

---

---

# Noise Filtering Techniques for Electrospray Quadrupole Time of Flight Mass Spectra

Jürgen Kast

The Biomedical Research Center, University of British Columbia, Vancouver, British Columbia, Canada

Marc Gentzel and Matthias Wilm

EMBL, Heidelberg, Germany

Keith Richardson

Micromass, Manchester, United Kingdom

---

The sensitivity of protein identification by peptide sequencing using a nanoelectrospray ion source is limited by our ability to identify peptide ions in the mass spectrum. Their intensity must be higher than the chemical noise level to allow a rapid localization in the spectrum. Multiply-charged peptide ions on or below this level can only be found because their isotopic pattern is denser than that of the mostly singly-charged chemical background ions. However, to find peptides by looking for multiply-charged isotope clusters can be very timeconsuming and may lead to misassignments of the first isotope. Here we present a software-based method to increase the signal to noise ratio of ion signals in an electrospray spectrum. The software has two elements, one to reduce the noise level and a second to increase the intensity of ion peaks. Both methods together generate a spectrum in which the signal to noise ratio of ion signals is considerably improved. Peptide ions previously hidden in the chemical background are dismantled and can now be localized and selected for fragmentation. The method has been used successfully to identify low level proteins separated by one dimensional gel electrophoresis. (J Am Soc Mass Spectrom 2003, 14, 766–776) © 2003 American Society for Mass Spectrometry

---

---

Protein sequencing by mass spectrometry is a key technology for biological research [1–3]. Large scale protein identifications by tandem mass spectrometry are mostly done by the direct analysis of protein mixtures after digestion with on-line coupled capillary HPLC quadrupole time of flight or ion trap mass spectrometers [4–8]. For the sequencing of an individual, gel separated protein, a nanoelectrospray-based approach is still a very sensitive, fast, and reliable method. When working with affinity purified proteins it often happens that the amount of protein available is so low that the sensitivity of the mass spectrometric analysis is limiting. In our hands the highest sensitivity is achieved by a nanoelectrospray approach analyzing the unseparated peptide mixture [9, 10]. The peptides extracted from the in-gel digest are exposed to the smallest possible surface and the smallest possible amount of reverse phase material. They are eluted under acidic conditions in at least 50% organic phase directly into the nanoelectrospray emitter. The elution

volume is often smaller than 1  $\mu$ l, which is nevertheless several times the column volume. Inability to identify peptide ions in the mass spectrum limits the overall sensitivity. Once an  $m/z$  value of a multiply-charged peptide is assigned, a fragment spectrum can be acquired even if the signal strength of the original ion is on or below the noise level. Since the chemical noise in an electrospray spectrum is composed mostly of singly-charged ions, fragments above the  $m/z$  value of a multiply-charged precursor will be clearly visible. Using the sequence tag algorithm for protein identification, these fragments can be sufficient to find the underlying peptide sequence in a database search [11]. Therefore the main concern when sequencing small quantities of proteins with nanoelectrospray is finding the ion signals of multiply-charged peptides in the mass spectrum.

On our triple quadrupole machine we use precursor ion scans for the immonium ion of isoleucine and leucine for this purpose with considerable success [12]. Peptides containing these common amino acids can be detected with very high sensitivity since background ions do not produce this fragment with any preference. Unfortunately, precursor ion scans are not available with the same sensitivity on quadrupole time of flight

---

Published online May 29, 2003

Address reprint requests to Dr. M. Wilm, Bioanalytical Research Group, EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. E-mail: wilm@embl-heidelberg.de

mass spectrometers because of their different instrument geometry. Here, we present a software-based method to increase the signal to noise ratio of ions in electrospray quadrupole time of flight spectra. The principal purpose of this spectrum processing is to enable a rapid identification of precursor ions and their charge states for further analysis by tandem mass spectrometry. To achieve this we chose to accept a limited number of artifacts from the data processing since the core of the analysis is the interpretation of the fragment spectrum, which remains unaffected.

## Experimental

### Chemicals

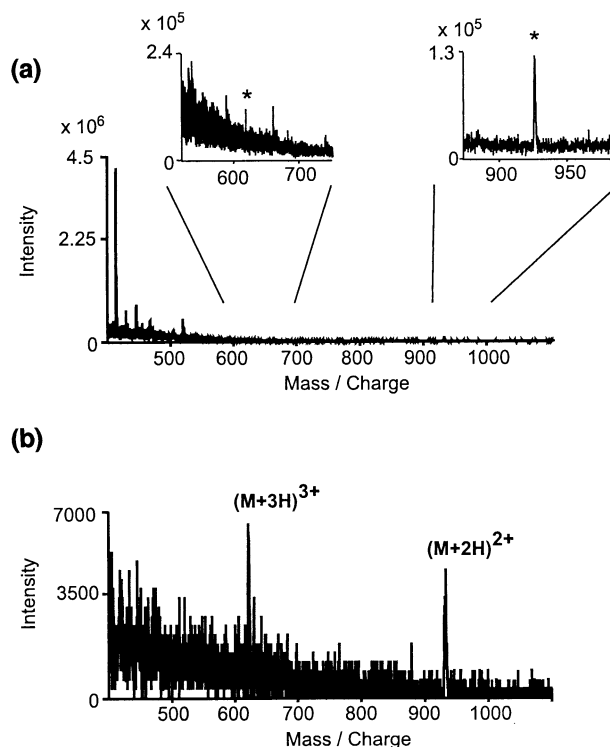
All chemicals used in this study were of analytical grade.  $^{18}\text{O}$ -labeled water was bought from Phychem (Düren, Germany) and purified before use by a self-made micro-distillation device. The protein was digested with unmodified trypsin from Boehringer Mannheim (Mannheim, Germany). Synthetic peptides were made in house.

### Sample Preparation

The protein band was cut from a one dimensional gel, reduced, alkylated, and in-gel digested using trypsin. For de novo sequencing the digestion buffer contained 50%  $^{18}\text{O}$  labeled water to partially incorporate a heavy oxygen isotope at the C-terminus of each peptide. Under the conditions chosen, no double incorporation of  $^{18}\text{O}$  atoms into peptides was observed (<5%). After overnight digestion the peptides were extracted and lyophilized [10]. For analysis, the peptide mixture was dissolved in 1  $\mu\text{l}$  80% formic acid and rapidly diluted with 9  $\mu\text{l}$  water. The peptide mixture was desalted on 100 nl of self-assembled Poros R2 and R3 columns in pulled glass capillaries and eluted directly into a nano-electrospray needle using two times 0.3  $\mu\text{l}$  60% methanol 5% formic acid for mass spectrometric analysis [9, 10]. All other samples were dissolved in 50% acetonitrile, 5% formic acid, and 45% water.

### Mass Spectrometry

Peptides were analyzed on a Q-TOF1 mass spectrometer equipped with a nano-electrospray ion source (Micromass, Manchester, UK). The interface temperature was set to 50 °C to limit the evaporation rate of the solvent inside the needle. After acquisition of a mass spectrum the spectrum was exported from MassLynx (Micromass, Manchester, UK) as an ASCII file with no processing applied and then filtered by a procedure programmed in IGOR Pro (Wavemetrics, Lake Oswego, OR). Subsequently, the processed mass spectrum was reimported into MassLynx to assign manually peptides of interest. Finally, fragment spectra of peptides that were absent from the control sample were acquired.



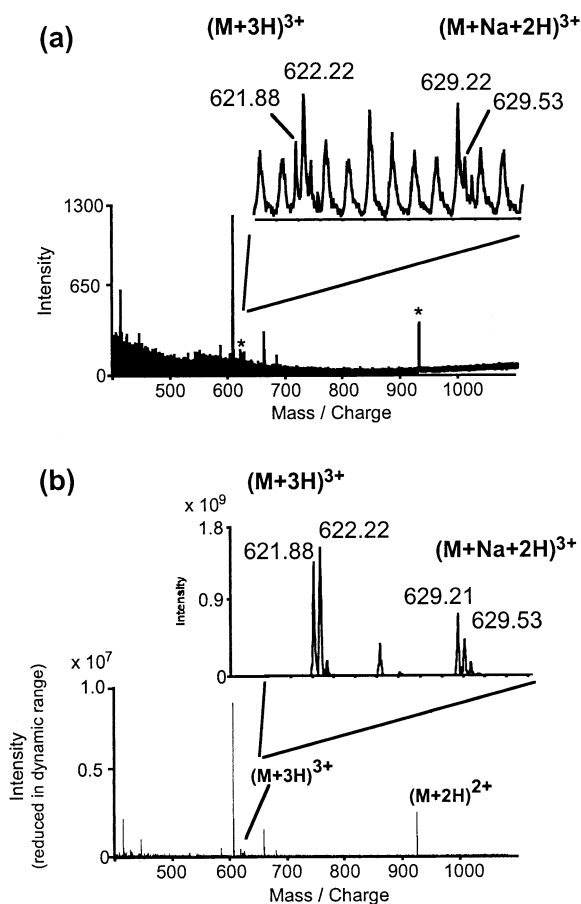
**Figure 1.** Nanoelectrospray spectrum and precursor ion scan spectrum of the peptide TPPAVLSDSGSYFLYSK on a triple quadrupole mass spectrometer. The peptide has a concentration of 10 fmol/ $\mu\text{l}$ . (a) shows the mass spectrum, (b) the precursor ion scan spectrum on 86 Da, the immonium ion of isoleucine and leucine. The precursor ion scan spectrum demonstrates a higher specificity for the detection of the peptide. The overall intensity is reduced, but the signal to noise ratio is improved for the triply-charged ion.

### Data Processing

Mass spectra were exported from MassLynx in ASCII format, processed by IGOR Pro and reimported into MassLynx to identify the peptide ions. In the IGOR Pro procedure, the imported mass spectrum was interpolated to obtain an equally-spaced dataset with 0.01 Th step width. For noise filtering, the spectrum was Fourier transformed using an IGOR Pro Fast Fourier Transformation algorithm. Maxima in the Fourier spectrum were eliminated and the mass spectrum was regenerated by a reverse Fast Fourier Transformation function in IGOR Pro. Negative intensities were discarded. For peak enhancement, the spectrum is interpolated with a 0.01 Th step width, increased by 1 to avoid multiplications by 0 and autocorrelated using Eq 1 (see below). The processed mass spectrum was saved as an ASCII file and reimported into MassLynx on a PC and into Bio-MultiView (Sciex, Toronto, Canada) on a Macintosh.

## Results and Discussion

The sensitivity of protein identification from unseparated peptide mixtures using a nano-electrospray-based



**Figure 2.** Nano-electrospray spectrum of the peptide TTPAV-LDSDGSYFLYSK on a quadrupole time of flight mass spectrometer. The peptide has a concentration of 10 fmol/ $\mu$ l, the same as in Figure 1. (a) shows the mass spectrum, (b) the processed spectrum after two algorithms have been applied to improve the signal to noise ratio of ion peaks. The signal to noise ratios of the doubly and triply-charged peptide ions in the unprocessed spectrum are very similar to those in the spectrum acquired with the triple quadrupole mass spectrometer. The doubly-charged peptide is clearly visible whereas the intensity of the triply-charged one is only slightly higher than the chemical background noise. It can be assigned in the original spectrum only because of its denser isotopic spacing. The noise is completely removed in the processed spectrum and all isotopes of the peptide and its sodium adduct are clearly visible. In contrast to highly sensitive precursor ion scans on triple quadrupole mass spectrometers the resolution has not been compromised by the procedure.

approach is limited by the ability to identify peptide ions in the mass spectrum. On triple quadrupole machines, precursor ion scans have been used very successfully to locate isoleucine or leucine containing peptides in the spectrum [12]. Figure 1 gives an example of the efficiency of this approach. The peptide TTPAV-LDSDGSYFLYSK was diluted to a concentration of 10 fmol/ $\mu$ l in 50% acetonitrile, 5% formic acid. The peptide is visible as a doubly and triply-charged ion in the MS scan and the precursor ion scan for 86 Da. But in the latter, the signal to noise ratio of the triply-charged ion is higher and the precursor scan demonstrates a high specificity for the peptide excluding singly-charged

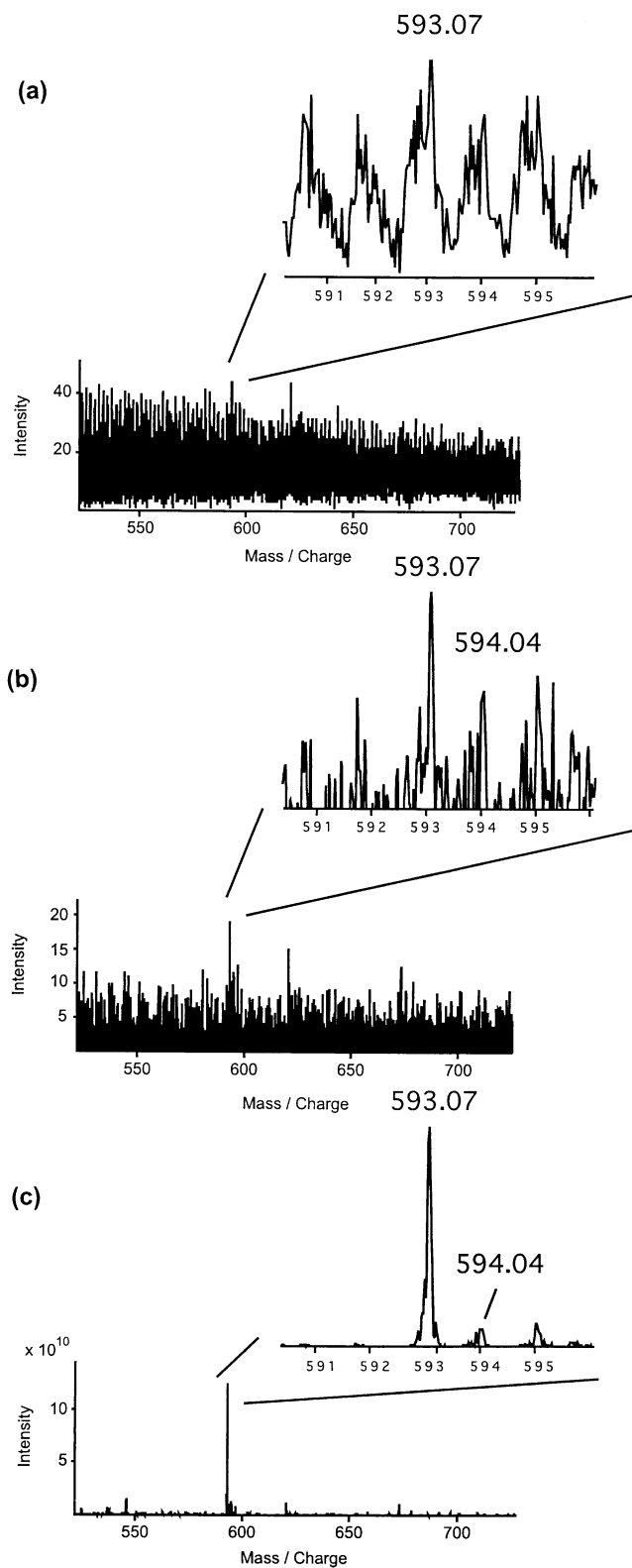
contaminating ions. The doubly-charged peptide ion has roughly the same signal to noise ratio in the MS scan and the precursor ion scan. At its relatively high  $m/z$  value of 932.2 Thomson (Th) the chemical background noise is lower than in the more typical region for tryptic peptides between 400 and 700 Th. At 932 Th the filtering effect of the precursor ion scan is less pronounced. Using precursor ion scans on 86 Da it is possible to localize peptides very efficiently on the  $m/z$  scale when using triple quadrupole machines.

On our quadrupole time of flight mass spectrometer the precursor ion scan is not available with the same sensitivity. Therefore we investigated other possibilities to increase the signal to noise ratio of peptides in the spectrum. Figure 2 demonstrates the final result of a software-based method. The spectrum was generated from the same 10 fmol/ $\mu$ l solution as the spectra in Figure 1. The triply-charged ion and its sodium adduct are isotopically resolved and cleaned of interference from background ions. Figure 3 demonstrates the algorithms separately on a singly charged ion. The method is not based on any chemical property of the ions, so it is not specific for peptides. This is not especially helpful for protein identifications but supports its general applicability. In an experiment to identify proteins, multiply-charged precursors would be fragmented first following the experience that they most often represent peptides from the digested proteins and that their fragment spectra are more likely to lead to the identification of the protein. The filtering affects the resolution of the spectrum to only a minor degree, so multiply-charged peptides can be recognized by their isotopic spacing.

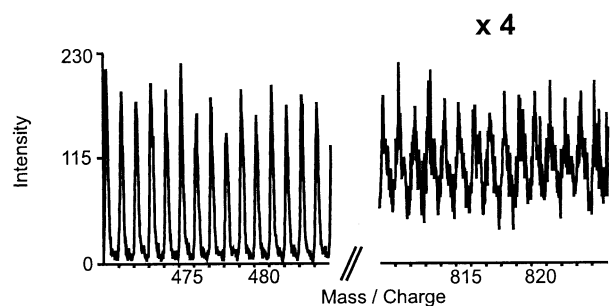
The processing of the spectrum consists of two steps: First, removal of regular noise by Fourier transform-based filtering; second, peak enhancement by correlating the spectrum with itself.

### Noise Filtering by Fourier Transformation

Fourier transform-based filtering is often used to improve the signal to noise ratio in digital datasets—in particular for image processing [13]. High pass or low pass filters can be implemented. In our application we use Fourier transform-based filtering to remove periodic noise from the mass spectrum. This noise most probably originates from solvent clusters generated by the electrospray ionization process and which survive the ion desolvation in the interface of the mass spectrometer. This noise is a typical feature of electrospray spectra. In MALDI spectra we have seen this kind of background only in spectra acquired on MALDI Q-TOF machines using 2,5-dihydroxybenzoic acid (DHB) as matrix. For these spectra a relative large laser intensity is used, which favors the generation of matrix clusters. In so far as our Fourier transform-based filter is specifically adapted to the origin of the background noise in our spectra, it is not generally applicable. But its specificity is, at the same time, its strength for our purpose of finding peptide ions for fragmentation.

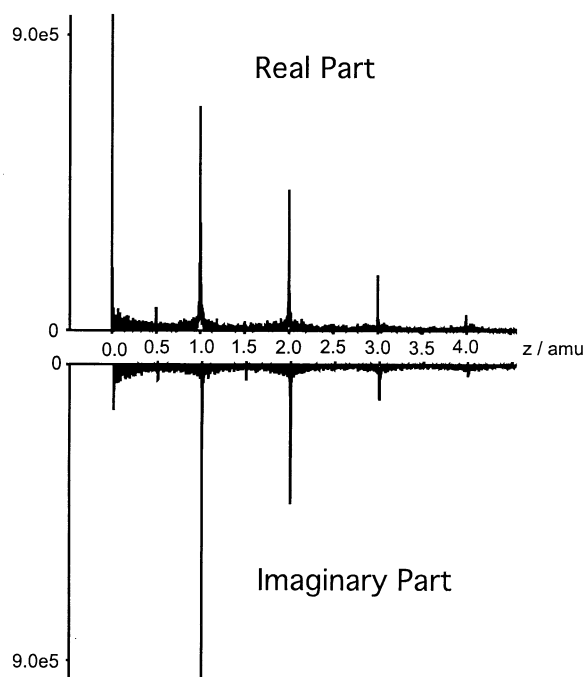


**Figure 3.** Nano-electrospray spectrum of the peptide IGETpYpGVVYK at a concentration of 10 fmol/ $\mu$ l. The effect of noise filtering and peak enhancement is demonstrated on a singly charged ion in the spectrum. (a) shows a part of the original spectrum, (b) the same part after Fourier transform based noise filtering, and (c) after peak enhancement. The first and the second isotope of the singly charged ion can only be guessed in the acquired spectrum, but they become clearly visible after the regular noise had been filtered out. The ion becomes a dominating peak after three cycles of the peak enhancement procedure. The last step distorts the isotopic ratio since intensive peaks are more enhanced than small peaks.

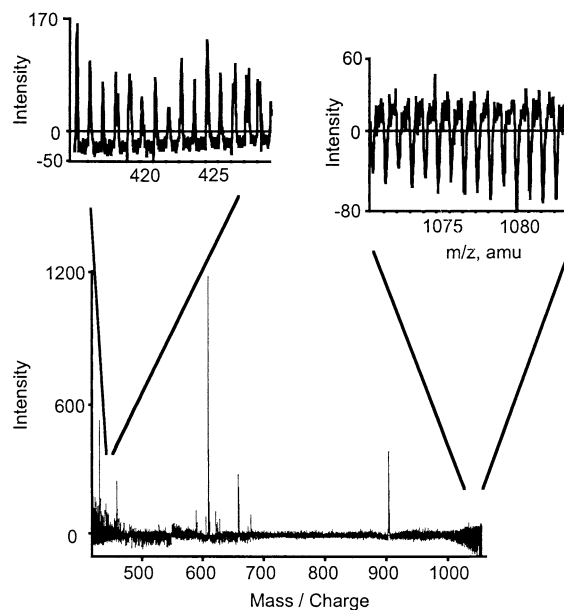


**Figure 4.** Noise characteristics in an electrospray spectrum. In the low  $m/z$  region the noise is very regular and of high intensity. In the high  $m/z$  region the regular component is still apparent but its overall intensity is reduced and an irregular contribution becomes more pronounced.

Figure 4 demonstrates the characteristic appearance of chemical noise in an electrospray quadrupole time of flight spectrum. Particularly in the low  $m/z$  region the background is very regular. The regularity is less pronounced in the high  $m/z$  region and its overall intensity decreases. Regularities in a spectrum are best seen after Fourier transformation. Figure 5 shows a typical Fou-



**Figure 5.** Fourier transformation of an electrospray spectrum. The upper panel shows the total of the real part of the Fourier transform, the lower panel the total of the imaginary part. The peaks in the spectrum make evident that there is a very regular component in the original mass spectrum occurring at every Thomson (peak at  $1 \text{ Th}^{-1}$  and its higher harmonics in the power spectrum). Often we observe additional, smaller peaks at  $(0.5 + n) \text{ Th}^{-1}$ ,  $n = 1, 2, \dots$ . Since the peptide ions are distributed in an irregular way over the spectrum these peaks must represent the chemical background noise. Their spacing indicates that the molecules or molecular clusters have multiple charge states. By selectively removing these peaks from the Fourier transform and retransforming it into the mass space, a background noise-filtered spectrum is obtained.

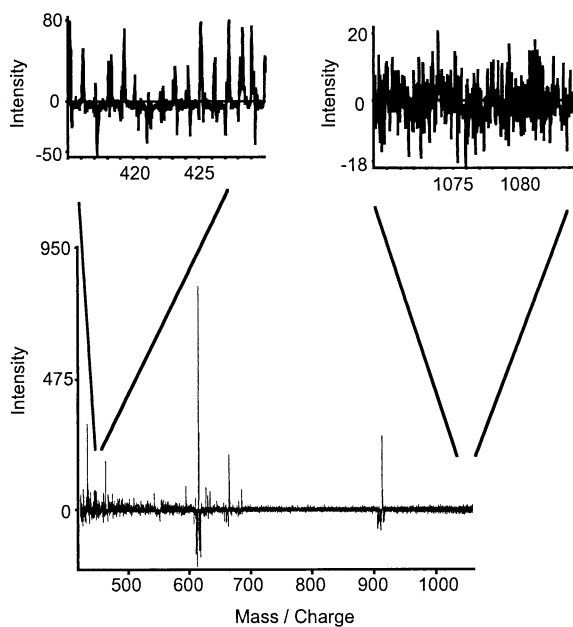


**Figure 6.** Noise-filtered mass spectrum after the major peaks of the Fourier transform are eliminated from a single Fourier power spectrum covering the entire mass spectrum. The overall intensity of the noise is reduced. However, in the low  $m/z$  region a considerable degree of regularity remains in the spectrum and in the low  $m/z$  region the noise is overcompensated and the filtering is not well adapted to the different noise characteristics.

rier transform of an electrospray spectrum. It highlights more clearly that the mass spectrum has very regular components. The regular background in an electrospray spectrum generates peaks at every Thomson (see Figure 2). This causes peaks in the Fourier transform at  $1 \text{ Th}^{-1}$  and its harmonics ( $2 \text{ Th}^{-1}$ ,  $3 \text{ Th}^{-1}$ ,  $\dots$ ) [14, 15]. Additionally, we often observed weaker peaks at  $(0.5 + n) \text{ Th}^{-1}$ ,  $n = 0, 1, 2, \dots$ . Peptide ions have an irregular distribution over the mass scale so these peaks represent regular background noise.

Our first attempt to eliminate the chemical background noise was simply to cut the peaks from the Fourier transform and do a reverse Fourier transformation to generate a filtered spectrum. This reduces the overall noise considerably, but since the noise characteristic is  $m/z$ -dependent, the adaptation to the local situation is insufficient. Figure 6 shows the result of such a procedure. The overall noise is reduced, as can be seen in a comparison with the original spectrum in Figure 2. But detailed examination reveals that in the low  $m/z$  region a regular component is left in the spectrum above the zero intensity line and that in the high  $m/z$  region the filtering is not very well adapted to the different noise characteristic.

To pay more attention to the varying noise characteristics, we chose to do Fourier transformations in separate windows, filter them, and combine the reverse Fourier transformations to get back a complete spectrum. When experimenting with different window sizes, a width of 10 Da appeared to give the best results for our spectra. A second improvement was realized by using  $|\sin(2\pi f)|^{0.25}$  as the filtering function for the



**Figure 7.** Retransformed mass spectrum after the major peaks have been smoothly removed from Fourier power spectra generated from 10 Th-wide windows of the original spectrum. In comparison to the filtered spectrum in Figure 6 the overall intensity has been reduced further. All regular features disappear from the low  $m/z$  region of the spectrum and the high  $m/z$  region does not show overcompensation or lack of adaptation of the filtering mechanism to the noise characteristics. There are slight over-compensations in the vicinity of abundant peaks, which are caused by their large contribution to the Fourier transform in the 10 Th-wide window around them. Considering the otherwise very positive effects of this filtering method to the readability of the spectrum, we accepted these limited artifacts.

Fourier transform. The sudden reduction of the intensities in the Fourier transform to 0 leads to regions of increased noise in the filtered mass spectrum (data not shown). The filtering function  $|\sin(2\pi f)|^{0.25}$  eliminates the maxima at  $0.5 \cdot n \text{ Th}^{-1}$  with  $n = 0, 1, 2, \dots$  in the Fourier transform in a smoother way. The resulting filtered mass spectrum is shown in Figure 7. The remaining noise signals in the low and the high  $m/z$  regions have a lower overall intensity in comparison with the filtered spectrum shown in Figure 6, and no regular features remain. The small window size of 10 Th leads to overcompensation of background noise in the vicinity of large peaks (see Figure 7). Choosing larger windows reduces the significance of this effect but spreads it. Only when operating with window sizes of 50 Da or more does this artifact disappear, but at the price that the adaptation to local noise characteristics is again insufficient. Considering the quite remarkable improvements in visibility of low-level ions, we accepted the localized overcompensation of noise in the vicinity of large peaks as a method-inherent artifact. Before continuing the spectrum processing, negative intensity values were discarded.

In summary, electrospray spectra can contain a considerable amount of chemical background noise. This noise is characterized by having a regular distribution

over the spectrum. This becomes obvious in the form of distinct peaks in the Fourier transform of an electrospray spectrum. When eliminating these peaks and regenerating the spectrum by reverse Fourier transformation, a filtered spectrum is obtained which no longer contains regular noise.

### Improvement of Signal to Noise by Peak Enhancement

The second type of filter that we use, the correlation, is mathematically a variation of matrix convolution filters used by James Carroll and Ronald Beavis to extract information from MALDI time of flight spectra [16]. A new spectrum is calculated from the original one. For every intensity on a  $m/z$  value in the new spectrum the entire environment of this  $m/z$  value in the original spectrum is considered. However, in our case the filter itself is spectrum-dependent and varies from point to point. We do not apply a static filter, but correlate the spectrum with itself enhancing peaks by intrinsic properties that differentiate them from random noise. We compared the noise-filtering effect with the matrix convolution filters applied by Carroll and Beavis [16]. We tried high pass, low pass and adaptive background correction filters [16]. For our data the correlation algorithm gave better results. Figure 8 gives an overview over the different methods tested. Carroll and Beavis filtered peaks from slowly varying background, but our problem was to find peaks in erratic, rapidly changing noise. The different requirements in these two cases are best met by different algorithms.

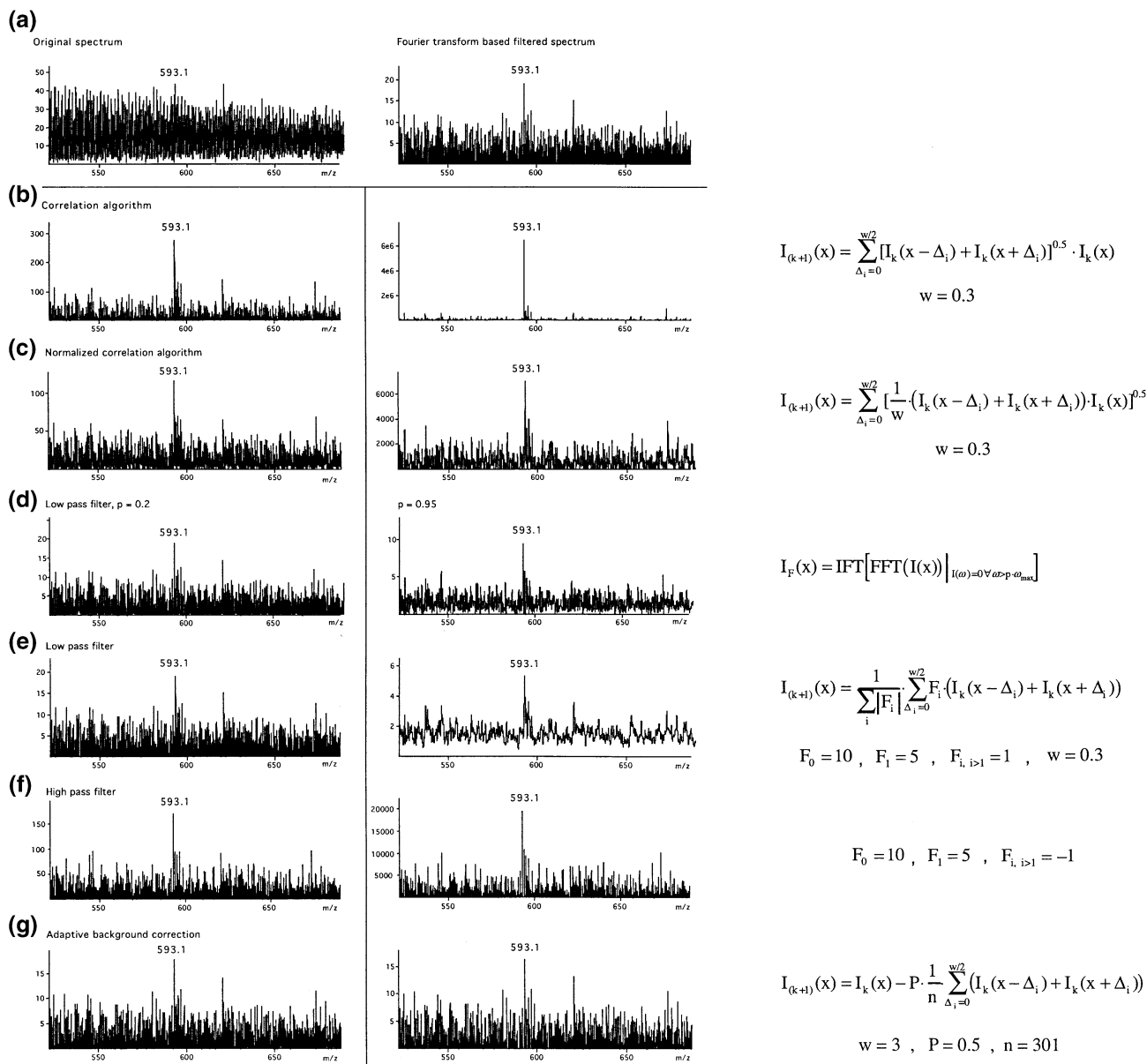
As Figure 7 demonstrates, there is some essentially random noise remaining in the spectrum after Fourier transform-based filtering. To improve the signal to noise ratio further, we chose an algorithm that enhances signals relative to random noise. A true ion peak is characterized by the fact that if there is an ion count in detection channel  $i$ , there is a higher than random probability that there are positive ion counts in the adjacent channels as well. This characteristic differentiates an ion signal from random noise. With a correlation algorithm, ion signals can be amplified out of random noise. We used the following cyclic method:

$$I_{(k+1)}(x) = \sum_{\Delta_i=0}^{w/2} [I_k(x - \Delta_i) + I_k(x + \Delta_i)]^{0.5} \cdot I_k(x) \quad (1.1)$$

or in integral form

$$I_{(k+1)}(x) = I_k(x) \cdot \frac{1}{w} \int_{x-w/2}^{x+w/2} I_k^{0.5} dx \quad (1.2)$$

where  $I_k(x)$  is the intensity of the mass spectrum at the  $m/z$  value  $x$  after  $k$  enhancement cycles.  $\Delta_i$  is a mass



**Figure 8.** Comparison of different noise filtering techniques. (a) shows the original spectrum and the spectrum after Fourier transform based filtering. (b), (c), (d), (e), (f), and (g) show the resulting spectrum after different filtering techniques have been applied to the filtered spectrum of (a). In the first column the filter had been applied once, in the second column three times, or as in (d) the high frequency cut off was at a lower frequency. (b) shows the effect of the correlation algorithm we propose in this article, (c) a normalized correlation algorithm, (d) a low pass filter implemented via the Fourier power spectrum, (e) a low pass filter using a cyclic algorithm, (f) a high pass filter, and (g) an adaptive background filter. The window size of the adaptive background filter is increased to 3 Da respecting its working principle to filter peaks from a relatively constant background. The maximal improvement in signal to noise is achieved by the correlation filter after three cycles. When comparing the spectra after one and three cycles of filtering the peak enhancement effect of the correlation algorithm is obvious. However, when considering the entire mass range, the intensity values span several orders of magnitude. Systematic peak annotations will be done better by a computer program than manually.

difference from the currently calculated  $m/z$  value  $x$ . The correlation is performed in a sliding window of width  $w$ . This window size corresponds to the average peak width. We chose a value of 0.3 Th. The intensity at every mass point  $x$  in the enhanced spectrum is calcu-

lated by summing up all the intensities within a window  $w$  and multiplying it with the current intensity at the mass point  $x$ . The intensities are divided by the square root of the sum of all the intensities in the window  $w$  to reduce the overall rate at which the

intensities increase. An undesired side-effect of the correlation algorithm is that the range of the intensity values increases by several orders of magnitude. Large peaks are more strongly amplified than small peaks. A normalized correlation algorithm did not generate comparable results in improving the signal to noise ratio (see Figure 8). This effect limits the number of enhancement cycles we can use to about five because the range of the intensity values expands beyond the maximal range of our calculation package and of our spectrum display program. For display purposes, we generate an additional spectrum whose range is adjusted by a parameter:

$$I_{\text{red}}(x) = (I(x) + 2)^{F_{\text{red}}}, \text{ with } F_{\text{red}} = \frac{\log(R)}{\log(I_{\text{max}}/I_{\text{min}})} \quad (2)$$

where  $I_{\text{red}}(x)$  is the mass spectrum intensity reduced in its range,  $I(x)$  the original, enhanced mass spectrum intensity,  $F_{\text{red}}$  the range reduction factor,  $R$  the chosen range,  $I_{\text{max}}$  the maximum intensity of the spectrum  $I(x)$ , and  $I_{\text{min}}$  the smallest intensity of the spectrum larger than 0. For the spectrum displayed in Figure 2 we chose a range  $R$  of 100,000 for the enhanced spectra shown in the overview. The partial spectra in the "blow ups" were taken from the unreduced enhanced spectra to demonstrate the true increase in signal to noise achievable with this method. We used the method to demonstrate its potential on a mixture of five peptides, diluted to concentrations between 20 fmol/ $\mu\text{l}$  and 200 fmol/ $\mu\text{l}$ . Three of the peptides were identifiable in the original spectrum (see Figure 9) but only with a very low signal to noise ratio. After filtering and enhancement all five peptides could be easily identified as multiply charged ions by their isotopic spacing, some of them in several charge states and as a sodium adduct. In addition, a reasonable number of singly charged molecules became visible, as judged by the presence of a second isotope. No obvious artifact generated by the spectrum processing was visible. When acquiring a precursor ion scan of the same sample on our triple quadrupole mass spectrometer only the two more abundant peptides could be detected.

Figure 2 to Figure 9 demonstrate the efficacy of the software-based methods to increase the signal to noise level of peaks in electrospray spectra on an example where the ions can still be identified in the original spectrum, though with some difficulty. Figure 10 demonstrates that ion peaks can be made visible when they are not identifiable in the original spectrum. The doubly phosphorylated peptide IGE<sup>T</sup>pYpGVVYK was diluted to a concentration of 10 fmol/ $\mu\text{l}$ . At this concentration the signal of the doubly-charged ion at 673.2 Th was completely covered by chemical noise (see Figure 10a). After Fourier transform-based filtering of the periodic noise, the ion starts to be visible (Figure 10b) and it becomes clearly discernible after five cycles of

enhancement (Figure 10c). Its fragment spectrum has the typical appearance of a tandem MS spectrum generated from a precursor with an intensity on or below the chemical noise level. The ions above the  $m/z$  value of the precursor are specific for the fragmented peptide whereas the region below it is dominated by ions generated from the chemical background.

Figure 11 shows another application of the filtering techniques. One method to sequence peptides de novo is based on 50%  $^{18}\text{O}$ -labeling at their C-termini. C-terminal fragment ions can be identified in the fragment spectrum by this label. The differential scanning technique can be used to improve their identification [17]. Two fragment spectra of the same peptide are acquired, one selecting the entire  $^{16}\text{O}/^{18}\text{O}$  isotopic cluster and a second selecting only the  $^{18}\text{O}$ -labeled isotopes. The identification of C-terminal fragment ions is now based on two criteria, their  $^{16}\text{O}/^{18}\text{O}$  isotopic cluster in the first fragment spectrum and the missing  $^{16}\text{O}$  isotopes in the second. The differential scanning method is one of the most specific techniques to identify C-terminal fragment ions in a tandem mass spectrum spectrum of a peptide. To use it, it is necessary to know precisely the location of the first isotope of the precursor ions so that the selection window for the second acquisition can be set accordingly. Figure 11 demonstrates that for ions at about the level of the chemical noise this can be done very successfully with our noise filtering techniques.

## Conclusions

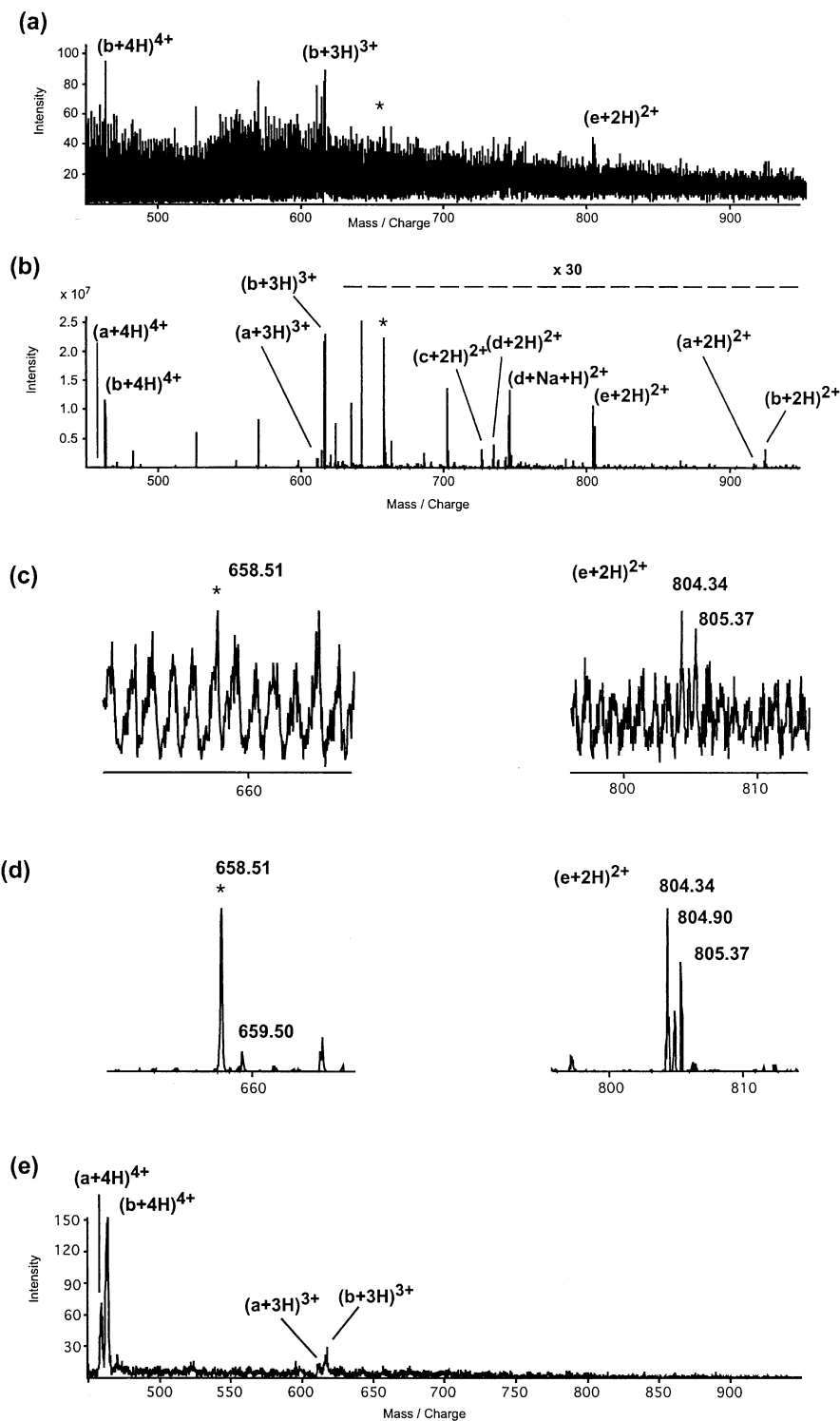
On the quadrupole time of flight mass spectrometer the software-based spectrum filtering techniques have the same functionality for low level protein sequencing as the precursor ion scan for the immonium ion of isoleucine and leucine on the triple quadrupole mass spectrometer. They allow the rapid identification of peptide ions of low level samples on or slightly below the chemical noise level. In contrast to the precursor ion scan, the filtering is not specific for peptides. All ions that are not part of the regular background noise have improved signal to noise ratios. However, the missing specificity is compensated by the isotopic resolution of the scans, which allows the direct determination of the charge states. Most often, multiply-charged ions are peptides from the digested proteins. Using these techniques, we reach the highest sensitivity for sequencing individual gel separated proteins on the quadrupole time of flight mass spectrometer by the nanoelectrospray approach.

## Acknowledgments

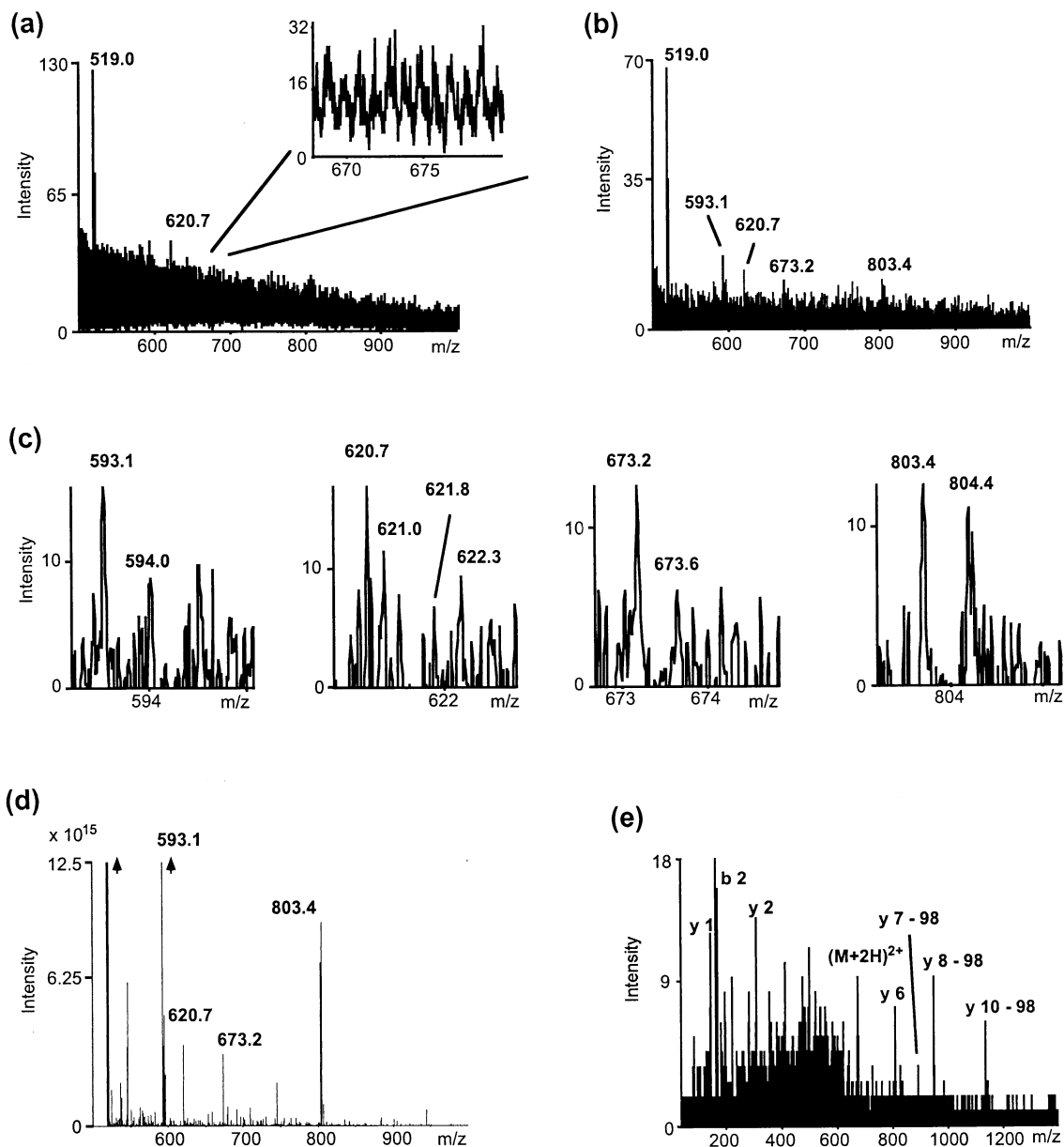
This research was supported by a grant of the Ministerium für Bildung und Forschung (BMBF) of the German Government, BioFuture 0311862.

The authors express many thanks to David Thomas for the critical reading of the manuscript.

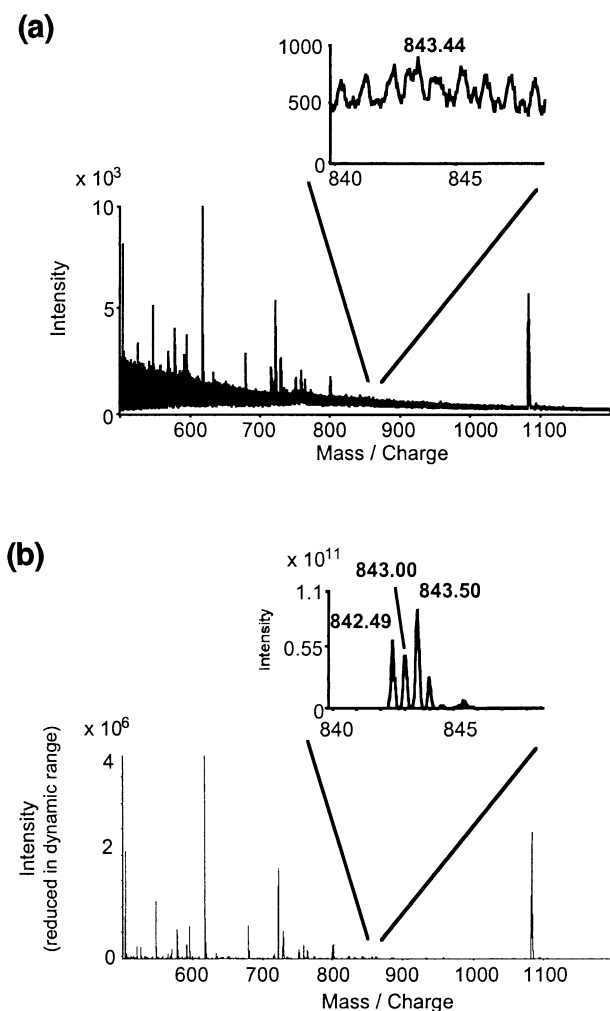




**Figure 9.** The effect of the noise filtering techniques applied to a mixture of peptides. Five peptides were diluted to concentrations between 20 fmol/ $\mu$ l and 200 fmol/ $\mu$ l [(a) YGGSKAEAAARALAKHG-MGG, (b) YGGSKAEAAARALAKHGFGG, (c) GMSAAELAAAQAAAGY, (d) GFSAAELAAAQAAAGY, (e) ADAENIKAAAQAGASGY]. (a) shows the original spectrum acquired on the Q-TOF1 mass spectrometer, (b) the same spectrum after Fourier transform based noise filtering and three cycles of peak enhancement. All five peptides become clearly visible and can be identified as multiply charged ions based on their isotopic spacing. All unlabeled peaks in the filtered spectrum appear to be singly charged ions judged from the presence of a second isotope. No abundant artifact originating from the extensive spectrum processing could be found. (c) shows the precursor ion spectrum for 86, the immonium ion of isoleucine/leucine, from the same sample acquired with our API III triple quadrupole instrument. Only the two more abundant peptides a and b are visible.



**Figure 10.** The analysis of the doubly phosphorylated peptide IGEGTpYpGVVYK at a concentration of 10 fmol/ $\mu$ l. (a) shows the original spectrum, (b) the noise-filtered spectrum, (c) individual ions from the Fourier transform based filtered spectrum, (d) the spectrum after peak enhancement, and (e) the fragment spectrum of the doubly charged phosphopeptide ( $m/z = 673.2$  Th). The doubly-charged ion at 673.2 Th can not be identified in the original spectrum. The ion is discernible after Fourier transform-based noise filtering (b), (c), but it would be very time-consuming to find the peak between all of the remaining noise signals. The ion becomes clearly visible after five cycles of peak enhancement (d). After enhancement its signal to noise ratio is 2000. It is flanked by singly-charged contaminating ions. These low-level singly-charged ions in this mass range are quite common. They are visible to the same degree in a precursor ion scan spectrum for the immonium ion of isoleucine/leucine on triple quadrupole mass spectrometers. A doubly charged ion is identified by the isotopic spacing in the Fourier transform based filtered spectrum. This allows the elimination of singly charged ions and occasional artifacts of the peak amplifying method from further consideration. After the doubly-charged precursor ion is localized, a fragment spectrum can be acquired (e). The fragment spectrum has the typical appearance of a tandem MS spectrum from a precursor with an intensity on or below the chemical noise level. The region below the  $m/z$  value of the precursor is dominated by ions generated from the chemical background, but the upper half of the spectrum contains fragment ions specific for the peptide. These ions can lead to the identification of the peptide sequence in a database search when the sequence tag algorithm is used.



**Figure 11.** Noise removal techniques applied to the electrospray spectrum of a peptide mixture. The C-termini of the peptides are labeled to 50% with  $^{18}\text{O}$ . (a) shows the original spectrum, (b) the noise filtered and enhanced spectrum. The isotopic envelopes of low-level peaks are clearly visible after Fourier transform-based filtering and three cycles of peak enhancement. The first isotope of the clusters can be assigned unambiguously. The identification of the peaks, the assignment of the charge state, and the localization of the first isotope can not be done in the original spectrum.

## References

1. Aebersold, R.; Watts, J. D. The Need for National Centers for Proteomics. *Nat. Biotechnol.* **2002**, *20*, 651.
2. Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.; Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edlmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. Functional Organization of the Yeast Proteome by Systematic Analysis of Protein Complexes. *Nature* **2002**, *415*, 141–147.
3. Ho, Y.; Gruhler, A.; Heilbut, A.; Bader, G. D.; Moore, L.; Adams, S. L.; Millar, A.; Taylor, P.; Bennett, K.; Boutilier, K.; Yang, L.; Wolting, C.; Donaldson, I.; Schandorff, S.; Shewnarane, J.; Vo, M.; Taggart, J.; Goudreault, M.; Muskat, B.; Alfarano, C.; Dewar, D.; Lin, Z.; Michalickova, K.; Willems, A. R.; Sassi, H.; Nielsen, P. A.; Rasmussen, K. J.; Andersen, J. R.; Johansen, L. E.; Hansen, L. H.; Jespersen, H.; Podtelejnikov, A.; Nielsen, E.; Crawford, J.; Poulsen, V.; Sorensen, B. D.; Matthiesen, J.; Hendrickson, R. C.; Gleeson, F.; Pawson, T.; Moran, M. F.; Durocher, D.; Mann, M.; Hogue, C. W.; Figeys, D.; Tyers, M. Systematic Identification of Protein Complexes in *Saccharomyces cerevisiae* by Mass Spectrometry. *Nature* **2002**, *415*, 180–183.
4. Yates, J. R., III; Link, A. J.; Schieltz, D. Direct Analysis of Proteins in Mixtures. Application to Protein Complexes. *Methods Mol. Biol.* **2000**, *146*, 17–26.
5. McCormack, A. L.; Schieltz, D. M.; Goode, B.; Yang, S.; Barnes, G.; Drubin, D.; Yates, J. R. Direct Analysis and Identification of Proteins in Mixtures by LC/MS/MS and Database Searching at the Low-Femtomole Level. *Anal. Chem.* **1997**, *69*, 767–776.
6. Tolley, L.; Jorgenson, J. W.; Moseley, M. A. Very High Pressure Gradient LC/MS/MS. *Anal. Chem.* **2001**, *73*, 2985–2991.
7. Vissers, J. P.; Blackburn, R. K.; Moseley, M. A. A Novel Interface for Variable Flow Nanoscale LC/MS/MS for Improved Proteome Coverage. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 760–771.
8. Choudhary, J. S.; Blackstock, W. P.; Creasy, D. M.; Cottrell, J. S. Interrogating the Human Genome Using Uninterpreted Mass Spectrometry Data. *Proteomics* **2001**, *1*, 651–667.
9. Wilm, M.; Mann, M. Analytical Properties of the Nano-electrospray Ion Source. *Anal. Chem.* **1996**, *68*, 1–8.
10. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal. Chem.* **1996**, *68*, 850–858.
11. Mann, M.; Wilm, M. Error-Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Anal. Chem.* **1994**, *66*, 4390–4399.
12. Wilm, M.; Neubauer, G.; Mann, M. Parent Ion Scans of Unseparated Peptide Mixtures. *Anal. Chem.* **1996**, *68*, 527–533.
13. Gonzalez, R. C.; Wintz, P., *Digital Image Processing*, 2nd ed.; Addison-Wesley: Reading, MA, 1987.
14. Grosshans, P. B.; Shields, J. P.; Marshall, A. G. Comprehensive Theory of the Fourier Transform Ion Cyclotron Resonance Signal for All Ion Trap Geometries. *J. Chem. Phys.* **1991**, *94*, 5341–5352.
15. Senko, M. W.; Beu, S. C.; McLafferty, F. W. Automated Assignment of Charge States from Resolved Isotopic Peaks for Multiply Charged Ions. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 52–56.
16. Carroll, J. A.; Beavis, R. C. Using Matrix Convolution Filters to Extract Information from Time-of-Flight Mass Spectra. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1683–1687.
17. Uttenweiler-Joseph, S.; Neubauer, G.; Christoforidis, S.; Zerial, M.; Wilm, M. Automated de Novo Sequencing of Proteins Using the Differential Scanning Technique. *Proteomics* **2001**, *1*, 668–682.