patterns in the dataset by plotting principal component values in 2-dimensional or 3-dimensional plots.
Peptide mass fingerprinting	A method for protein identification by mass spectrometry in which the molecular masses of unfragmented peptide ions are matched to theoretical peptide masses.

Term	Definition
Peptidotypic	Peptidotypic fragment ions are specific to a particular peptide sequence (analogous to proteotypic peptides).
Phosphoproteomics	The branch of proteomics concerned with analysis of phosphorylated proteins and peptides.
PPI	Protein-protein interaction. The physical association of two or more proteins.
PRM	Parallel reaction monitoring. A targeted proteomics method similar to SRM. Specific peptides are selected by m/z in the first stage of a tandem mass spectrometer and fragmented. Unlike SRM, parallel reaction monitoring detects all target fragment ions in one, concerted MS analysis using a high-resolution mass analyser.
Protein correlation profiling	A spatial proteomics method for determining the subcellular localisation of proteins. Proteins are partially resolved according to subcellular location via density gradient centrifugation. Fractions are analysed by label-free MS/MS to construct protein relative abundance profiles spanning the density gradient. Subcellular localisation is determined by matching the profiles of proteins with unknown localisation to those of well-characterized marker proteins for specific locations.
Proteoform	The specific molecular form of a protein product arising from a single gene. This distinguishes proteins arising from a particular gene that differ due to genetic variation, alternative splicing and post-translational modification.
Proteome	The complete set of proteins expressed in a cell, tissue or biological system. Proteomes vary with time and biological context.
Proteomics	The large-scale study of proteomes. This term incorporates the large-scale investigation of protein expression, turnover, post-translational modification, localisation and interactions between proteins.
Proteotypic peptides	Peptides whose sequence is unique to a specific protein. Proteotypic peptides are used for protein identification.
PSM	Peptide spectrum match. A peptide hit identified by matching an acquired MS2 spectrum to an MS2 spectrum in a sequence database (for example, containing theoretical MS2 spectra generated by *in silico* digestion of a proteome).
PTM	Post-translational modification. The diverse processing events applied to proteins following their translation. PTM most commonly refers to the covalent attachment of a functional group to a protein. For example, phosphorylation or acetylation.

Term	Definition
QconCAT	A method for absolute quantification of target proteins. A synthetic gene is designed to express heavy isotope labelled concatemers of proteotypic peptides for the protein/s of interest. These labelled polypeptides are used as internal standards for targeted proteomics.
Quantitative proteomics	The systematic analysis of relative or absolute abundance of proteins.
RPLC	Reversed-phase liquid chromatography. A chromatographic technique for separating mixtures using a non-polar stationary phase and polar mobile phase.
RPPA	Reverse phase protein array. An antibody-based proteomics technology for analysing protein expression or post-translational modification in a large number of samples. Protein lysates are printed onto microarray slides and probed with antibodies to specific proteins.
S/N	Signal-to-noise ratio. In mass spectrometry, a measure of instrument sensitivity. The maximum signal intensity of an analyte is divided by noise calculated as the standard deviation or root mean square of the baseline surrounding the signal peak.
SCX	Strong cation exchange chromatography. A technique for separating mixtures by net surface charge. A negatively charged stationary phase binds positively charged analytes in a mobile phase. SCX can be used to enrich phosphopeptides from a proteolytic digest for phosphoproteomics analysis. At low pH, tryptic peptides typically carry a net charge of +2 or higher. In contrast, phosphorylated peptides carry a lower net charge (0/+1).
SDS	Sodium dodecyl sulphate. An anionic surfactant used to solubilize protein samples.
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis. An electrophoretic technique for separating proteins according to their molecular mass. Proteins are denatured by the anionic detergent sodium dodecyl sulphate (SDS). SDS binds to polypeptides in proportion to their molecular mass, imparting a uniform charge density (negative charge) to polypeptides. Samples are separated in a polyacrylamide gel matrix by applying an electric field. Polypeptides migrate towards the anode as a linear function of their molecular mass. Mass-resolved protein samples are commonly analysed by Western blotting or non-specific protein staining.
Secretome	The entire set of proteins secreted by a cell, tissue or organism.
Shotgun proteomics	See Discovery proteomics.

Term	Definition
SILAC	Stable isotope labelling by amino acids in cell culture (SILAC) is a MS-based quantitative proteomics technique in which proteins are labelled via the incorporation of stable isotopic forms of amino acids during cell growth.
SIMAC	Sequential elution from IMAC. A phosphoproteomics strategy for separating monophosphorylated peptides from multiply phosphorylated peptides. Monophosphorylated and multiply phosphorylated peptides are eluted from IMAC under acidic or basic conditions, respectively. Monophosphorylated peptides are subsequently enriched from unphosphorylated peptides using TiO_2 chromatography.
Soft ionization	Ionization techniques that produce gas-phase ions with minimal fragmentation of the analyte. Examples include ESI and MALDI.
Spectral counting	A semiquantitative label-free proteomics method in which the total number of MS2 spectra derived from a protein is used as a measure of protein abundance.
SPS-MS^3	Synchronous precursor selection-MS^3. A tandem mass spectrometry method for data acquisition, which uses isolation waveforms with multiple frequency notches to isolate and fragment multiple MS2 product ions simultaneously. SPS-MS^3 is employed for quantitation via isobaric labelling methods, such as TMT. SPS-MS^3 reduces ratio distortion caused by coselection, while concomitantly increasing the number of reporter ions for quantitation in MS3 spectra over the standard MS^3 method.
SRM	Selected reaction monitoring. A targeted proteomics method typically performed on a triple quadrupole mass spectrometer. A specific peptide precursor ion is selected by m/z in the first mass analyser, fragmented in a collision cell and a specific fragment ion selected by m/z in the second mass analyser. These precursor/fragment ion pairs, termed transitions, are used to selectively monitor target proteins.
SRMAtlas	A database of targeted proteomics assays (SRM/MRM).
Static modification	A type of modification that can be specified when carrying out peptide identification by sequence database searching. Static or fixed modifications are applied to every instance of the specified amino acid or terminus and do not increase the computational complexity of a search.
STRING	Search tool for the retrieval of interacting genes/proteins. A database of known and predicted protein-protein interactions.
Subproteome	A subdivision of a proteome. For example, the protein complement of an organelle.
SWATH-MS	In this data-independent acquisition MS method, all precursor ions in a predefined m/z window are selected for MS/MS concurrently. The mass spectrometer repeatedly cycles through consecutive precursor isolation windows to acquire convoluted fragment ion spectra for all detectable analytes in a wide m/z range.

Term	Definition
TAP	Tandem affinity purification. A common method for affinity purification that utilizes a dual tag to sequentially purify a fusion protein. In the original method, a TAP tag is fused to the C-terminus of a protein. This tag consists of a calmodulin-binding peptide, TEV protease cleavage site and Protein A (N-terminal to C-terminal). The TAP-tagged protein is affinity purified using an IgG matrix, cleaved with TEV protease and further purified using calmodulin beads.
Targeted proteomics	Targeted proteomics methods analyse predefined subsets of proteins at high sensitivity using mass spectrometry. Examples include selected reaction monitoring (SRM). Contrast with Discovery proteomics methods.
TCA	Trichloroacetic acid. Used for protein precipitation.
TCEP	Tris-(2-carboxyethyl)phosphine. A reducing agent commonly used to reduce protein disulphide bonds.
TiO_2	Titanium dioxide. Used for phosphopeptide enrichment by metal oxide affinity chromatography (MOAC).
TMT	Tandem mass tag. Amine-reactive isobaric tagging reagents for protein quantification by tandem mass spectrometry.
Top-down proteomics	One of the two main strategies for protein analysis within proteomics (contrast with Bottom-up proteomics). Intact proteins are ionized and analysed by tandem mass spectrometry. Fragment ions are used for protein identification.
Total ion current (TIC)	The summed signal intensity of all ions in a mass spectrum or in a specified m/z range of a mass spectrum.
Total ion current chromatogram	A chromatogram in which the total ion current in a series of mass spectra is plotted against retention time.
Trypsin	Protease that hydrolyses peptide bonds at the C-terminal side of lysine and arginine residues. Used for sample preparation in bottom-up methods. May be used in combination with alternative proteases, such as Lys-C, to increase sequence coverage.
Variable modification	A type of modification that can be specified when carrying out peptide identification by sequence database searching. Variable or dynamic modifications are those which are sometimes present on a specific amino acid or terminus. For example, methionine oxidation, which causes a mass shift of 16 Da for a peptide containing a single oxidized methionine. Variable modifications increase the computational complexity of a search.

Term	Definition
Western blotting	A technique for protein analysis. Protein samples separated by gel electrophoresis are transferred onto the surface of a membrane. The membrane is probed with a primary antibody that recognizes one or more epitopes specific to a particular protein. Primary antibodies are detected with species-specific secondary antibodies. Most commonly, secondary antibodies are conjugated to horseradish peroxidase (HRP), which generates luminescence when incubated with a cleavable chemiluminescence reagent. This signal is proportional to the amount of protein and is typically detected with photographic film or a CCD camera.
XIC	Extracted ion chromatogram. A chromatogram in which the signal intensity of one or more specified m/z values are plotted against retention time.