

SIMPATIQCO: A Server-Based Software Suite Which Facilitates Monitoring the Time Course of LC–MS Performance Metrics on Orbitrap Instruments

Peter Pichler,^{*,†,‡,§} Michael Mazanek,^{§,||} Frederico Dusberger,[†] Lisa Weilmöck,[⊥] Christian G. Huber,[⊥] Christoph Stingl,[#] Theo M. Luider,[#] Werner L. Straube,[†] Thomas Köcher,[†] and Karl Mechtler^{*,†,||}

[†]Research Institute of Molecular Pathology, Vienna, Austria

[‡]Christian Doppler Laboratory for Proteome Analysis, University of Vienna, Austria

^{||}Institute of Molecular Biotechnology, Vienna, Austria

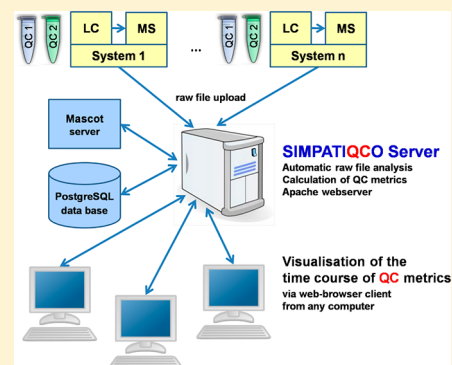
[⊥]Department of Molecular Biology, University of Salzburg, Austria

[#]Department of Neurology, Erasmus MC, Rotterdam, The Netherlands

S Supporting Information

ABSTRACT: While the performance of liquid chromatography (LC) and mass spectrometry (MS) instrumentation continues to increase, applications such as analyses of complete or near-complete proteomes and quantitative studies require constant and optimal system performance. For this reason, research laboratories and core facilities alike are recommended to implement quality control (QC) measures as part of their routine workflows. Many laboratories perform sporadic quality control checks. However, successive and systematic longitudinal monitoring of system performance would be facilitated by dedicated automatic or semi-automatic software solutions that aid an effortless analysis and display of QC metrics over time. We present the software package SIMPATIQCO (SIMPLE AuTomatIc Quality CONTROL) designed for evaluation of data from LTQ Orbitrap, Q-Exactive, LTQ FT, and LTQ instruments. A centralized SIMPATIQCO server can process QC data from multiple instruments. The software calculates QC metrics supervising every step of data acquisition from LC and electrospray to MS. For each QC metric the software learns the range indicating adequate system performance from the uploaded data using robust statistics. Results are stored in a database and can be displayed in a comfortable manner from any computer in the laboratory via a web browser. QC data can be monitored for individual LC runs as well as plotted over time. SIMPATIQCO thus assists the longitudinal monitoring of important QC metrics such as peptide elution times, peak widths, intensities, total ion current (TIC) as well as sensitivity, and overall LC–MS system performance; in this way the software also helps identify potential problems. The SIMPATIQCO software package is available free of charge.

KEYWORDS: shotgun proteomics, quality control, performance metrics, mass spectrometry, software, Orbitrap



INTRODUCTION

Proteomics workflows share one essential requirement: they necessitate appropriate quality control (QC). Stringent QC not only requires an evaluation of outcomes, rather QC starts with adequate experimental design; it incorporates monitoring of processes, protocols, and reagents; and it includes aspects such as adequate training of personnel and a surveillance of the proper functioning of instrumentation.^{1–3} For proteomics workflows, QC must be established at four steps: sample preparation, liquid chromatography (LC), mass spectrometry (MS), and data interpretation.³ A recent study showed that a major reason for the failure of a number of laboratories to reliably identify proteins present in a relatively simple test sample was due to inadequate data processing; the same study suggested that this could be overcome by standardized operating procedures involving appropriate protocols for data interpretation.⁴

This is reassuring as a bioinformatics data interpretation workflow, once established, can be maintained easily because software algorithms continue to work reproducibly with moderate requirement for maintenance.^{5–7} In contrast, all the other above-mentioned steps are subject to variation from experiment to experiment, and suboptimal system performance needs to be identified and corrected promptly. For this reason, software which facilitates an effortless monitoring of the performance of LC and MS systems is highly desirable.

Improvements in LC and MS instrumentation contribute to the remarkable performance of today's LC–MS equipment, empowering researchers to implement novel experimental protocols. However, challenging applications require measures

Received: February 19, 2012

Published: October 22, 2012

to ensure constant and optimal system performance. Highly sophisticated QC approaches were reported previously such as the NIST MSQC pipeline, which outputs 46 QC metrics,⁸ or QuaMeter a software which can compute up to 42 QC metrics for instruments from different vendors.⁹ QC software has also been reported for the detection of LC runs that constitute outliers¹⁰ and for computation and visualization of QC metrics from individual LC runs,¹¹ and a real-time QC solution for monitoring electrospray quality and LC backpressure was described recently.¹² However, to our knowledge no currently available QC software allows an automated upload of QC runs to a dedicated server postacquisition, followed by seamless computation of QC metrics based on peptide identifications and instrument parameters, storage of QC results in a database, and the possibility for convenient plotting of the time course of QC metrics via a web browser. Such a longitudinal analysis of QC data from multiple LC runs could potentially reveal suboptimal LC–MS system performance as a deviation outside the range of a QC metric when the LC–MS system is operating under optimal conditions.

We therefore developed SIMPATIQCO (*SIM*ple *Au*Tomat*Ic* Quality *C*ontrol), a software suite specifically designed to aid routine longitudinal monitoring of QC metrics and troubleshooting when using LTQ Orbitrap, Q-Exactive, LTQ FT, and LTQ instruments. SIMPATIQCO is a server-based application, designed to process raw files from the above-mentioned Thermo Scientific instruments that can be uploaded either manually (individually or in batch) via a web browser interface or in an automated manner via a “hot folder” (Windows network share or ftp). As Xcalibur can be configured to run a batch file postacquisition, raw files can be copied instantaneously from mass spectrometers in the laboratory to the “hot folder” on the SIMPATIQCO host for automatic processing. SIMPATIQCO then extracts information from the raw file, starts a Mascot search to identify spectra, and calculates QC metrics. Results are stored in a database and can be viewed for individual QC runs as well as plotted over time via a web browser interface from any computer in the laboratory.

To demonstrate the successful usage and general utility of the software suite, we present longitudinal QC data from six instruments and four laboratories in the Supporting Information (Figure S-1). Three of the four laboratories installed a SIMPATIQCO server in-house, while one laboratory uploaded data to a SIMPATIQCO server installed and operated by one of the other laboratories (which is an option for laboratories that do not have the resources to maintain the server themselves).

We also propose an LC–MS quality control workflow based on the routine surveillance of QC metrics with two types of QC samples that are intercalated with regular samples to monitor two disparate aspects of overall system performance: a sample of very low complexity designed to monitor sensitivity using an analyte (peptide) concentration only modestly above the limit of detection (LOD) of a system performing close to optimum; and another sample of very high complexity that poses a challenge to the speed of data acquisition.¹³ Metrics derived from these samples thus reflect the sensitivity as well as the speed and overall performance of the LC–MS system. Of note, SIMPATIQCO can be configured to accommodate other types of samples, as disparate QC samples are in use within the community. Indeed, some of the above-mentioned participating laboratories preferred to use different standard proteins and

synthetic peptides and different concentrations in their QC samples.

The purpose of SIMPATIQCO is to provide a set of key QC metrics that are generated in an effortless manner and that can be displayed conveniently. SIMPATIQCO facilitates a longitudinal surveillance of the time course of system performance and within-laboratory reproducibility, and it helps pinpoint potential suboptimal system performance. Identification of outliers with regard to the performance metrics should prompt an immediate enquiry so that the reason for deviant performance can be identified and resolved. In this way, SIMPATIQCO may aid proteomics laboratories achieve and maintain optimum quality of data acquisition.

METHODS

SIMPATIQCO was developed for deployment (installation) on a Windows computer. Raw files from QC runs intercalated with regular samples are uploaded to a server. SIMPATIQCO extracts data such as ion injection times, TIC, and lock mass information, and estimates the overfill or underfill ratio of the automatic gain control (AGC) target value for each spectrum.¹⁴ MS/MS spectra are submitted to a Mascot server for peptide identification. SIMPATIQCO then calculates QC metrics that reflect the performance of LC, electrospray, and MS such as peak widths and areas, peptide elution times and intensities, TIC, the number of identified peptide–spectrum matches per minute, and the sequence coverage of “proteins of interest”. For each QC metric the range indicating optimal system performance is learned from the data and displayed using a background band colored in green, whereas yellow or red colored bands highlight deviant system performance. We calculate robust statistics on QC metrics, eliminating the requirement to manually remove outliers (“bad runs”). Results are stored in a PostgreSQL database. PHP scripts and Apache webserver allow comfortable display of results from any computer in the laboratory.

A description of the QC samples and the LC and MS methods used by the four laboratories is provided in the Supporting Information, illustrating that SIMPATIQCO is compatible with different QC work flows established in various laboratories. A detailed overview of the components and design of the SIMPATIQCO software package is also provided in the Supplementary Materials and Methods section within the Supporting Information. Further information, download links, installation guideline, manual, and FAQs are available online at <http://ms.imp.ac.at/>.

RESULTS AND DISCUSSION

Obtaining mass spectrometry data for state-of-the-art proteomics studies requires QC measures. In our view, the key to achieve and maintain optimum data quality is successive, routine, and if possible automated rather than sporadic and manual monitoring of system performance. In the following we present a QC workflow as well as a software package aimed to aid the maintenance of constant and optimum LC–MS system performance.

Traditionally the LC has been a major source of concern, and close surveillance of LC performance is certainly an important prerequisite to ensure high quality proteomics data. However, in our experience there is likewise a requirement to monitor the performance of the MS instrumentation as closely as that of the LC. For instance with the newer generations of Orbitrap

instruments that have enhanced sensitivity partly due to increasing the overall gas throughput into the system,¹⁵ there may be the need to clean various components of the instrument more often in order to maintain optimum performance, as deposition of particles onto components of the ion optics can lead to a decrease in ion transmission.

Both LC and MS instruments offer a variety of instrument-specific system tests. These tests may provide very valuable hints during troubleshooting. For instance, the latest version of Velos Orbitrap instrument software (Tune 2.7 SP1) offers several diagnostic procedures that have been added to assist in determining whether a deposition of charged particles is occurring and which part of the ion optics most likely has the contamination. However, several types of system evaluation procedures on Orbitrap instruments may require switching from the nanosource to the ESI source, interrupting measurements in addition to the time required for the tests themselves. For this reason system tests are costly to perform on a daily or more than daily basis. In addition there is a risk that the mounting and unmounting of system parts to perform these tests might lead to a deterioration of system performance aside from the time requirements. Cleaning procedures and the associated bake-out of the instrument lead to prolonged instrument downtime, so that there is a requirement to estimate appropriate timing of such maintenance procedures. We therefore reasoned that it would be desirable to have a system evaluation test at hand that can be performed without any need to interrupt the running LC–MS system.

We hence suggest that the more time-consuming instrument-specific tests such as those mentioned above be performed only at defined time points dedicated to system maintenance and during setup or troubleshooting of an LC–MS system. In these situations, considerable time and effort is invested into optimizing system performance. Manual evaluation of raw files and analysis with tools such as rawMeat (<http://vastsci.com/rawmeat/>), NIST MSQC,⁸ QuaMeter,⁹ LogViewer,¹¹ and MaxQuant¹⁶ can provide valuable information at this stage. In addition parameters for methods should be tested in order to evaluate their effect on system performance. These time-consuming optimization steps are necessary because monitoring the time course of performance metrics can only be regarded as an indicator of optimum system performance when the values of the metrics under optimal conditions have been established in the first place. After these initial steps, we suggest a routine analysis of two types of QC samples in short time intervals to monitor QC metrics such as sensitivity as well as speed and overall performance of the entire LC–MS system. The QC samples are measured intercalated within the regular experiments, and analyzed automatically with the help of SIMPATIQCO so that there is no interruption of system operation.

The general concept is based on the consideration that relevant system alterations involving a significant change in system performance should lead to a change of QC metrics derived from either or both of the two QC samples. As we want the QC samples to indicate whether the LC–MS system is in an adequate status for measuring actual samples, the choice of the two QC samples is such that one QC sample is designed to evaluate sensitivity (QC 1) whereas the other one is intended to monitor overall system performance and speed (QC 2). The choice of a suitable amount of the QC 1 sensitivity sample is important. One may first perform a dilution series to establish the LOD and then decide to inject an amount only modestly above this value. When the injected amount of QC 1 sample is

too high, the test would constitute a less sensitive indicator of impaired system performance. However a concentration that is too low would also be suboptimal because this would lead to on/off observations rather than a more robust performance required for the calculation of meaningful performance statistics on the QC metrics.

Specifically the two QC samples that we suggest are as follows:

1. A sensitivity analysis, typically based on measuring a small amount of a well-defined sample such as, e.g., 1 fmol of BSA, 2.5 fmol of a mixture of 10 phosphopeptides, and 500 fmol of cytochrome *c* over a short, e.g., a 0.5 h gradient time (see Supplementary Materials and Methods section of the Supporting Information). We suggest to measure this “QC 1” sample whenever there have been alterations to any part of the LC–MS system, after calibration of the MS (change to ESI source), and in general at least once or if possible several times per day.
2. A performance and speed analysis, usually based on the injection of a highly complex sample such as 0.1 μg or 1 μg of HeLa measured over a longer gradient time (for instance 3 h). This “QC 2” sample is measured at least once per week, whenever there have been alterations to any component of the LC–MS system, and before the analysis of biological samples that require high speed of acquisition and optimal overall performance.

Other types of QC samples such as, e.g., other proteins, synthetic peptides spiked at different concentrations, *Escherichia coli*, *Pyrococcus furiosus*, or yeast, are conceivable and can be configured for automatic analysis with SIMPATIQCO. A schematic overview of the SIMPATIQCO software suite is shown in Figure 1.

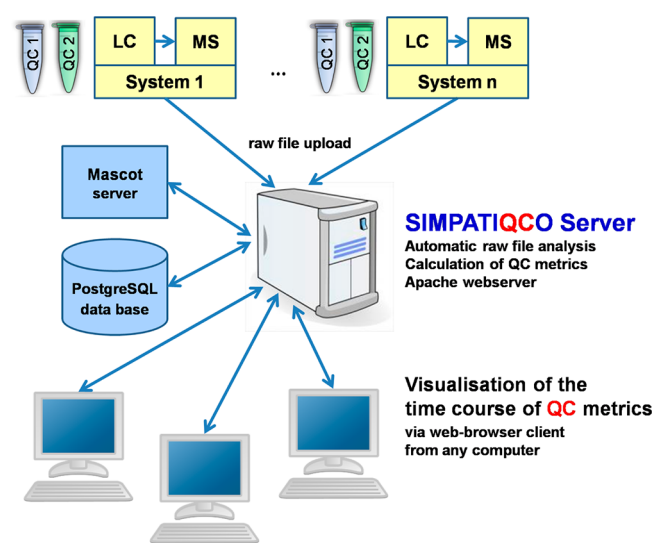


Figure 1. Components and flow of information using SIMPATIQCO. Two types of QC samples are analyzed on each LC–MS system: QC 1 contains a simple digested protein mixture (e.g., cytochrome *c*, BSA, and synthetic phosphopeptides) that is measured 1–2 \times /day for an evaluation of sensitivity and LC metrics such as peptide retention times. QC 2 is a HeLa digest that is analyzed 1–2 \times /week to evaluate system speed and overall performance. After upload of the respective raw files to the SIMPATIQCO server, QC metrics are calculated and stored in a PostgreSQL database. The database is linked to a webserver and can be queried to visualize the time course of QC metrics via a web browser.

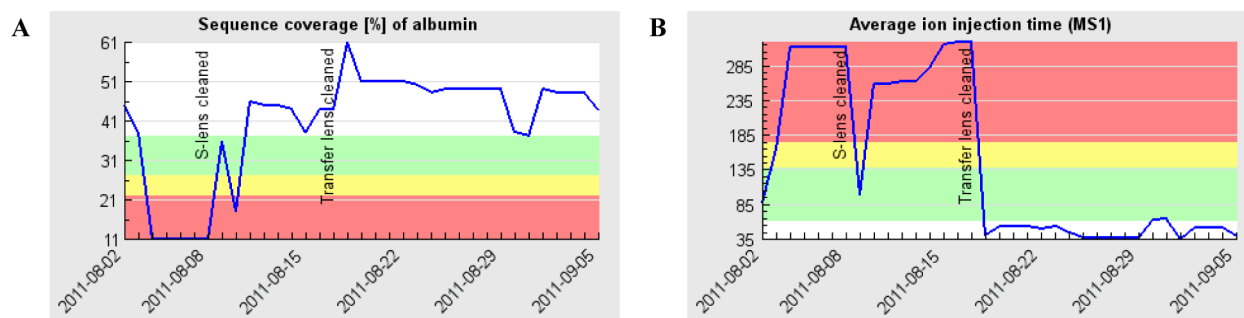


Figure 2. Example of SIMPATIQCO output: BSA sequence coverage (A) and average MS^1 ion injection times (B) on a Velos Orbitrap system monitored over time. System maintenance and service messages were entered manually by an operator and illustrate the relationship between the two QC metrics and system maintenance efforts such as venting and cleaning of the S-lens and exit lens (log entry “S-lens cleaned”) by laboratory staff (which improved BSA sequence coverage but led to an only transient reduction in average MS^1 injection times) or venting and cleaning of the transfer lens between the high pressure and the low pressure cells by a service engineer. Green band: within $1 \times$ median absolute deviation (MAD). Yellow: $1-2 \times$ MAD. Red: outside $2 \times$ MAD.

A single SIMPATIQCO server can be used to upload, analyze, and visualize data from several LC–MS systems (Figure 1 and Figure S-2 in the Supporting Information). The server can be accessed from any web browser within the local network. The four main menu categories are as follows: “Raw files” allows access to analyses of individual uploaded raw files, and manual upload and deletion of raw files. “Time course” permits visualization of the time course of QC metrics based on information stored in the database. A virtual logbook can be accessed under “Log” where the date and a short description of events such as system maintenance or cleaning of the ion optics can be entered for each system. “Settings” can be used to set the URL of the Mascot server used for searching MS/MS spectra and to configure predefined “Peptide lists” and “Proteins of interest” as well as “Run-types” and settings for the calculation of statistics from performance metrics. “Peptide lists” allow a definition of peptides whose peak elution time, apex intensity, peak width, and area can be plotted via the “Time course” menu. “Proteins of interest” define proteins for which the sequence coverage is shown when clicking on a raw file in the “Raw files” menu, and for which the time course of sequence coverage can be plotted in the “Time course” menu. “Run-types” define settings specific for a certain instrument and QC sample, such as precursor and fragment ion tolerance, Mascot “instrument type” (ESI trap, ETD trap, etc.), the FASTA database, enzyme, and Mascot peptide ion score cutoff. In addition the “Settings” menu allows defining the time period for the calculation of statistics on QC metrics (default setting: inclusion of all runs), and whether the green, yellow, and red background bands should be calculated based on a certain multiple (e.g., $1 \times$, $2 \times$, or $3 \times$) of the standard deviation or the median absolute deviation (MAD).

After data upload, SIMPATIQCO derives several QC metrics from the raw files, searches MS/MS spectra via Mascot, and then stores all extracted data and results in a database that can be queried remotely by a web browser client. This is advantageous because it provides researchers and technical personnel a convenient access to the QC data reflecting the current performance of each LC–MS system in the laboratory. In addition, SIMPATIQCO allows operators to enter “system maintenance & service messages” into a virtual logbook, for instance when the S-lens was cleaned or when a service engineer performed system maintenance work on the instrument. The respective messages can be superimposed on the time course plots so that the effect of system maintenance can be

related to changes in the QC metrics (Figure 2). In this way, SIMPATIQCO can help identify impaired system performance, monitor whether system maintenance procedures were successful, and thus help achieve and maintain close to optimum LC–MS system performance to ensure high quality data acquisition.

For each QC metric, SIMPATIQCO learns the range reflecting adequate or inadequate system performance from the uploaded data. This requires that the LC–MS system was operated under adequate or possibly optimum conditions for the majority of QC runs. Via the SIMPATIQCO “Raw files” menu, “bad” runs can be deleted manually. However it may be difficult to judge which data points actually represent outliers. We therefore decided to implement robust statistics to find the range for each metric that indicates adequate system performance, calculating median and MAD (median absolute deviation) values as an alternative to average values and standard deviation. Contrary to average values and standard deviations that are sensitive to outliers, median values and MAD provide a more robust estimation of adequate system performance that is insensitive to outliers. This strategy therefore renders manual outlier removal unnecessary. The default setting is that these statistics are calculated for all LC runs of a certain run type that have been uploaded to the database. However as soon as a time period that likely reflects “optimum system performance” has been identified, operators may define and configure this period in the “Settings” menu as a fixed reference time for calculation of statistics of QC metrics such as median, MAD, average values, and standard deviation. In this way, the above-mentioned statistics reflect the ranges during “optimum system performance”. For longitudinal monitoring of the time course of QC metrics, the values of the respective metric are plotted against a colored background where a green color band indicates optimum system performance (median \pm MAD) whereas deviant system performance is highlighted by yellow (between 1 and $2 \times$ MAD) or red color (outside $2 \times$ MAD). These plots of the time course of QC metrics allow convenient and rapid pinpointing of adequate or likely suboptimal system performance.

SIMPATICO automatically calculates and displays the following QC metrics:

1. Metrics for which the time course can be visualized (Figure 2, Figures S-1 and S-4 in the Supporting Information)
 - Number of MS^1 and MS/MS scans
 - MS^1 and MS/MS scans where lock-mass was detected

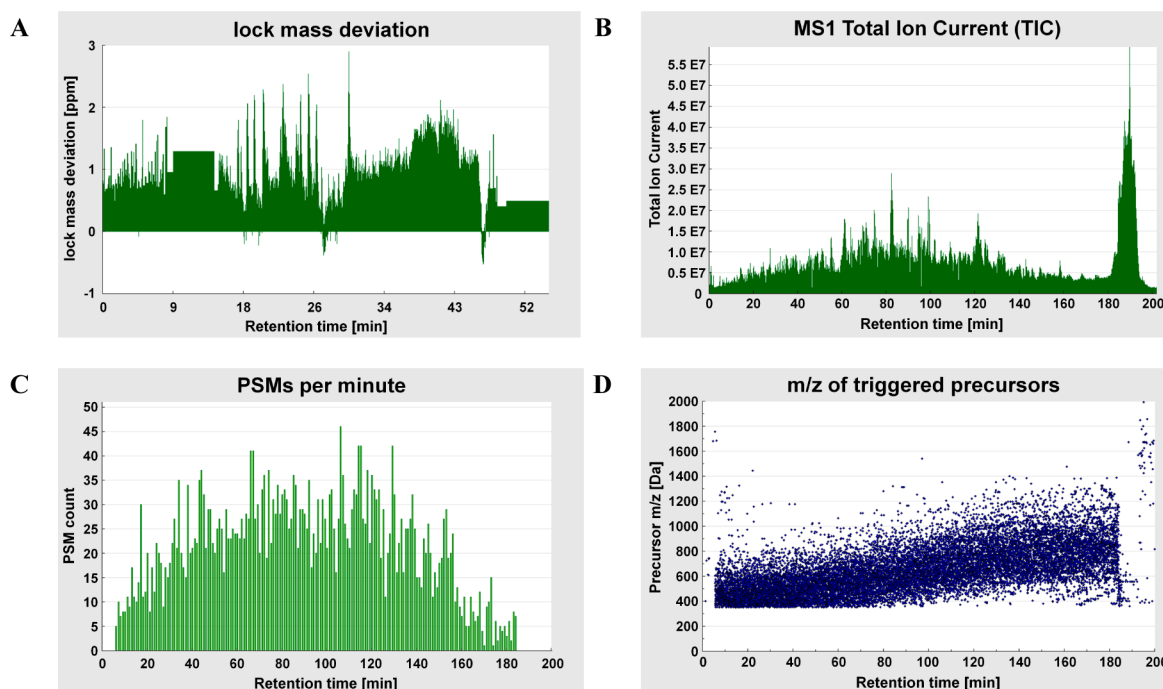


Figure 3. Lock mass deviation from a QC 1 sensitivity run (CID): peculiar pattern, plateaus indicate that lock mass was not detected (A). TIC from a QC 2 performance run (ETD): considerable fraction of TIC between 180 and 200 min (B). Different QC 2 performance run (ETD): PSMs per minute (C) and m/z of triggered precursors (D) plotted over the entire retention time range of the raw file.

- MS¹ and MS/MS scans where maximum injection time was reached
- Average MS¹ lock mass deviation
- Average MS¹ and MS/MS ion injection time and total scan time
- Total number of identified PSMs and proteins
- Sequence coverage of proteins that match an entry in the “proteins of interest” list
- For definable “peptide lists” (e.g., for QC 1 sample: cytochrome *c* peptides and synthetic phosphopeptides): peptide peak elution time, apex intensity, peak width and area

2. Metrics that can be plotted over the retention time of individual raw files (Figure 3)

- MS¹ and MS/MS ion injection time and scan time
- Lock mass detection and deviation
- MS¹ and MS/MS TIC
- MS¹ and MS/MS “(%) of target value”
- Number of PSMs identified per minute
- m/z of triggered precursors

The above-mentioned metrics are calculated and can be visualized for both types of QC runs. However, we define peptide lists only for the “sensitivity” QC 1 sample because the calculation of retention times is more precise for this sample. Taken together, these QC metrics provide a quick overview and help in monitoring the entire LC–MS system, including performance of the LC, electrospray, and MS.

Metrics Monitoring LC

LC performance can be evaluated via the peak apex retention times and intensities, peak widths, and peak areas of peptides on definable “peptide lists” (Figure S-1 in the Supporting Information). These metrics are mandatory to monitor

particularly when reproducible retention time and constant ionization efficiency are required, e.g., for label-free quantification.

There are myriads of reasons for shifts in retention times such as alterations in solvent solutions, temperature shifts, dead volume, and leaking connections. In general malfunctioning of almost every part of the LC may lead to a change in retention times; for this reason this metric is a valuable measure of LC performance. Retention time shifts may also indicate that the LC fails to generate an adequate flow rate, e.g., because of partial occlusion of the trap column or the analytical column that the system cannot compensate. In these cases, a change in the pressure curves on the LC can confirm the cause. The large amount of cytochrome *c* (500 fmol) in the QC 1 sample permits a rapid inspection of cytochrome *c* peptide retention times with LC software whenever a UV detector is used. The “peptide list” feature of SIMPATIQCO allows monitoring of this metric even without a UV detector directly from mass spectrometry data. Additional usage of a UV detector may still be helpful, e.g., for the identification of peak broadening due to dead volume in fluidic connections between the UV cell and the electrospray emitter. Moreover “peptide lists” permit monitoring the retention time, e.g., of the phosphopeptides in the QC 1 sample whose concentration would be too low to be detectable on the UV trace.

In addition, peak widths can be plotted for peptides on a SIMPATIQCO “peptide list”. Again there are many causes for peak broadening such as the presence of post-column dead volume or a deterioration of the analytical column. In the absence of a UV detector, SIMPATIQCO “peptide lists” provide a possibility to monitor peak widths, which are an excellent indicator of LC system performance.

Monitoring peptide peak intensities may also provide an indirect indication of impaired LC performance and peak broadening that typically goes along with a lowering of peak intensities. Alternatively, decreased peak intensities and peak areas may suggest either an incorrect amount of injected sample

or sample loss on the part of the LC, or a problem on the part of the MS such as less effective ionization or loss of ions before reaching the detector or loss of detector sensitivity.

Metrics Monitoring Electrospray

Electrospray quality and stability can be evaluated indirectly by its impact on MS¹ and MS/MS ion injection times, number of acquired MS¹ and MS/MS scans, and suppressed (or absent) peak intensities and areas of peptides on SIMPATIQCO “peptide lists”. All these metrics can be plotted over time to help provide information including the spray quality and stability. In addition, MS¹ and MS/MS TIC as well as the ratio “(%) of target value” can be plotted over the retention time for individual raw files. High ion injection times, low number of MS¹ and MS/MS scans, a low TIC, and the presence of fluctuations of TIC between MS¹ scans can suggest problems with electrospray, and will often prompt an exchange of the electrospray emitter. High injection times, low number of MS¹ and MS/MS scans, and low TIC can also be a sign of electron multiplier gain calibration being out of calibration.

Metrics Monitoring MS Sensitivity and Overall Speed and Performance

For the QC 1 sensitivity sample that includes cytochrome *c*, BSA, and 10 phosphopeptides, plotting the sequence coverage of BSA is recommended in the first place, as this value reflects overall LC–MS system sensitivity. The status of a system may be considered as appropriate for an analysis of, e.g., routine samples (such as the identification of proteins from immunoprecipitation or tandem affinity purification experiments) when the BSA sequence coverage is adequate. QC 1 samples can be measured with disparate types of fragmentation techniques, e.g., collision-induced dissociation (CID), electron transfer dissociation (ETD), or higher energy collisional dissociation (HCD) depending on the methods used in the ensuing regular samples.

Time course plots suggest that the average MS¹ ion injection time (calculated over the entire QC 1 run) may also constitute a sensitive indicator of MS system performance. An increase of the average MS¹ ion injection time can be reverted by appropriate actions such as instrument calibration, or cleaning of the contaminated parts of the ion optics (Figure 2). In a series of several QC 1 runs the average MS¹ ion injection time already increased while the BSA sequence coverage still remained within the range of stochastic variation, suggesting that in certain situations the average MS¹ ion injection time may constitute an even more sensitive indicator of subtle changes before a drop in sensitivity becomes apparent (Figure S-3 in the Supporting Information). We therefore monitor both BSA sequence coverage and the MS¹ ion injection time closely to schedule the appropriate maintenance or MS system tests for an evaluation of possible contamination of the ion path when the average MS¹ ion injection time starts to rise or when the BSA sequence coverage drops below normal.

As QC 1 sensitivity samples are analyzed once or even more than once per day, the metrics derived from these samples reflecting LC performance, electrospray, and MS sensitivity are monitored closely.

The QC 2 speed and performance sample, a HeLa digest, is analyzed at least once per week and in addition before an analysis of complex samples that require optimum system speed. Similar to QC 1 samples, QC 2 runs can be performed separately for each type of fragmentation method used for measuring regular samples on the respective instrument (e.g., CID, ETD, or HCD). The QC 2 sample is a complex protein digest, providing complementary information with regard to

system acquisition speed and performance (Figure S-1 A in the Supporting Information). The number of peptide-spectrum matches and identified proteins indicate whether the LC–MS system is in an appropriate condition for the analysis of complex samples. Hence these numbers constitute a sensitive indicator of overall LC–MS system performance for such samples.

MS¹ and MS/MS ion injection times can also be plotted over the retention time of a QC 2 raw file to determine how frequently the ion injection time reached the maximum value, suggesting that the AGC target value might not have been reached for the respective scan. In addition SIMPATIQCO calculates “(%) of target value”, an estimator of the number of ion charges present in each scan divided by the AGC target value. More information on this metric can be found in the Supporting Information.

Finally, the average lock-mass deviation may serve as an indicator when the next mass calibration should be scheduled.

QC runs are repeated when outliers are observed so that impaired system performance should become apparent as a series of outliers. Whenever the pattern of the time course of QC metrics suggests suboptimal performance, an investigation of possible causes is undertaken. SIMPATIQCO also assists with this task. For instance, one may wish to plot the lock mass¹⁷ deviation, MS¹ TIC, the number of PSMs identified per minute, and the *m/z* of triggered precursor ions over the course of the QC run (Figure 3).

Although SIMPATIQCO can assist in finding the reason of suboptimal system performance, in many situations LC-specific and MS-specific system evaluation tests need to be performed to find out the actual cause. In severe cases, prolonged MS¹ ion injection times are observed even when spraying calibration mixture solution with the ESI source mounted. SIMPATIQCO also offers the possibility to upload such data, for instance 100 off-line acquired MS¹ scans of a calibration mixture. A special type of QC-sample “run-type” can be defined for such data and configured so that the data is not searched with Mascot. When such data sets are uploaded on a regular basis, SIMPATIQCO allows convenient monitoring of the time course of average ion injection times of MS¹ scans of a calibration mixture acquired off-line with the ESI source, which should be below 0.1 ms on Velos Orbitrap instruments (Figure S-4 in the Supporting Information).

Among MS-specific system evaluation procedures, the so-called “Multipole Flight Time evaluation” test, now called the “Source Optics Flight Time evaluation” test, along with the “MP0 Flight Time evaluation” test, can help provide an indication as to whether there is a need to clean some element of the source ion optics (including the S-Lens, Exit Lens, MP00, Lens 0, or MP0). In cases where cleaning of the source optics does not lead to a sustained improvement of the results of the QC metrics monitored by SIMPATIQCO, other parts of the ion optics can be considered as a cause. In this case another system evaluation procedure called the “Ion Optics Charging Evaluation” can help pinpoint a contaminated device along the entire optics path including the Exit Lens, MP00, Lens 0, MP0, Lens1, MP1, Gate Lens, Front Lens, and Center Lens, by sequential exposure to a negative ion beam. The Center Lens between the high and low pressure cells can also be evaluated via the “transfer efficiency evaluation” or “transfer lens calibration check”. Notably, appropriate tuning of multipole offsets particularly with regard to MP0 can have a strong effect on resisting contamination effects. With Tune 2.7 SP1, inadvertent setting of low voltages due to inappropriate tuning is prevented to

ensure robust ion transmission even in case of moderate contamination.

The first samples analyzed after starting the LC–MS system again following troubleshooting and maintenance procedures should always be further QC samples. At times, more than one cleaning approach may be required until the source of impaired system performance can be identified and resolved (Figure 2), which should then become apparent as long-term sustained optimal performance, e.g., with regard to BSA sequence coverage and MS¹ ion injection times and other performance metrics.

CONCLUSION

We developed SIMPATIQCO (SIMple AuTomatIc Quality COntrol), a software that aids the successive and routine surveillance of sensitivity, speed, and other QC performance metrics of LC–MS systems. The software is designed for use with LTQ Orbitrap, Q-Exactive, LTQ FT, and ion trap instruments manufactured by Thermo Scientific. SIMPATIQCO incorporates data extraction, MS/MS spectra search, and storage of results within a PostgreSQL database coupled to an Apache webserver, as well as calculation of QC metrics that can be visualized remotely from a web browser client. Taken together the available QC metrics allow monitoring of each step of data acquisition. The time course of QC metrics can be plotted over a background colored in green, yellow, or red reflecting optimal or impaired system performance as learned from the uploaded data by calculation of robust statistics (Figure 2 and Figure S-1 in the Supporting Information). This helps operators identify suboptimal LC–MS system performance, prompting an immediate system workup. SIMPATIQCO also offers further features such as plots of the number of PSMS per minute, lock mass deviation, peptide elution times, apex intensities, peak widths, and areas that aid operators in tracking possible causes (Figure 3 and Figures S-1 and S-4 in the Supporting Information). The benefit of monitoring the time course of QC metrics currently becomes more widely recognized. For instance the Association of Biomolecular Resource Facilities (ABRF) PRG 2012 study aims at evaluating within-laboratory instrument reproducibility (study announcement on the webpage <http://www.abrf.org/>). SIMPATIQCO may be used by research laboratories and core facilities alike to achieve this goal. The software package is available free of charge and can be downloaded from the webpage <http://ms.imp.ac.at/> together with installation instructions and a manual. We hope that SIMPATIQCO will help proteomics laboratories achieve and maintain optimal quality of proteomics data acquisition.

ASSOCIATED CONTENT

Supporting Information

Supplementary figures and supplementary materials and methods. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*P.P.: Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, 1030 Vienna, Austria; e-mail, peter.pichler@imp.ac.at; phone, +43-1-79044-4293; fax, +43-1-79044-110. K.M.: Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, 1030 Vienna, Austria; e-mail, karl.mechtler@imp.ac.at; phone, +43-1-79044-4280; fax, +43-1-79044-110.

Author Contributions

[§]Equally contributing first authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Jae Schwartz, Jean-Jacques Dunyach, and Andreas Huhmer (Thermo Scientific) for helpful suggestions, explanations, and critical review of the manuscript. Dorothea Anrather, Rainer Gith, and Gustav Ammerer kindly provided data for one of the four participating laboratories, using the SIMPATIQCO server operated by the IMP. In addition we are grateful to our colleagues Susanne Opravil, Ines Steinmacher, Johannes Fuchs, Michael Schutzbier, Gabriela Krssakova, and Mathias Madalinski for technical assistance. We also thank Andreas Schmidt, Nikolai Mischerikow, Thomas Taus, Johannes Stadlmann, Elisabeth Roitinger, and Christoph Jüschke for useful discussions and critical reading of the manuscript. This work was supported by the Christian Doppler Research Association CDG (Christian Doppler Forschungsgesellschaft), the Austrian Science Foundation FWF (Fonds zur Förderung der wissenschaftlichen Forschung) SFB-F3402 Chromosome Dynamics, and the European Commission via the FP7 projects MEIOsys and Prime-XS.

REFERENCES

- (1) Martens, L.; Vizcaino, J. A.; Banks, R. Quality control in proteomics. *Proteomics* **2011**, *11* (6), 1015–6.
- (2) Kinsinger, C. R.; Apffel, J.; Baker, M.; Bian, X.; Borchers, C. H.; Bradshaw, R.; Brusniak, M. Y.; Chan, D. W.; Deutsch, E. W.; Doman, B.; Gorman, J.; Grimm, R.; Hancock, W.; Hermjakob, H.; Horn, D.; Hunter, C.; Kolar, P.; Kraus, H. J.; Langen, H.; Linding, R.; Moritz, R. L.; Omenn, G. S.; Orlando, R.; Pandey, A.; Ping, P.; Rahbar, A.; Rivers, R.; Seymour, S. L.; Simpson, R. J.; Slotta, D.; Smith, R. D.; Stein, S. E.; Tabb, D. L.; Tagle, D.; Yates, J. R.; Rodriguez, H. Recommendations for mass spectrometry data quality metrics for open access data (corollary to the Amsterdam Principles). *J. Proteome Res.* **2012**, *11* (2), 1412–9.
- (3) Kocher, T.; Pichler, P.; Swart, R.; Mechtler, K. Quality control in LC-MS/MS. *Proteomics* **2011**, *11* (6), 1026–30.
- (4) Bell, A. W.; Deutsch, E. W.; Au, C. E.; Kearney, R. E.; Beavis, R.; Sechi, S.; Nilsson, T.; Bergeron, J. J. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nat. Methods* **2009**, *6* (6), 423–30.
- (5) Nesvizhskii, A. I. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *J. Proteomics* **2010**, *73* (11), 2092–123.
- (6) Granholm, V.; Kall, L. Quality assessments of peptide-spectrum matches in shotgun proteomics. *Proteomics* **2011**, *11* (6), 1086–93.
- (7) Vaudel, M.; Burkhart, J. M.; Sickmann, A.; Martens, L.; Zahedi, R. P. Peptide identification quality control. *Proteomics* **2011**, *11* (10), 2105–14.
- (8) Rudnick, P. A.; Clauser, K. R.; Kilpatrick, L. E.; Tchekhovskoi, D. V.; Neta, P.; Blonder, N.; Billheimer, D. D.; Blackman, R. K.; Bunk, D. M.; Cardasis, H. L.; Ham, A. J.; Jaffe, J. D.; Kinsinger, C. R.; Mesri, M.; Neubert, T. A.; Schilling, B.; Tabb, D. L.; Tegeler, T. J.; Vega-Montoto, L.; Variyath, A. M.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Carr, S. A.; Fisher, S. J.; Gibson, B. W.; Paulovich, A. G.; Regnier, F. E.; Rodriguez, H.; Spiegelman, C.; Tempst, P.; Liebler, D. C.; Stein, S. E. Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Mol. Cell. Proteomics* **2010**, *9* (2), 225–41.
- (9) Ma, Z. Q.; Polzin, K. O.; Dasari, S.; Chambers, M. C.; Schilling, B.; Gibson, B. W.; Tran, B. Q.; Vega-Montoto, L.; Liebler, D. C.; Tabb, D. L. QuaMeter: Multivendor Performance Metrics for

LC-MS/MS Proteomics Instrumentation. *Anal. Chem.* **2012**, *84* (14), 5845–50.

(10) Matzke, M. M.; Waters, K. M.; Metz, T. O.; Jacobs, J. M.; Sims, A. C.; Baric, R. S.; Pounds, J. G.; Webb-Robertson, B. J. Improved quality control processing of peptide-centric LC-MS proteomics data. *Bioinformatics* **2011**, *27* (20), 2866–72.

(11) Sweredoski, M. J.; Smith, G. T.; Kalli, A.; Graham, R. L.; Hess, S. LogViewer: A Software Tool to Visualize Quality Control Parameters to Optimize Proteomics Experiments using Orbitrap and LTQ-FT Mass Spectrometers. *J. Biomol. Tech.* **2011**, *22* (4), 122–6.

(12) Scheltema, R. A.; Mann, M. SprayQc: A Real-Time LC-MS/MS Quality Monitoring System To Maximize Uptime Using Off the Shelf Components. *J. Proteome Res.* **2012**, *11* (6), 3458–66.

(13) Kocher, T.; Pichler, P.; Swart, R.; Mechtler, K. Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. *Nat. Protoc.* **2012**, *7* (5), 882–90.

(14) Olsen, J. V.; Ong, S. E.; Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–14.

(15) Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.; Wouters, E. R.; Senko, M.; Makarov, A.; Mann, M.; Horning, S. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol. Cell. Proteomics* **2009**, *8* (12), 2759–69.

(16) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–72.

(17) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **2005**, *4* (12), 2010–21.