MIAPE: Mass spectrometry

Chris F Taylor[1], Pierre-Alain Binz[2,3]*, Ruedi Aebersold [4], Michel Affolter[5], Robert Barkovich[6], Eric W. Deutsch[7], David M. Horn[8], Andreas Hühmer[9], Martin Kussmann[5], Kathryn Lilley[10], Marcus Macht[11], Matthias Mann[12], Dieter Müller[13], Thomas A. Neubert[14], Janice Nickson[15], Scott D. Patterson[16], Roberto Raso[17], Kathryn Resing[18], Sean L. Seymour[19], Akira Tsugita[20], Ioannis Xenarios[21], Rong Zeng[22], Randall K. Julian, Jr. [23]

[1] European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK
[2] Swiss Institute of Bioinformatics, Proteome Informatics Group, Rue Michel-Servet 1, CH-1211 Geneva 4, Switzerland
[3] GeneBio SA, Av. de Champel 25, Geneva, Switzerland
[4] Institute for Molecular Systems Biology, ETH Zurich, HPT E 78, Wolfgang-Pauli-Str. 16, 8093 Zürich, Switzerland
[5] Nestlé Research Center, Nestec Ltd., Vers-chez-Les-Blanc, 1000 Lausanne 26, Switzerland
[6] Affimetrix, Inc., 3420 Central Expressway, Santa Clara, CA 95051, USA
[7] Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103, USA
[8] Agilent Technologies, 5301 Stevens Creek Blvd., Santa Clara, CA 95051, USA
[9] Thermo Fisher Scientific, 355 River Oaks Parkway, San Jose, CA 95134, USA
[10] Cambridge Centre for Proteomics, University of Cambridge, Cambridge, Cambridgeshire, CB2 1QW, UK
[11] Bruker Daltonik GmbH, Bremen, Germany
[12] Dept. Proteomics and Signal Transduction, Max-Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany
[13] Novartis Institutes for BioMedical Research, Genome and Proteome Sciences, Systems Biology, WSJ-088.702, CH-4056 Basel, Switzerland
[14] Skirball Institute of Biomolecular Medicine and Department of Pharmacology, New York University School of Medicine, New York, NY 10016, USA
[15] Pathways, DECS, AstraZeneca, Alderley Park, UK
[16] Amgen Inc., Molecular Sciences, One Amgen Center Drive MS 38-3-A, Thousand Oaks, CA, 91320-1799, USA
[17] Kratos Analytical (Shimadzu), Manchester, UK
[18] Dept. of Chemistry & Biochemistry, University of Colorado, Boulder, CO 80309-0215, USA.
[19] Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404, USA
[20] Proteomics Research Laboratory, Tokyo Rikakikai Co., Tsukuba, Japan
[21] Swiss Institute of Bioinformatics, Vital-IT group, Quartier Sorge – Bâtiment Génopole, CH-1015 Lausanne, Switzerland
[22] Institutes for Biological Sciences, Chinese Academy of Sciences, Graduate School of the Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China
[23] Indigo BioSystems, Inc., Indianapolis, IN, USA

* Corresponding author

Abstract
"MIAPE - Mass spectrometry" (MIAPE-MS) is one module of the Minimal Information About a Proteomics Experiment (MIAPE) documentation system. MIAPE is developed by the Proteomics Standards Initiative of the Human Proteome Organisation (HUPO-PSI). It aims at delivering a set of technical guidelines representing the minimal information required to report and sufficiently support assessment and interpretation of a proteomics experiment. This MIAPE-MS module is the result of a joint effort between the Mass Spectrometry group of HUPO-PSI and the proteomics community. It has been designed to specify a minimal set of information to document a mass spectrometry experiment. As for all MIAPE documents, these guidelines evolve and are made available on the PSI website at the url http://psidev.info.

MIAPE: Mass Spectrometry

Version 2.24, 28th May, 2008.

This module identifies the minimum information required to report the use of a mass spectrometer in a proteomics experiment, sufficient to support both the effective interpretation and assessment of the data and the potential recreation of the work that generated it.

Introduction

This document is one of a collection of technology-specific modules that together constitute the Minimum Information about a Proteomics Experiment (MIAPE) reporting guidelines produced by the Proteomics Standards Initiative. MIAPE is structured around a parent document that lays out the principles to which the individual reporting guidelines adhere. In brief, a MIAPE module represents the minimum information that should be reported about a data set or an experimental process, to allow a reader to interpret and critically evaluate the conclusions reached, and to support their experimental corroboration. In practice a MIAPE module comprises a checklist of information that should be provided (for example about the protocols employed) when a data set is submitted to a public repository or when an experimental step is reported in a scientific publication (for instance in the materials and methods section). The MIAPE modules specify neither the format that information should be transferred in, nor the structure of the repository/text. However, PSI is not developing the MIAPE modules in isolation; several compatible data exchange standards are now well established and supported by public databases and data processing software in proteomics (for details see the PSI website www.psidev.info).

The modern mass spectrometer is a rather complex instrument with many operational parameters; the data sets generated are similarly complex, and often rather voluminous. These guidelines for the reporting of mass spectrometry do not prescribe that all of that information be captured; and given the diversity of instruments currently available, the utility of such detail is clearly open to question.

However, it is possible to specify parameters that are representative of the way in which the mass spectrometer was used, to contextualise the data generated and thereby enable a better-informed process of assessment and interpretation.

These guidelines cover both the operation of a mass spectrometer and the generation of mass spectra from the 'raw' data. They do neither cover the delivery of sample to the mass spectrometer, nor the interpretation of spectra by search engines; these details are captured in separate MIAPE modules, the latest versions of which can be obtained from the HUPO Proteomics Standards Initiative website (http://psidev.info/). Note also that these guidelines do not cover all the available components of a mass spectrometer (for example, some of the less frequently used ion sources); subsequent versions of this document will have expanded coverage, as it will almost certainly be the case for all the MIAPE modules.

The following section, detailing the reporting guidelines for the use of a mass spectrometer, is subdivided as follows:

1. General features; summary statistics such as instrument manufacturer; the software used to run the machine and the parameters applied to it.
2. Ion sources; for example, matrix-assisted laser desorption ionisation (MALDI), electrospray ionisation.
3. All major components after the ion source; for example, ion traps, collision cells, time-of-flight tubes, detectors (including Fourier Transform Ion Cyclotron Resonance detection). Note that where a collision cell is an ion trap (including FT-ICR cells), the requirements for the relevant components should be combined.
4. The data resulting from the procedure; the method of generation of peak lists and the location of the raw data from which they were generated; the method by which quantitation

was performed (where appropriate) and the resulting quantitative data set.

Reporting guidelines for mass spectrometry

1. General features

a) Global descriptors
- Date stamp (as YYYY-MM-DD)
- Responsible person (or institutional role if more appropriate); provide name, affiliation and stable contact information
- Instrument manufacturer and model
- Customisations (summary)

b) Control and analysis software
- Software name and version
- Switching criteria (tandem only)
- Isolation width (global, or by MS level)
- Location of 'parameters' file

2. Ion sources

As each spectrum is acquired using only one ionisation source, select the one that applies

a) Electrospray Ionisation (ESI)
- Supply type (static, or fed)
- Interface manufacturer, model and catalog number (where available)
- Sprayer type, coating, manufacturer, model and catalogue number (where available)
- Relevant voltages where appropriate (tip, cone, acceleration)
- Other parameters if discriminant for the experiment (such as nebulising gas and pressure)

b) MALDI
- Plate composition (or type)
- Matrix composition (if applicable)
- Deposition technique
- Relevant voltages where appropriate (Grid , acceleration)
- PSD (or LID/ISD) summary, if performed
- Operation with or without delayed extraction
- Laser type (*e.g.* nitrogen) and wavelength (nm),
- Other laser related parameters, if discriminating for the experiment (such as pulse energy (μJ), attenuation, focus diameter (μm), pulse duration (ns at FWHM), frequency (Hz) and average shots fired per spectrum)

3. Post-source component

As a MS experiment performed on one instrument cannot be acquired using all existing analysers and detectors, select the elements that apply

a) Ion optics, 'simple' quadrupoles, hexapoles
- *No parameters to be captured*

b) Time-of-flight drift tube (TOF)
- Reflectron status (on, off, none)

c) Ion trap
- Final MS stage achieved

d) Collision cell
- Gas type and pressure (bar)
- Collision energy

e) FT-ICR
- *As for 'Ion trap' (3c) and 'Collision cell' (3d) combined, no further parameters required*

f) Detectors
- Detector type
- Detector sensitivity

4. Spectrum and peak list generation and annotation

For this section; if software other than that listed in 1b (Control and analysis software) is used to perform a task, the producer, name and version of that software must be supplied in each case

a) Spectrum description
- Location of source ('raw') file including file name and type
- Identifying information for the target area (MALDI-like methods only)
- MS level for this spectrum
- Ion mode for this spectrum
- Precursor *m/z* and charge, with the full mass spectrum containing that peak (for MS level 2 and higher)

b) Peak list generation
- Parameters triggering the generation of peak lists from raw data, including filtering for exclusion of peak lists from raw spectra, where appropriate
- Acquisition number (from the 'raw' file) of all acquisitions combined in the peak list, the total number combined and whether summed or averaged
- Smoothing; whether applied, parameters
- Background threshold, or algorithm used
- Signal-to-noise estimation and method

- Percentage peak height for centroiding; or algorithm used, if appropriate
- Whether charge states were calculated, spectra were deconvoluted and peaks were deisotoped (with methods described as appropriate)
- Relative times for all acquisitions combined in the peak list (electrospray only)
- Base peak *m/z*, where appropriate
- Metastable peaks removed, if applicable
- *m/z* and intensity values

c) Quantitation for selected ions (in addition to 4a) and 4b)
Only applicable if a quantitation experiment has been performed

- Experimental protocol, canonical reference where available with deviations
- Number of combined samples and MS runs analysed
- Quantitation approach (*e.g.* integration)
- Normalisation technique
- Location of quantitation data, with file name and type (where appropriate)

Summary

The MIAPE: MS minimum reporting guidelines for the use of a mass spectrometer specify that a significant degree of detail be captured, for mass spectrometry, spectral data and its subsequent processing. Providing the information required by this document will enable both the effective interpretation and assessment of mass spectral data and potentially, the recreation of the work that generated it. Much of the required information should be reusable from existing files, or exportable from the instrument; we anticipate further automation of this process.

These guidelines will evolve. To contribute, or to track the process to remain 'MIAPE-compliant', browse to the website at http://psidev.info.

Appendix One. The MIAPE: MS glossary of required-parameter classifications

Classification	*Definition*
1. General features — (a) Global descriptors	
Date stamp	The date on which the work described was initiated; given in the standard 'YYYY-MM-DD' format (with hyphens).
Responsible person or role (or institutional role if more appropriate); provide name, affiliation and stable contact information	The (stable) primary contact person for this data set; this could be the experimenter, lab head, line manager *etc*... Where responsibility rests with an institutional role (*e.g.* one of a number of duty officers) rather than a person, give the official name of the role rather than any one person. In all cases give affiliation and stable contact information. This information can be made available as part of an authors' list or in an acknowledgment section
Instrument manufacturer, model	The manufacturing company and model name for the mass spectrometer.
Customisations	Any significant (*i.e.* affecting behaviour) deviations from the manufacturer's specification for the mass spectrometer.
1. General features — (b) Control and analysis software	
Software name and version	The instrument management and data analysis package name, and version; where there are several pieces of software involved, give name, version and role for each one. Mention also upgrades not reflected in the version number.
Switching criteria (tandem only)	The list of conditions that cause the switch from survey or zoom mode (MS^1) to or tandem mode (MS^n where n > 1); e.g. 'parent ion' mass lists, neutral loss criteria and so on.
Isolation width (global, or by MS level)	For tandem instruments (*i.e.* multi-stage instruments such as triple quads and TOF-TOFs, plus ion traps and equivalents) the total width (*i.e.* not half for plus-or-minus) of the gate applied around a selected precursor ion m/z, provided for all levels or by MS level.
Location of 'parameters' file	The location and name under which the mass spectrometer's parameter settings file for the run is stored, if available. Ideally this should be a URI+filename, or most preferably an LSID, where feasible.
2. Ion sources — (a) Electrospray Ionisation (ESI)	
Supply type (static, or fed)	Whether the sprayer is fed (by, for example, chromatography or CE) or is loaded with sample once (before spraying).
Interface manufacturer, model and catalogue number (where available)	Where the interface was bought from, plus its name and catalogue number; list any modifications made to the standard specification. If the interface is entirely custom-built, describe it or provide a reference if available.
Sprayer type, coating, manufacturer, model and catalogue number (where available)	Where the sprayer was bought from, plus its name and catalogue number; list any modifications made to the standard specification. If the sprayer is entirely custom-built, describe it briefly or provide a reference if available.
Relevant voltages where appropriate (tip, cone, acceleration)	Voltages that are considered as discriminating from an understood standard measurement mode, or important for the interpretation of the data. These might include the voltage applied to the sprayer tip, the voltage applied to the sampling cone, the voltage used to accelerate the ions into the rest of the mass spectrometer (mass analysis + detection) by MS level.
Whether in-source dissociation performed	State whether in-source dissociation was performed (increased voltage between sample orifice and first skimmer).
Other parameters if discriminant for the experiment (such as nebulising gas and	Where appropriate, and if considered as discriminating elements of the source parameters, describe these values.

pressure)	
2. Ion sources — (b) MALDI	
Plate composition (or type)	The material of which the target plate is made (usually stainless steel, or coated glass); if the plate has a special construction then that should be briefly described and catalogue and lot numbers given where available.
Matrix composition (if applicable)	The material in which the sample is embedded on the target (*e.g.* alpha-cyano-4-hydroxycinnamic acid).
Deposition technique	The method of laying down (matrix and) sample on the target plate (including matrix concentration and solvents applied); for example, matrix+sample in single deposition; or matrix, then matrix+sample (if several matrix substances are used, name each); where chromatographic eluent is directly applied to the plate by apparatus, or for other approaches, describe the process and instrumentation involved very briefly and cross-reference.
Relevant voltages where appropriate	Voltages considered as relevant for the interpretation of the data. This might include the grid voltage (applied to the grid that sits just in front of the target), the acceleration voltage (used to accelerate the ions into the analyser part of the mass spectrometer (mass analysis + detection), etc.
PSD (or LID/ISD) summary, if performed	Confirm whether post-source decay, laser-induced decomposition, or in-source dissociation was performed; if so provide a brief description of the process (for example, summarise the stepwise reduction of reflector voltage).
Operation with or without delayed extraction	State whether a delay between laser shot and ion acceleration is employed.
Laser type (*e.g.* nitrogen) and wavelength (nm)	The type of laser and the wavelength of the generated pulse (in nanometres).
Other laser related parameters, if discriminating for the experiment (such as pulse energy (μJ), attenuation, focus diameter (μm), pulse duration (ns at FWHM), frequency (Hz) and average shots fired per spectrum	Other details of the laser used to shoot at the matrix-embedded sample if considered as important for the interpretation of data; this might include the pulse energy in micro Joules, focus diameter in microns, attenuation details, pulse duration in nanoseconds at full-width half maximum, frequency of shots in Hertz and average number of shots fired to generate each combined mass spectrum.
3. Post-source componentry — (a) Ion optics, 'simple' quadrupoles, hexapoles	
No parameters to be captured	These components (focusing elements and ion guides) require no description at present.
3. Post-source component — (b) TOF drift tube	
Reflectron status (on, off, none)	Whether a Reflectron is present, and if so, whether it is used.
3. Post-source component — (c) Ion trap	
Final MS stage achieved	The final MS level achieved in generating this data set with an ion trap or equivalent (*e.g.* MS^10).
3. Post-source component — (d) Collision cell	
Gas type and pressure (bar)	The composition and pressure of the gas used to fragment ions in the collision cell (TOF-TOF, linear trap, Paul trap, or FT-ICR cell).

Collision energy	The specifics for the process of imparting a particular impetus to ions with a given *m/z* value, as they travel into the collision cell for fragmentation. This could be a global figure (e.g. for tandem TOF's), or a complex function; for example a gradient (stepped or continuous) of *m/z* values (for quads) or activation frequencies (for traps) with associated collision energies (given in eV).
3. Post-source component — (e) FT-ICR	
No further parameters required	This component requires no description beyond what is captured in (3c) and (3d).
3. Post-source component — (f) Detectors	
Detector type	Short phrase describing the type of detector used in the machine (*e.g.* microchannel plate, channeltron *etc.*).
Detector sensitivity	An appropriate measure of the sensitivity of the described detector (*e.g.* applied voltage).
4. Spectrum and peak list generation and annotation — for this section, if software other than that listed in 1b (Control and analysis software) is used to perform a task, the producer, name and version of that software must be supplied in each case.	
4. Spectrum and peak list generation and annotation — (a) Spectrum description	
Location of source ('raw') file including file name and type	The location and filename under which the original raw data file from the mass spectrometer is stored, if available. Also give the type of the file where appropriate, or else a description of the software or reference resource used to generate it. Ideally this should be a URI+filename, or most preferably an LSID, where feasible.
Identifying information for the target area	Either a spot number, or some other form of coordinates if more appropriate, that link the spectrum to the part of the plate shot at while acquiring data (MALDI-like methods only).
MS level for this spectrum	The MS level (*e.g.* MS^2) at which this spectrum was acquired.
Ion mode for this spectrum	The ion mode (positive or negative), which is assumed to be the same for all contributing acquisitions.
Precursor *m/z* and charge, with the full mass spectrum containing that peak (for MS level 2 and higher)	For tandem spectra only; the precursor m/z value and the charge state of the precursor ion should be given; to be accompanied by the whole spectrum generated when the precursor ion was selected.
4. Spectrum and peak list generation and annotation — (b) Peak list generation	
Parameters triggering the generation of peak lists from raw data, where appropriate	The total ion count or S/N threshold for a spectrum and the minimum number of ions detected in that scan, for it to be a candidate for grouping in a peak list; plus the mass tolerance (Da) on the precursor ion masses for MS/MS spectra.
Acquisition number (from the 'raw' file) for all acquisitions combined in the peak list, total number and whether summed or averaged.	Where available, the reference numbers of all the scans (as numbered in the raw file) that were combined to produce a peak list, the total number of acquisitions combined to produce the peak list, and whether the peak list was produced by summing or averaging the scans that are listed.
Smoothing; whether applied, parameters	Any peak smoothing should be described, along with the parameters supplied to the algorithm.
Background threshold, or algorithm used	The intensity or S/N cut-off used to filter background noise; or a description of the algorithm used to gate the noise, if complex.
Signal-to-noise estimation and method	The ratio of signal to noise for each *significant* peak in a peak list; significance is defined as being above a given intensity (which should be supplied) or being otherwise of interest; the method of calculation should also be named (if available).
Percentage peak height for centroiding; or	The percentage peak height at which centroids are calculated; if a more complex algorithm is used to perform the

algorithm used, if appropriate	process, it should be named here.
Whether charge states were calculated, spectra were deconvoluted and peaks were deisotoped (with methods described as appropriate)	Firstly, the use of any of these three techniques should be made explicit; secondly, wherever a piece of software other than that named in *1(b)* has been used, the software's manufacturer, and its version, should be provided. For charge state determination, describe the type of data analysed (*i.e.* zoom or full scan, centroid or profile data). For deisotoping declare whether the final peaks are monoisotopic (briefly describe the algorithm if available) or average mass.
Relative times for all acquisitions combined in the peak list	The times relative to the start of the MS run for all acquisitions that were combined in the peak list so that those acquisitions may later be correlated to a chromatogram (continuously-fed electrospray sources only).
Base peak m/z, where appropriate	If the intensities are scaled to that of a 'base peak' then the *m/z* of that base peak should be given.
Metastable peaks removed, if applicable	Add a comment if the analysis software has removed peaks resulting from metastable transitions from the spectrum.
m/z and intensity values	The actual data (*m/z versus* intensity); as described in the preceding sections.
4. *Peak list generation and annotation — (c) Quantitation for selected ions (in addition to 4a) and 4b)*	
Experimental protocol, canonical reference where available with deviations	Which methodology is being used for quantitation (*e.g.* duplex stable isotope labelling, multiplex isobaric tag labelling, label-free method based on spectral count, etc.); if a methods paper for the technique exists, a reference to it should be given and any significant deviations noted.
Number of combined samples and MS runs analysed	The number of experimental classes and MS runs (including number of replicates) that are represented, each with its own different tag.
Quantitation approach (e.g. integration)	Whether the measured value is the area under the selected ion current, max peak height or something else.
Normalisation technique	Describe briefly the normalisation strategy employed; *e.g.* take ratios, then normalise to a global average.
Location of quantitation data, giving the file name, type and Uniform Resource Indicator	The location and filename under which the quantitation data from the statistical analysis are stored, if available. Ideally this should be a URI+filename, or most preferably an LSID. Also give the type of the file where appropriate, or else a description of the software used to generate it.