HUMOS: How to Understand My Orbitrap Spectrum? – An interactive web-based tool to teach the basics of mass spectrometry-based proteomics

Julia A. Bubis, Vladimir Gorshkov, Anja M. Billing, and Frank Kjeldsen

J. Proteome Res., Just Accepted Manuscript • DOI: 10.1021/acs.jproteome.0c00395 • Publication Date (Web): 19 Aug 2020

Downloaded from pubs.acs.org on August 21, 2020
HUMOS: How to Understand My Orbitrap Spectrum? – An interactive web-based tool to teach the basics of mass spectrometry-based proteomics

Julia A. Bubis¹,²,³, Vladimir Gorshkov³, Anja M. Billing³, Frank Kjeldsen³ *

¹ V. L. Talrose Institute for Energy Problems of Chemical Physics, N. N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia
² Moscow Institute of Physics and Technology (State University), 141700 Dolgoprudny, Russia
³ Department of Biochemistry and Molecular Biology, University of Southern Denmark, DK-5230 Odense M, Denmark

* Corresponding author: frankk@bmb.sdu.dk; +45 6550 2439

Abstract

The Orbitrap mass analyzer can provide high mass accuracy and throughput, which has significantly improved proteome research and made this type of instrumentation one of the most frequently applied in proteomics. The efficient use of Orbitrap mass spectrometers requires training. Students in the field of proteomics can benefit from a deeper understanding of the Orbitrap technology in order to comprehend mass spectral interpretation, troubleshooting, and judgment of experimental settings. Unfortunately, the cost of high-end mass spectrometers limits the implementation of this type of equipment in educational laboratories. Guided by these concerns we developed an eLearning web-application called HUMOS aimed to help teaching Orbitrap mass spectrometry. Although a typical proteomics experiment includes the use of several different technologies, such as liquid chromatography, mass spectrometry, and bioinformatics, the learning objectives of HUMOS is focused on mass spectrometry. HUMOS models a mass spectrum of a peptide mixture allowing to investigate the influence of mass spectral acquisition parameters. By changing the parameters and observing the differences students can learn more about mass
spectral resolution, duty cycle, the throughput of the analysis, ion accumulation, spectral dynamic range, and get familiar with advanced spectral acquisition methods, such as BoxCar. **HUMOS** is an open-source software published under Apache license; the live installation is available at [http://humos.bmb.sdu.dk](http://humos.bmb.sdu.dk)

Keywords: proteomics, mass spectrometry, peptides, Orbitrap, eLearning tool

### 1. Introduction

Today mass spectrometry (MS) has become the leading technology in many biological research areas, including proteomics, metabolomics, clinical and pharmaceutical research, forensics, food chemistry, and archaeology.\(^1\)--\(^9\). The typical MS approach in proteomics (proteome characterization) is termed **bottom-up proteomics**\(^10\). It includes protein digestion to peptides, followed by liquid chromatographic (LC) separation and fragmentation of peptides in the gas-phase (MS\(^2\)). The purpose of MS\(^2\) is to fragment peptides into a collection of sequence-informative product ions from which the peptide sequence can be deduced either by manual interrogation or by database searching. Confidently identified peptides can then be matched to specific proteins from a proteome of interest. Mass spectrometry-based proteomics has been further strengthened over the past decades with the introduction of high-resolution hybrid mass spectrometers allowing for accurate mass determination in the ppm range. Because of improved instrumentation, researchers today perform routinely large-scale proteomics studies including characterization of post-translational modifications and quantitative measurements.

Since the introduction of the Orbitrap mass analyzer in 2005\(^11\), this type of instrumentation has become very popular in the proteomics field, as it offers a combination of high mass accuracy and high throughput. Thus, new Orbitrap MS practitioners will benefit from obtaining a solid introduction to key parameters of this instrumentation that is visual and intuitive. In order to efficiently apply Orbitrap-based proteomics in practice, there are certain concepts, for example, duty cycle, space-charge effect, mass spectral resolution, that have to be mastered. It certainly would benefit the training if a student could perform experiments using different parameters of the Orbitrap mass spectrometer in order to better understand their interrelation and the effect of
their values. But both the cost and complexity of hybrid Orbitrap mass spectrometers complicate the use of this equipment in educational laboratories. This was the motivation guiding us through the development of the **HUMOS** application. Typical bottom-up proteomics experiment workflow consists of three major parts – sample preparation, LC-MS analysis followed by computational analysis. We aim here to create an educational tool to introduce students to the most important concepts related to the mass spectrometric part of the workflow, in this particular case an Orbitrap mass spectrometer. This part of the workflow is rarely covered in existing educational resources and therefore we believe that HUMOS can have a significant educational impact. **HUMOS** is intended for students with a basic knowledge of mass spectrometry (e.g. knowledge of isotopes, ionization of analytes, mass-to-charge ratio, ion abundance) and protein/peptide chemistry. This manuscript is divided into the following sections: **What is HUMOS?** (section 2) describes the design and interface of **HUMOS** from the user point of view, **Learning Objectives** and **Sample Questions** (sections 3, 4) are aimed to aid the teacher in using **HUMOS** in a class. A detailed technical description of **HUMOS** is provided in the Supplementary Material.

## 2. What is HUMOS?

**HUMOS** is a web application that is built using Dash ([https://plot.ly/dash/](https://plot.ly/dash/)), a Python framework to create web applications. Thus, the application must be deployed (installed) on a web server while users can access it using any modern web browser both from stationary and mobile devices. No installation is necessary on the user side since all computations and visualizations are performed on the server-side. The source code of **HUMOS** is open and published ([https://github.com/SimpleNumber/HUMOS](https://github.com/SimpleNumber/HUMOS)) under the Apache 2.0 license ([https://www.apache.org/licenses/LICENSE-2.0](https://www.apache.org/licenses/LICENSE-2.0)). Additionally, a fully functional **HUMOS** installation is available at [http://humos.bmb.sdu.dk](http://humos.bmb.sdu.dk). We advise you to open this web site while reading further.
2.1. HUMOS interface

The interface of HUMOS can be divided into several parts (panels) (marked with letters in Figure 1), each panel either allows changing some parameters of the simulation or shows a visual representation of the simulation results. Visual representations are responding dynamically to changes in the parameters. The main panel (Figure 1C) represents the mass spectrum populated with a number of peptide ions. Each peptide in this mass spectrum is modeled in different charge states as a complete isotopic cluster (i.e. including ion signals corresponding to natural isotopic variants (\(^{13}\text{C}, ^{2}\text{H}, ^{15}\text{N}, \text{etc} \)) of the peptide molecule). The peptides are selected randomly among non-redundant peptides from the complete tryptic \emph{in silico} digest of the human proteome and used in all simulations.

The \textbf{Information table} summarizing the \textbf{ion accumulation time} (the time necessary to collect the requested number of ions), the number of \textbf{accumulated ions} (the actual number of ions collected), and the \textbf{scan time} (the shortest time necessary to obtain the mass spectrum, depending on the time of ion accumulation, ion transfer to the Orbitrap, and dwell time in the Orbitrap to obtain the requested resolution) is shown in Figure 1A.

The \textbf{Dynamic range} panel in Figure 1B contains two plots. The first one displays the dynamic range (the difference between highest and lowest abundant components) expressed in orders of magnitude for the modeled peptide mixture and spectra included in the simulation. The second one shows the percentage of peptides in the modeled mixture that is detected with the current settings.

Next to the main mass spectrum, there is a panel (Figure 1D) showing theoretical and observed elution profiles of the most abundant peptide in the model peptide mixture with the current gradient length. Annotations on the plot show peptide sequence, \(m/z\) and charge of peptide ion, and the number of points in the observed elution profile. This plot aids in understanding the importance of the duty cycle for LC-MS experiments.
Below there are three panels with controls. Figure 1E contains controls related to general parameters, Figure 1F – controls related to MS\(^1\) spectra, and Figure 1G – controls related to MS\(^2\) spectra. A detailed description of all settings is provided below.

The **Mass Spectral Resolution** panel (Figure 1H) shows the mass spectrum of two adjacent reporter ions for a Tandem Mass Tags (TMT) 10-plex\(^{12}\) – 127N and 127C differing only by 6.3 mDa and serves to illustrate more evidently the effect of mass spectral resolution.

The **Cycle Time** plot in Figure 1I shows the order of events and which parts of the mass spectrometer are used during a single acquisition cycle (a single iteration through survey (MS\(^1\)) scan and several fragmentation (MS\(^2\)) scans). Next to it, one can assess the throughput of the analysis with the current settings by observing the numbers of MS\(^1\) and MS\(^2\) spectra acquired. It should be noted, however, that the reported values are optimistic estimations since they do not include overhead time spend on tasks such as prescan accumulation. Overhead costs depend on the instrument type and experimental design; thus, we did not include them in the simulation.

The main objective of **HUMOS** is to allow a visual and easy way to learn about the interconnection between modeled parameters (see **Learning Objectives** (section 3) and **Sample Questions** (section 4)).
**Figure 1** The interface of HUMOS
• Peptide distribution

This setting (Figure 1E) allows the students to generate three different abundance distributions of the peptide mixture (one can think of it as a mole quantity of each peptide). It should be noted that such a feature is not possible with any mass spectrometer and serve in HUMOS entirely to allow studying different distributions. There are three options available:

a. **Equimolar** produces a mass spectrum using equal amounts of each peptide. The charge state distribution of a peptide, however, can be different, thus differently charged species (2+, 3+, ...) will have different abundances. In other words, the (equal) amount of every peptide is further split into unequal amounts of each charge state.

b. **Regular** sets all peptide abundances to imitate those experimentally observed in a proteomics experiment with, for example, cell lysates. Charge state distribution is applied on top of abundance distribution, in the same way, as described for equimolar distribution.

c. **Regular with majors** is identical to regular, with the addition, that the quantity of a few peptides is greatly enhanced, i.e. the majority of the molecular species in the mass spectrum belong to these few major peptides. This kind of peptide distribution is common, for example, in plasma proteomics.

• Total Ion Current

Allows changing the ion current (number of ions per second) that is used in the simulation. The total ion current is calculated using all ions in the simulation.

• Acquisition Method

There are two options for mass spectrum acquisition: **Usual MS1** and **BoxCar** (Figure 1E). The BoxCar method \(^{13}\) was developed recently as a solution to a fundamental problem in trapping mass spectrometers – limited ion capacity. Briefly, the problem arises from the fact that only a limited number of ions can be accumulated (trapped) to produce a mass spectrum, thus, ions present in higher quantities will inevitably suppress the accumulation of low
abundant ion species, as a result, there is a limitation on the ratio between the abundances of the highest and the lowest ions in a mass spectrum (it is considered to be around four orders of magnitude in modern mass spectrometers \(^{14}\)). In the regular MS\(^1\), all ions are accumulated simultaneously and, consequently, all ions compete. To mitigate the problem, during BoxCar acquisition ions are selected in narrow “packages” according to their \(m/z\), while the resulting mass spectrum is only calculated in silico by a combination of these packages. The latter allows relaxing the competition between high and low abundant ion species and, therefore, efficiently increases the dynamic range of the mass spectrum analysis. For more detailed information about BoxCar, please refer to the original publication \(^{13}\).

When the BoxCar option is switched on, both regular MS\(^1\) and BoxCar spectra are displayed to facilitate a comparison between the two methods.

- **Resolution**

  This setting found in Figures 1F and 1G allows changing the mass spectral resolution for MS\(^1\) and MS\(^2\) spectra, respectively. The values of the mass spectral resolution are identical to the current generation of the Orbitrap mass spectrometers and the selected resolution applies to \(m/z\) 200 and scales for other \(m/z\) values using the dependency \(\frac{M}{\Delta M} \sim (m/z)^{-1/2}\) as described by Makarov \(^{15}\). The special case of IT resolution in MS\(^2\) allows simulation using an ion trap for fragment spectra acquisition, otherwise, all spectra are acquired in an Orbitrap. Ion trap detection is available in the most advanced “tribrid” mass spectrometers, such as Orbitrap Fusion, Fusion Lumos, and Eclipse.

- **AGC Target and Max Injection Time**

  These parameters in Figure 1F and 1G allow changing the AGC (Automatic Gain Control) target and the maximum injection time used to acquire virtual mass spectra. The AGC target for MS\(^2\) spectra is not used during the simulation and cannot be changed. It is important to note that the highest possible value (1e+07) is beyond the upper limit of the current Orbitrap-based mass spectrometers. AGC target is the maximum total number of ions that can be collected prior to ion detection in the Orbitrap. This number of ions is split among all ion species,
produced by all allowed combinations of a peptide, charge state, and isotopic composition. The maximum injection time is the time allowed to be spent to accumulate ions and potentially reach the AGC target value. These two parameters work together, in simple terms, ion accumulation is stopped either when the maximum allowed number of ions (AGC target) is reached, or when the maximum allowed time for ion accumulation (maximum injection time) is reached, in the latter case the actual number of accumulated ions will be less than the AGC target. For a detailed discussion, please refer to \textsuperscript{16,17}.

- **TopN / TopSpeed**

This setting (Figure 1G) controls the number of fragmentation events (MS\textsuperscript{2} scans) that are scheduled for each parent ion (MS\textsuperscript{1}) scan. The checkbox below controls \textbf{Parallelization}. With the use of parallelization different parts of the mass spectrometer can perform their tasks at the same time, for example, accumulation of ions for the proceeding mass spectrum can happen at the same time as ions of the current mass spectrum are being analyzed. The first option – TopN – limits the number of fragmentation events to a user-defined value. During a real proteomics experiment, TopN value is used as the highest allowed number of dependent (MS\textsuperscript{2}) scans to be performed for each precursor (MS\textsuperscript{1}) scan, thus, sometimes the instrument will perform less than TopN dependent scans. On the contrary, in HUMOS TopN value is used as the average number of dependent scans for each precursor scan. Although, for a sample of high complexity (for example, full proteome) and optimal instrument settings, the average number is expected to be close to the maximum TopN. The second option – TopSpeed – allows controlling the number of fragmentation events indirectly, the number of dependent scans adjusted to reach user-defined duty cycle length. This option is available on the latest generations of Orbitrap-instruments, such as tribrids and Exploris series.

3. **Learning objectives**

Using HUMOS the students can become familiar with the following basic concepts related to mass spectrometry and proteomics:
3.1. Mass resolution

For best learning yield a basic understanding of the term mass spectral resolution, i.e. the definition and how it is calculated is advised. With HUMOS the students will be able to learn which factors influence the mass resolution in proteomics experiment (both MS\(^1\) and MS\(^2\) scans), and the benefits and drawbacks of using high/low resolution. In particular, which resolution is necessary for mass spectrometry applications involving isotopes (Figure 1H), such as in TMT-10plex quantitation.

3.2. Dynamic range

The students will be able to familiarize themselves with the concept of the dynamic range of the sample and the mass spectrum, study the influence of highly abundant peptides on the detection of low abundant peptides, and learn about a state-of-the-art method to deal with dynamic range limitation – the BoxCar acquisition method.

3.3. Ion accumulation

The students will be able to investigate the relationship between the number of accumulated ions and mass spectrum quality e.g. number of observed peptides, the number of isotopic peaks, and their relative intensity. Furthermore, it is possible to increase the understanding of the interrelation of AGC target and maximum injection time.

3.4. Duty cycle and throughput of the analysis

The students will be introduced to the concept of duty cycle and will be able to perform experiments with various mass spectral acquisition settings to experience their influence on the duty cycle and, thus, on the throughput of the analysis.

4. Training questions

This section contains a set of exemplary training questions that can be used in the class. These questions can be addressed either by individuals or small groups and later be collectively
discussed in plenum. The questions are intended to highlight the most important observations, stimulate discussions, and ensure that learning objectives are met. Each question is provided with an answer, however, the answer itself is not intended to be definite or exhaustive but rather should serve as guidance for a discussion.

**Q.1: If the “equimolar” spectrum is generated from a mixture with equal amounts of peptides, then why the distribution of peptide ion charge states is not the same?**

*Hint: Consider the impact of individual peptide charge state distributions. Select two different peptides (different m/z values of the same charge state), then identify and evaluate their ion abundances as 2+ and 3+ ions, respectively.*

A.1: Although the amount (total number of molecules) of each peptide is identical, each peptide might be presented in several different charge states. Their relative abundance is determined both by the number, order, and types of amino acids within the peptide sequence. Thus, different charge states of the same peptide will have different abundances as we observe in the equimolar spectrum.

**Q.2: What is the difference in observed peptide repertoire between the peptide distributions (equimolar, regular, etc.)? How the number of detected peptides depends on the distribution?**

A.2: The **Equimolar** distribution allows observing the most complete repertoire of the peptide mixture. Even with differences in charge state distribution, peptides at equal amounts produce a close to evenly intense mass spectrum. It is important to mention that such equimolar mixtures are rarely if ever observed in real proteomics experiments.

The **Regular** peptide distribution has a much greater difference in peptide abundance; thus, the most frequent peptides have a greater propensity to be detected. As a result, less frequent peptides can be missed, especially when a low number of ions is to be accumulated because of low AGC target, short accumulation time, or low ion flux (number of ions per unit of time).
The presence of enhanced major peptides (Regular with majors) makes it even less likely to detect other low abundant peptides. This mixture serves to simulate samples like blood plasma that spans ten orders of magnitude between the least and the most abundant proteins.

By comparing the percentage of detected peptides between the distributions using the default settings, one can see that the Equimolar mass spectrum always results in complete coverage (100%), while Regular and Regular with majors usually show values slightly below 100%. The exact value might differ due to randomness in ion sampling.

HUMOS allows the user to switch between the different peptide abundance distributions and visually assess how the same set of peptides is visualized in the mass spectrum.

Q.3: What mass resolution should be used in a standard bottom-up experiment and why? How does higher resolution influence your experiment?

A.3: The resolution to be selected should be based on experimental needs. Higher resolution, in general, allows detecting more peaks both by improving signal-to-noise ($S/N \sim \sqrt{R}$) and resolving peaks having isobaric masses (ions with the same nominal mass but a small difference in exact mass). Higher mass resolution is also necessary for higher mass accuracy measurements of peptide and fragment ions. The penalty of acquiring higher mass resolution in Orbitrap mass spectrometry relates to a required longer spectrum acquisition time and, thus, the lower overall throughput of the analysis (i.e. fewer peptides can be analyzed per unit of time).

Typically, one would run a bottom-up proteomics experiment with 60,000 or 120,000 resolution in MS\textsuperscript{1} scans, to record more peptides with high-mass accuracy. When aiming for a maximum throughput of the analysis, the resolution for MS\textsuperscript{2} spectra should be set to the lowest acceptable value (i.e. 7,500 or 15,000 for modern Orbitrap-based instruments), however, when more sensitivity, resolving power, and mass accuracy are desired, the resolution in MS\textsuperscript{2} can be increased. Kelstrup et al\textsuperscript{19} provides more practical advice on balancing the sensitivity and the throughput, using an Orbitrap hybrid mass spectrometer. Throughput can be increased even more by applying additional mass analyzers, such as ion trap, to acquire fragmentation spectra, the cost of it is, however, a significant decrease in mass spectral resolution and mass accuracy.
HUMOS allows users to experiment with the resolving power in MS\(^1\) and MS\(^2\) scans and observe the influence of this parameter on the number of detected peptides and throughput of the analysis (the number of MS\(^1\) and MS\(^2\) scans per hour, length of the duty cycle – the time between consecutive MS\(^1\) scans). One can also zoom in on any peptide ion in the main panel (Figure 1C) and see how the width of peaks decrease with increased resolution.

Q.4: **What is the minimum necessary resolution to distinguish 127N and 127C TMT reporter ions?**

A.4: The accurate masses of 127N and 127C reporter ions, are 127.124761 Da and 127.131081 Da, respectively. The resolution is defined as \( \frac{M}{\Delta M}\), where \( \Delta M \) is the full width of the peak at half maximum (FWHM), to get acceptable separation of the peaks one should require the FWHM to be twice lower than the distance between the peaks. Thus, we can calculate the minimal resolution at mass 127 as \( \frac{2 \cdot 127}{127.131081 - 127.124761} \approx 40,200 \). Since the resolution in modern Orbitrap-based instruments is calculated at mass 200 and \( \frac{M}{\Delta M} \sim (m/z)^{-1/2} \), the resolution should be approximately 32,000 at mass 200. In practice, one would like to have complete baseline separation, which would require a minimum resolution of 45,000.

Using HUMOS, the user can see that the resolution of 15,000 is insufficient for separation, at 30,000 one can observe better but incomplete separation, while 60,000 allows complete baseline separation.

Q.5: **How does the accumulation of more ions (higher AGC target) change the mass spectrum?**

*Hint: check overall “shape” of the mass spectrum and then focus on some low-abundance ion clusters*

A.5: Since the accumulation of ions is a stochastic process, certain variations of observed abundances are expected. However, due to the large (>100,000 with default settings) number of ions accumulated, the ratio between major ion signals in the mass spectrum (the “overall shape”) does not change much even for the lowest AGC target. On the contrary, the abundances of minor components are more sensitive to stochastic sampling, thus the changes in the shape of low-
abundant isotopic clusters due to poor ion statistics and vanishing of low abundant peaks can be observed.

As expected, the effect is the least prominent for **Equimolar** peptide distribution, since all peptides are “major”, while **Regular with majors** distribution is affected the most.

In **HUMOS**, users can do a repeated sampling of the ions to generate an updated mass spectrum, for example, by clicking on the same AGC target value several times. They can observe that with the lowest AGC target value the abundance of major peaks is changing slightly. Later they can select one of the “small” peaks and observe the changes in the shape of the isotopic cluster and disappearing of some peaks. One can perform the same test using higher AGC target values and different peptide distributions.

**Q.6: In terms of observed ion dynamic range (difference in the span of most and least abundant ions) could it be advantageous to accumulate as many ions as possible? Evaluate how the observed mass spectral dynamic range relates to the peptide mixture dynamic range. Try different peptide distributions.**

**A.6:** Ion accumulation is only one of the factors, along with sample type, ionization method, and ion detection, that specify the observed dynamic range. Accumulation of more ions, in general, improves the detection of low abundant ion species since they have an opportunity to pass the detection limit. Thus, by this consideration alone, it should be preferred to accumulate as many ions as possible, given the available flux of ions. However, the detection of such a high dynamic range is only possible with the “ideal” mass analyzer. The Orbitrap mass analyzer has certain physical and electronic limitations, that restrict the detectable dynamic range (it is reported that current generation of Orbitrap mass analyzers allows 1:5000 intraspectral dynamic range \(^{14}\)), additionally in the case of high ion loads, one will inevitably face the negative effects of space charging (discussion on space charging can be found elsewhere \(^{20,21}\)). Consequently, in practical applications, it is necessary to balance the ion load with the detector capabilities and negative effects of high ion loading (space-charge effect).
The comparison can be based on the position (which abundances are included) and the magnitude of dynamic range as presented in Figure 1B. As one can observe mass spectral dynamic range covers only the upper part of the peptide mixture dynamic range (i.e. only the most abundant ions are included) and the magnitude rarely matches between them. The closest match happens when using the **Equimolar** distribution. The dynamic range of the peptide mixture increases in the following order – **Equimolar, Regular, Regular with majors**. With the increase in the magnitude of sample dynamic range the magnitude of spectral dynamic range increases as well, however, it peaks at about four orders of magnitude. Thus, the observed mass spectrum provides close to complete image of the ion population only for the samples with low dynamic range, in other cases a significant part of the ion population will not be reflected in the spectrum.

**Q.7: How does lower ion current influence mass spectrum characteristics? Is there a way to compensate for the observed effect? Which peptide distribution is affected the most? What consequences can it have?**

*Hint: Keeping the same AGC target and maximum injection time change the value of total ion current and observe the changes in the number of accumulated ions, accumulation time, dynamic range, and percentage of detected peptides, next change the values for AGC target and max injection time and observe the difference.*

A.7: Lowering the total ion current leads to longer accumulation times, however as long as the reduction in ion flux does not impair the accumulation of the preset number of ions (i.e. AGC target can be reached), all characteristics of the mass spectrum remains the same. Reduction in the number of accumulated ions (below AGC target) results in lowering the observed percentage of peptides and mass spectral dynamic range. To compensate for the low ion flux, one can increase the maximum injection time; longer injection time might allow collecting enough ions to receive a high-quality mass spectrum. The strongest effect of low ion current can be observed with the **Regular with majors** distribution since the ion population is dominated by a small number of high abundant ions, these ions will be the first to pass the detection limit, while low abundant ones will not be accumulated in necessary quantities. All summed together, it shows the importance of
keeping the ion flux high enough to get rich mass spectra, especially when analyzing samples with “major” peptides.

Q.8: Starting with a high number of ions (regular distribution, total ion current 1e9, max injection time 50 ms) restrict the number of ions to a small value (total ion current 1e5, max injection time 5 ms) in several steps. How does the visual of the mass spectrum changes? What max injection time is necessary at the lowest ion current to get a high-quality mass spectrum? How does the cycle time change?

A.8: While lowering the number of accumulated ions one can observe losing information from low abundant peaks in the mass spectrum, only the most abundant peaks remain visible. At a very low number of ions (< 1e3) the abundances even for the most abundant peaks are greatly distorted (i.e. the observed abundances do not represent the real ones even in the relative sense). At the lowest value of the ion current, one needs to use very high max injection time > 2000 ms to get a mass spectral quality close to the starting one. Long ion accumulation times significantly increase the cycle time, and, thus, reduce the throughput of the analysis.

This exercise illustrates the situation typical for fragmentation spectra. Even with a high sample load, low abundant peptides will require accumulation for longer times to collect enough ions for a high-quality fragmentation spectrum. Without any restriction, necessary accumulation time will in most cases result in an unreasonably long duty cycle. This fact is the rationalization behind the use of the maximum injection time parameter. Designing an experiment, a researcher should find a compromise between sensitivity (high mass spectrum quality) and throughput of the analysis.

Q.9: How MS^2 resolution, maximum injection time, and parallelization influence the duty cycle? Is it possible to select the optimal settings for the most efficient use?

A.9: Under the DDA experiment the mass spectrometer operates in cycles, first parent ion (MS^1) scan, followed by several fragmentation (MS^2) scans. The time between the consecutive parent ion scans is named the duty cycle time. Fragment scans are used for peptide identification, the number of them determines how many peptides can be identified, while the parent ion scans are used for
detecting the peptides to be fragmented and later for their quantification. Thus, the length of the
duty cycle is an important characteristic of instrument performance.

A higher resolution leads to longer scan times and, thus, increases the cycle time; longer maximum
injection times have the same effect. In HUMOS cycle time is modeled with the assumption that ion
accumulation for MS$^2$ always terminates at maximum injection time – the worst-case scenario. In
real proteomics experiments at least some peptide precursors will have enough ion flux to reach
the AGC target before the maximum injection time. The portion of such precursors will depend on
the sample and experimental setup; however, it is usually kept low in a high-throughput
proteomics experiment. It is up to the researcher to decide whether scan speed (longer cycle)
should be traded for mass spectral quality (more ions accumulated).

Parallelization is the ability of the instrument to operate different parts of it simultaneously. This
feature is feasible in most modern instruments. The effect is clearly visible during MS$^2$ – the
accumulation of ions happens at the same time as mass spectrum acquisition. HUMOS allows
modeling two different scenarios – 1) the same mass analyzer (Orbitrap) is used for both MS1 and
MS2 spectra acquisition; 2) two different mass analyzers (Orbitrap and ion trap) are used. Using
an extra mass analyzer allows better parallelization and shrinks the cycle time significantly. It is
worth noting, that the second option requires more sophisticated instrumental design and is
available only in high-end mass spectrometers.

With certain combinations of resolution and maximum injection time, one can observe gaps in the
time when a mass analyzer or the ion source are busy, thus, it is practical to align the load for all
instrument elements. For example, if a high resolution of fragment scans is necessary, one can use
a much larger maximum injection time without any negative effect on the duty cycle.

Q.10: Change the duration of the duty cycle either directly (TopSpeed mode) or indirectly
(TopN mode). What changes can be observed in the Peptide Elution Profile plot?

A.10: The general trend is that the observed peptide elution profile resembles the theoretical one
increasingly better with a decrease in duty cycle duration. The close similarity between these
profiles allows a more precise estimation of the elution peak height/area (important for label-free
quantitation) and peptide elution time. It is important to emphasize that changes in the elution peak shape and position of sampling points will influence the correlation between observed and theoretical elution profiles. Since neither of these two factors can be fully controlled in a real experiment, a common recommendation is to select the duty cycle duration that achieves at least 5 – 7 sampling points per average eluting peak. The elution profile width will grow proportionally with the length of LC gradient and, thus, for longer LC-MS runs the required conditions can be reached for longer duty cycles.

Q.11: Describe the changes to the peptide abundances in the BoxCar method compared to the usual MS\textsuperscript{1} method. Which peptide distribution shows the most changes?

A.11: With the same number of total accumulated ions, the BoxCar mode permits observing a greater variety of ion species – ions that are low in abundance in regular MS\textsuperscript{1} mode become visible/detectable. It is important to emphasize that depending on the selected parameters the intensity of high abundant ions can become lower in the BoxCar mode. In general, this leads to the so-called equalization of the mass spectrum – the difference in the observed abundances is becoming less. It is especially evident when switching between different peptide distributions since the equimolar one does not change much in the BoxCar mode. BoxCar makes other distributions look more like equimolar. Consequently, the distribution that is the least similar to equimolar is influenced the most.

BoxCar aims to increase the dynamic range of spectra; however, it is important to note, that the improvement of the dynamic range occurs only after computational processing of the raw spectra obtained by the instrument.

HUMOS displays BoxCar spectra as they will look before the processing to emphasize the changes in ion intensity and allow a deeper understanding of BoxCar mode of action. By clicking on the trace icons in the main graph legend it is possible to show and hide spectra to make the equalization more evident. Dynamic range plot (Figure 1B) allows observing the dynamic ranges reached in individual BoxCar scans and comparing them to MS\textsuperscript{1} and model peptide mixture. One can clearly observe the improvement in mass spectral dynamic range (typically ~ 0.5 orders of
magnitude) obtained using the BoxCar method for Regular and Regular with majors abundance distributions.

Q.12: Test BoxCar mode with different ion distributions (equimolar, typical, etc.), AGC targets, and injection times. What do you observe?

A.12: The largest changes in the BoxCar mode are observed in the “sparse” regions of the spectrum, i.e. with a smaller number of ion signals and/or many low abundant ions. When compared to regular MS\(^1\) mode, BoxCar acquisition results in a large increase in intensities of ions in these regions, while in “dense” regions, having high abundant ions or a lot of ions with the same abundance, the abundances do not improve much or even become lower. The most improvement measured as additional peaks becoming visible in the mass spectrum can be observed using smaller AGC targets. With larger AGC target values most of the peaks are present both in regular MS\(^1\) and BoxCar spectra, however, the intensity in BoxCar is generally higher, leading to a larger number of isotopic peaks, more well-defined peak shapes and intensities (re-sampling of spectrum results in less variation). All together it improves the quality of the mass spectrum. Another prominent observation is that an increase in maximum injection time might be necessary to observe the improvement in BoxCar mode.

It is important to understand, that effectively BoxCar allows reaching higher AGC targets without facing the negative effects of it, the ions are collected more uniformly across the same scan range and split into several narrow boxes thus reducing competition between high and low abundant ions. On the other hand, the flux of low abundant ions is lower and, thus, their accumulation will be slower.

Q.13: Compare the number of accumulated ions and ion accumulation time for MS\(^1\) and BoxCar. What can you observe? Reduce the ion current in BoxCar mode? What can you observe?

Hint: Do the number of accumulated ions and ion accumulation time reach target values? Can you come up with an explanation?
A.13: Ion accumulation time to reach a similar number of accumulated ions is typically about two orders of magnitude higher for BoxCar compared to regular MS\(^1\) scan, although it can happen that even with extremely long maximum injection times the target AGC won’t be reached in BoxCar mode. During the BoxCar acquisition, the full mass range is split into smaller chunks – boxes – both AGC target and maximum injection time is distributed equally between the boxes, for example, if the maximum injection time is set for 100 ms, and AGC target is 1e6, than, having 10 boxes, each one will be accumulated for 10 ms or up to 1e5 ions (whichever comes first). Since some boxes might have very low abundant ions or no ions at all, requested AGC value might be reached only with some unrealistic injection time or never at all.

Lowering ion current leads to a proportional decrease in the number of collected ions and an increase in ion accumulation time necessary to reach the same number of ions. In the case of BoxCar, this leads to impractically long max injection time (too long duty cycle) or a very small number of ions (poor mass spectrum quality). Low ion flux eliminates benefits provided by the BoxCar approach.

In HUMOS spectra are generated “noiseless”, while real mass spectra will always contain some chemical and electronic noise that will limit detection of low abundant ions, i.e. if an ion signal is significantly lower than noise level it will not be accumulated in detectable quantities even with infinite acquisition time. This simplification, however, does not affect the educational value of HUMOS, since competition between high and low abundant ions (both inside a box or inside of a full spectrum) is the major factor influencing dynamic range.

Q.14: Compare the instrument throughput (number of MS\(^1\) and MS\(^2\) scans) in BoxCar and MS\(^1\) mode? Test different TopN values. With the same duty cycle length (TopSpeed setting) compare the number of MS\(^2\) scans per cycle. What consequences this might have?

A.14: The throughput of the analysis in the BoxCar mode is significantly lower since each BoxCar scan requires both more scans and longer accumulation time for each scan. Modern Orbitrap-based mass spectrometers allow accumulation of ions in parallel with signal detection, i.e. relatively long accumulation of several ion boxes during BoxCar can be performed in “spare” time, while the previous portion of ions is dwelling in an Orbitrap. For example, using the most common resolution
of 120,000 will result in a scan time of 265 ms that should leave enough time for parallel accumulation of ions in the BoxCar mode (compare this time with typical ion accumulation time in BoxCar mode). However, even with complete parallelization, several high-resolution scans need to be acquired, making the duty cycle significantly longer. If the duty cycle is kept constant (TopSpeed mode) the number of fragmentation spectra per cycle is much lower in BoxCar mode. Both longer duty cycle and a smaller number of MS\(^2\) scans mean fewer precursors can be fragmented, leading to less identified peptides. Additionally, longer cycle time might affect the number of MS\(^1\) scans per chromatographic peak (as can be observed on Peptide Elution Profile panel) and, thus, quantification accuracy in label-free and parent ion-based quantification methods.

Q.15: What methods can you suggest for improving the duty cycle of the BoxCar scan? (Assume there are no technical limitations).

A.15: There are few suggestions listed in order of whimsicality.

1. Scheduling MS\(^1\) and MS\(^2\) scans in a different order can help get better parallelization. The time necessary for several high-resolution MS\(^1\) scans can be used to accumulate ions for MS\(^2\) scans, if an additional mass analyzer is available, not only accumulation but recording of mass spectra is possible.

2. Faster mass spectral acquisition (lower dwell time) will lower the time demand for high-resolution scans.

3. Several Orbitrap analyzers running in parallel will shrink the duty cycle as well. The last two methods are not specific for BoxCar and might face other limitations – such as ion flux.

**Supporting Information**

The following supporting information is available free of charge at ACS website http://pubs.acs.org

*Suppl. Material S1.* Detailed technical description of HUMOS
References


https://doi.org/10.1002/rcm.955.

https://doi.org/10.1074/mcp.r200007-mcp200.


https://doi.org/10.1002/mas.20186.


Acknowledgment

The authors are grateful for the IT Service of the University of Southern Denmark for providing the possibility to offer a public live demo of HUMOS and the members of the Protein Research Group at the University of Southern Denmark for the participation and valuable feedback in the evaluation of HUMOS as a teaching resource.
How to Understand My Orbitrap Spectrum?
How to Understand My Orbitrap Spectrum?
HUMOS: How to Understand My Orbitrap Spectrum?

A: Information Table

<table>
<thead>
<tr>
<th></th>
<th>MS1</th>
<th>BoxCar scan 1</th>
<th>BoxCar scan 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ion accumulation time, ms</td>
<td>0.50</td>
<td>44.18</td>
<td>55.83</td>
</tr>
<tr>
<td>Accumulated ions</td>
<td>5.0e+05</td>
<td>3.3e+05</td>
<td>2.9e+05</td>
</tr>
<tr>
<td>Scan time, ms</td>
<td>265.00</td>
<td>265.00</td>
<td>265.00</td>
</tr>
</tbody>
</table>

B: Dynamic Range

- Spectrum
- BoxCar scan 1
- BoxCar scan 2

C: Peptide Elution Profile

D: RT (s)

E: Peptide Distribution
- Equimolar
- Regular
- Regular with majors

F: MS1 Resolution
- 7500
- 15000
- 30000
- 50000
- 70000
- 100000
- 200000
- 300000
- 500000
- 1000000
- 10000000
- 100000000
- 1000000000
- 10000000000
- 100000000000
- 1000000000000

G: MS2 Resolution
- 30
- 60
- 90
- 120
- 150
- 180
- 210
- 240

H: Mass Spectral Resolution

The graph shows two adjacent TMT 10-plex reporter ions

I: Cycle Time

1.410 sec

#MS2: 15

MS1 Scans in 60 minutes: 2553
MS2 Scans in 60 minutes: 38295