Published in final edited form as:

Anal Chem. 2009 May 15; 81(10): 3965–3972. doi:10.1021/ac900294r.

Direct Quantitation of Peptide Mixtures without Standards using Clusters Formed by Electrospray Ionization Mass Spectrometry

Ryan D. Leib, Tawnya G. Flick, and Evan R. Williams*

Department of Chemistry, University of California, Berkeley, CA 94720-1460

Abstract

In electrospray ionization mass spectrometry, ion abundances depend on a number of different factors, including analyte surface activity, competition between analytes for charge, analyte concentration, as well as instrumental factors, including mass-dependent ion transmission and detection. Here, a novel method for obtaining quantitative information about solution-phase concentrations of peptide mixtures is described and demonstrated for five different peptide mixtures with relative concentrations ranging from 0.05% to 50%. In this method, the abundances of large clusters containing anywhere from zero to thirteen impurity molecules are measured and directly related to the relative solution-phase concentration of the peptides. For clusters containing ~15 or more peptides, the composition of the clusters approaches the statistical value indicating that these clusters are formed nonspecifically and that any differences in ion detection or ionization efficiency are negligible at these large cluster sizes. This method is accurate to within ~20% or better, even when the relative ion intensities of the protonated monomers can differ by over an order of magnitude compared to their solution-phase concentrations. Although less accurate than other quantitation methods that employ internal standards, this method does have the key advantages of speed, simplicity, and the ability to quantitate components in solution even when the identity of the components are unknown.

Introduction

The ability to accurately and rapidly measure concentrations of individual components in complex mixtures is a key challenge in chemistry. Preparation of small molecules of pharmaceutical interest can result in low molar fraction impurities that stem from unreacted starting materials or intermediates, products from competing side reactions and degraded or modified catalysts.^{1–3} Because these diverse minor components may have unintended metabolic effects, impurities must be quantified and if present in sufficient molar excess, removed or structurally identified and shown to be biologically inert.^{1–3} In proteomics, protein modifications can be indicative of biological activity as well as provide biomarkers for disease.^{4–8} Quantitation of these discrete modifications can be used to determine relative expression levels in living systems. Mass spectrometry (MS) is commonly employed to identify the components of pharmacological or biological mixtures,^{9–14} and is used to identify thousands of peptides originating from protein digests of whole cells, biological fluids, tissues and organisms.^{15–29}

*Address correspondence to: williams@cchem.berkeley.edu.

Quantitation of nonvolatile or thermally labile molecules using MS can be complicated by differences in ionization efficiency and matrix effects, in which analyte molecules compete for charge with other components in the mixture,^{30–34} as well as effects of instrumental parameters, including m/z dependent transmission and detection efficiency.^{35–38} In matrix assisted laser desorption/ionization (MALDI), matrix effects can occur as a result of cocrystallization of analytes or in the desorption/ionization process.^{31,39} In electrospray ionization (ESI), ion abundances and charge-state distributions can depend on many factors, including the relative surface activities, basicities, surface tension, concentrations and conformations of the analytes and types of solvent.^{30,40–52}

Relative ion abundances and charge states in ESI mass spectra have been related to the relative surface activities or hydrophobicities of the molecules or ions.^{30,40–43} The abundances of anions formed from a solution containing equimolar concentrations of anionic salts trend with the Hofmeister series, with greater dehydration free energies and lower surface affinity correlating to lower relative ion intensity in the gas phase.⁴⁰ The ESI intensity of tetraalkyl ammonium halides increases by over an order of magnitude when the hydrophobic chain length increases from methyl to butyl, an effect attributed to the relative solvophobicity of these ions.³⁰ Differences in charge-state distributions observed for similar size peptides has also been attributed to differences in analyte solvophobicity.³⁰ Enke and coworkers investigated the effect of peptide surface activity on ion formation by ESI and found that ion intensities of protonated tripeptides decreased with increasing sidechain polarity, indicating that differential partitioning between ions at surfaces vs. in the bulk can affect the observed ion intensity.⁴¹ Addition of basic solvents to peptide and protein containing solutions can result in lower charge state distributions and reduced signal for the peptide or protein ions.⁴⁴ The charge state distribution of protein ions formed from denaturing solutions occur at lower m/z ,^{48–50} which can result in ion abundances that depend on solution composition or pH due to m/z dependent detection biases. For low pH solutions in which proteins are denatured, Pan and McLuckey⁵³ showed that signal responses of different proteins were better correlated to their solution-phase concentration when normalized for the average charge state of the protein.

Because ion intensities depend on both ionization and instrumental effects, internal standards are typically used to obtain quantitative information from ESI-MS.^{54–64} Accurate quantitative information can be obtained by using isotopically labeled analyte molecules as an internal standard. Various methods of introducing isotopic labels have been successfully demonstrated, including chemical modification methods,^{10,54–58,65} metabolic incorporation,^{59–61} and “spiking” the sample with a synthetically labeled analyte.^{63,64} With isotope coded affinity tags (ICAT),^{10,54,55} chemically modified peptides can be selectively retained during an affinity-based chromatographic separation. The abundances of ICAT peptides can be then compared with peptides from other cells that have been modified using a linker with a different isotopic composition which makes possible accurate quantitation of protein expression between the different cell lines. Another common approach to quantitative measurement of relative cell line proteomic expression is stable isotope labeling with amino acids in cell culture (SILAC).^{60,61} In SILAC, a cell line grown in media containing an isotopically enriched amino acid is compared against a line grown in

unenriched media making possible direct comparison of expressed proteins. These approaches can yield quantitation accurate to on average better than ~5–10%.^{61,66} Mass defect tags, in which chemical modifications incorporating elements, such as Br, can be used to shift the mass of tagged peptides slightly off the exact masses of similar size peptides. This makes possible accurate quantitation even in complex mixtures.^{57,65} More generally, preparation of isotopically labeled species for each analyte can be quite time consuming, particularly when an unknown analyte must first be identified. Analog molecules that have similar ionization efficiencies and mass to the analyte can also be used to obtain quantitative information. Lebrilla and coworkers demonstrated that by using a deuterated heptameric maltosaccharide as an internal standard, the concentrations of heptameric fructosaccharides could be obtained to within 1% and other fructo-oligomers in the mixture within 3% using MALDI-MS.⁶⁷ Methods to quantitate enantiomeric mixtures using tandem MS or ion mobility MS have also been developed.^{68–75}

Here, a novel method to obtain quantitative information about relative peptide concentrations in solutions without using standards is demonstrated. In this method, the abundances of *nonspecific*, heterogeneous clusters of molecules are used to obtain accurate solution concentrations. This method is demonstrated on five different peptide mixtures with relative molar fractions ranging from 0.05% to 50%. This method does not require standards, is very fast, reasonably accurate, and can be used to obtain quantitative information about components whose identity is not known.

Experimental

Sample Preparation

Leucine enkephalin, methionine enkephalin, bradykinin, bradykinin fragment 1-8, and bradykinin fragment 2-9 were purchased from Sigma Aldrich Co. (St. Louis, MO); [ala]₃-leucine enkephalin (AAAYGGFL) and [lys]₆-leucine enkephalin (YGGFLK) were purchased from VWR (West Chester, PA). Standard 5.00 mM aqueous stock solutions were prepared by weighing the peptides on a microbalance (± 0.01 mg) and then correcting these masses for peptide content using information from the manufacturers' certificate of analysis. All other solutions (48% water, 48% methanol, 4% acetic acid) were prepared using these stock solutions.

Mass Spectrometry

Experiments were performed on a 9.4 T Fourier transform ion cyclotron resonance (FT/ICR) mass spectrometer with an external ESI source that is described elsewhere.^{76,77} Ions are generated by nanoelectrospray (nanoESI) using borosilicate capillaries pulled so that the tips have a ~2 μ m inner diameter (model P-87 capillary puller, Sutter Instruments, Novato, CA). A small volume of analyte solution, typically ~2–10 μ L, is loaded into the capillary, and a platinum wire is inserted into the solution and grounded. The borosilicate capillary is positioned ~2 mm from the source inlet to which a potential of –800 to –1200 V is applied. Ions are accumulated in a hexapole trap for 1.1 s and are subsequently introduced into the ion cell. A pulse of nitrogen gas is introduced through a piezoelectric leak valve to a pressure of $\sim 1 \times 10^{-6}$ Torr to assist trapping and thermalization of the ions. Ions from two

hexapole injections are accumulated in the cell prior to exciting and detecting the ions. The signal intensity of the peptide clusters is optimized by selectively tuning the DC potential for the external hexapole in a range from 2.8 to 4.0 V.⁷⁸ Three mass spectra composed of 50 coadded scans (Predator datastation, NHMFL, Tallahassee, Florida) were acquired at three different DC offset values for each analyte to minimize any effects of transmission efficiency. For experiments where the solution concentration was varied, each mass spectrum was acquired using a new borosilicate capillary to assess effects of tip-to-tip variations and to avoid any sample contamination.

Results and Discussion

ESI Mass Spectra

Solutions containing leucine enkephalin (LE) and [ala]₃-leucine enkephalin (A₃LE) were prepared to 2.4 mM total peptide concentration with the fraction of A₃LE ranging from 0.05% to 50%. A nanoESI mass spectrum of one of these solutions in which the molar fraction of A₃LE is 1.0%, obtained under typical experimental conditions, is shown in Figure 1a. The abundance of protonated A₃LE is 3.8% that of LE in this spectrum, a value much higher than the molar fractions of these peptides in solution. The normalized abundances of the protonated molecular ions as a function of molar fraction of A₃LE in these 2.4 mM solutions are shown in Figure 2a. At molar fractions below 1%, the normalized abundance of (A₃LE + H)⁺ relative to that of (LE + H)⁺ is about 4 times higher than the relative concentration of these peptides in solution. The deviation between the relative ion abundance and relative solute concentration increases with increasing fractional A₃LE concentration. At molar fractions of 20% and 50%, the [A₃LE + H]⁺/[LE + H]⁺ ratio is 1.7 and 7.3, respectively, corresponding to an approximately eight fold higher relative ion abundance of (A₃LE + H)⁺ than its relative concentration in solution.

There are many factors that contribute to ion abundances in ESI mass spectra, including instrument parameters, detection, transmission and ionization efficiencies and ion suppression in mixtures.^{30–38,40–50} The greater propensity to form protonated A₃LE compared to LE in these mixtures may be due to the greater surface activity or hydrophobicity of the former. Alanine is hydrophobic and addition of A₃ to LE should increase the surface activity of A₃LE relative to that of LE. Preferential formation of ions that have high surface activities have been reported previously.^{30,40,41} At high solute concentration (few hundred μ M), ion formation is limited by the excess charge on the ESI droplet surface,⁵¹ which appears to enhance this effect presumably because less charge is available for the less surface-active species.

In addition to the protonated molecular ions, abundant ions corresponding to the singly protonated dimer and trimer of LE as well as multiply charged higher order protonated clusters are clearly observed (Figure 1a). Heterogeneous clusters, such as the dimer (A₃LE + LE + H)⁺, are also formed. Cluster formation is common in ESI of concentrated solutions, e.g., 100 μ M.^{78–80} The abundances of the clusters in the ESI spectra can be enhanced by changing several different instrument parameters, including the DC offset potential of the hexapole. This effect is illustrated in Figure 1a-c in which the DC offset potential was 2.8, 3.4 and 4.0 V, respectively, with all other conditions the same. Higher order clusters, up to

(LE₂₉ + 5H)⁵⁺ and its heterogeneous analogs, are observed. The abundances of lower *m/z* ions are significantly attenuated at the higher hexapole DC offset potentials. In addition, the relative ion abundance of protonated A₃LE to LE is dramatically decreased. This ratio is 3.8%, 2.7% and <0.3% at these respective potentials. Other instrumental parameters can also influence relative ion abundances. These results illustrate that the relative abundance of these protonated molecular ions are an unreliable indicator of their solution molar fractions.

Effects of Solute Concentration

To investigate the effects of solute concentration on the relative abundances of protonated A₃LE to LE, ESI mass spectra were acquired for equimolar mixtures of these two peptides from 10 nM to 2.4 mM total peptide concentration. Different nanoESI tips were used for each solution to avoid sample contamination. Although mass spectra acquired using the same tip are typically reproducible to better than 1%, the use of different tips results in more significant variations in relative ion intensities.

To obtain an estimate of the tip-to-tip variability, three replicate measurements were made using three different tips at each solute concentration. A plot of the normalized ion abundance of protonated A₃LE, $[A_3LE + H]^+ / ([A_3LE + H]^+ + [LE + H]^+)$, as a function of solute concentration for this 50/50 mixture, is shown in Figure 2b. Error bars correspond to the standard deviation of three replicate measurements. At low concentrations, the uncertainty is higher due to the lower S/N. At higher concentrations, this uncertainty due to tip-to-tip variability is typically a few percent. The high uncertainty at 50 μM is due to one of the three mass spectra being significantly different than the other two, which shows that slight variations in seemingly identical nanospray capillaries can occasionally result in large variations in ion signals. At concentrations above 100 μM, (A₃LE + H)⁺ is approximately 8 times more abundant than (LE + H)⁺ whereas for solutions 10⁻⁶ M, these abundances are nearly equal, consistent with their equimolar solution-phase concentrations. Thus, the abundances of the protonated molecular ions are more representative of their solution-phase molar fractions when the overall solute concentration is at or below 1 μM. However, this is not always the case for other peptide mixtures.

Mass spectra of four other equimolar peptide mixtures were measured at three different solute concentrations: methionine enkephalin (ME) or [lys₆]-leucine enkephalin (KLE) with LE, and bradykinin fragment 1-8 (BK1-8) or bradykinin fragment 2-9 (BK2-9) with bradykinin (BK). The molar fractions derived from the relative protonated molecular ion intensities measured in these experiments are given in Table 1. For the ME/LE mixture, molecular ion abundances reflect the molar fraction in solution at all three concentrations. Replacement of leucine by methionine should have little effect on the basicity or surface activity of the peptide, consistent with the similar ionization efficiencies observed. In contrast, the abundances of protonated A₃LE and KLE do not reflect well their solution-phase concentrations. Molar fractions derived from these ion abundances are too high at 2.4 mM (80% and 71%, respectively) whereas KLE values are too low (38%) at 1.2 μM. Addition of a lysine residue to LE should significantly affect the physical properties of the peptide resulting in different ionization efficiencies for these two peptides.

For equimolar BK1-8/BK and BK2-9/BK solutions, the relative molecular ion intensities also depend on solute concentration. For BK2-9/BK, the ionic molar fraction is significantly higher than that in solution at 1.2 μM . These data illustrate that using the relative abundance of protonated molecular ions to determine the molar fraction of components in solution can result in significant deviation from the solution-phase concentration. This deviation can be reduced for solute concentrations below 1 μM in some cases. However, the ability to detect low molar fraction components in such dilute solutions can be limited by S/N. In these experiments, the S/N of molecular ions from 50 nM solutions is ~ 100 for $(\text{LE} + \text{H})^+$, which limits detection of minor components to $\sim 2\%$ without modifying the experiment to improve overall S/N.

Solute Concentration from Cluster Ion Intensities

A wide range of both homogeneous and heterogeneous cluster ions can typically be formed at high peptide concentrations (Figure 1). If the individual constituent molecules are incorporated into these clusters statistically, then the molar fraction of the minor component, $F_m\%$, can be expressed as a binominal expansion for any two clusters of the same number of subunits, n , (equation 1);

$$F_m\% = \sqrt[n]{\frac{P(X_h, Y_k)}{p^k \bullet \left(\frac{n!}{h!(n-h)!}\right)}} \quad (1)$$

where $P(X_h, Y_k)$ is the probability of forming a cluster corresponding to h units of component X , and k units of component Y such that $h + k = n$, and p is the observed fraction corresponding to the homogeneous cluster.

To illustrate this process, an expansion of the mass spectrum from Figure 1a showing clusters for $n = 19$ obtained from the 2.4 mM 1% A_3LE with LE solution is shown in Figure 3. The normalized abundances of $(\text{LE}_{19} + 5\text{H})^{5+}$ and $(\text{LE}_{18} + \text{A}_3\text{LE} + 5\text{H})^{5+}$ are 100.0 and 10.2, respectively. No incorporation of additional A_3LE molecules, i.e., $(\text{LE}_{17} + \text{A}_3\text{LE}_2 + 5\text{H})^{5+}$, is observed for this cluster size at this molar fraction. To obtain the fractional incorporation, $P(\text{A}_3\text{LE}_1, \text{LE}_{18})$, the intensity of the cluster that has incorporated a single A_3LE is divided by the sum of that cluster and the homogeneous cluster, i.e., $[\text{LE}_{18} + \text{A}_3\text{LE} + 5\text{H}]^{5+} / ([\text{LE}_{18} + \text{A}_3\text{LE} + 5\text{H}]^{5+} + [\text{LE}_{19} + 5\text{H}]^{5+}) = 10.2 / (100.0 + 10.2) = 0.093$. From this value and the cluster size, an average value for the molar fraction of $0.87\% \pm 0.32\%$ is obtained from the binomial expansion (equation 1) for all three DC offset values. This average molar fraction corresponds to the experimentally determined molar fraction derived from the $n = 19$ clusters. Similarly, molar fraction values are obtained for each cluster size and these values are shown as a function of cluster size in Figure 4.

The molar fractions obtained from the cluster data vary significantly for clusters with $n < 10$. For example, molar fractions of $\sim 3.7\%$ and 0.2% are obtained from the $n = 5$ and 6 clusters, respectively. However, molar fractions determined from the $n = 17 - 29$ clusters are similar. The average of the values for the $n = 17 - 29$ clusters is $1.0 \pm 0.3\%$, consistent with the

solution molar fraction (1.0%). This suggests that these large clusters are nonspecific and that differences in composition have only a minor effect on relative ionization efficiencies.

Results for Different Molar Fractions

Molar fractions determined as a function of cluster size were obtained for each solution with A₃LE ranging from 0.05% to 10% of the total peptide concentration with LE as described above. For these concentrations, there are typically two or three different clusters for a given size n , with distributions weighted towards clusters with higher extent of incorporation of the more abundant component. In these cases, the effect of noise on the observed ion intensities is small, and the binomial distribution correlates well with the solution molar fraction. This is illustrated in Figure 5, which shows representative cluster intensities as a function of cluster composition and solution molar fraction. For low molar fractions, the abundance of the homogeneous cluster is high (Figure 5a and b). However, for solutions containing 20% or more of the A₃LE with LE, higher order clusters that contain multiple A₃LE molecules are abundant in the mass spectra and these heterogeneous clusters increase with increasing molar fraction and cluster size. To obtain a discrete value using the binomial expansion, the homogeneous cluster must be observed. However, the homogeneous clusters at the larger cluster sizes are not typically observed for the 20% and 50% molar fraction solutions due to low S/N (Figure 5c). For these clusters, a weighted average can be used (equation 2);

$$F_m \% = \frac{\sum_h I_h \times \frac{h}{n}}{\sum_h I_h} \quad (2)$$

where I is the intensity of each observed cluster consisting of n total peptide molecules with h molecules of the minor fraction component. This weighted average method works well for symmetrical distributions, such as those observed at high molar fractions or large cluster sizes, even at relatively low S/N because the center of the distribution can be accurately measured (Figure 5c). The asymmetrical distributions observed for lower molar fractions and small clusters are more difficult to accurately measure when the signal-to-noise is low (Figure 5a,b). In these cases, the weighted average of these clusters is systematically low. Thus, both methods are used: a binomial expansion when both the homogeneous peak and a heterogeneous cluster are observed and a weighted average when no homogeneous cluster is detected. The binomial expansion method should be less sensitive to effects of low S/N because only the high intensity homogeneous and single impurity clusters are used (Figure 5a and b).

The A₃LE % molar fraction calculated using the cluster intensities for $n = 17$ as a function of % molar fraction in solution is shown in Figure 6. These data can be fit to a straight line ($r^2 = 0.97$) with a slope of 0.92. This value is less than 1.0, which indicates that the concentration of A₃LE determined by this method is systematically low. There is also some uncertainty in the solute concentration due to the potential for water absorption by the peptides during weighing, although this uncertainty is likely to be small. These data suggest

that there may be a slight preference for incorporation of LE into these clusters, although this effect appears to be minor.

In contrast to the results obtained from the cluster data, a linear fit of the molar fractions determined from the protonated molecular ion abundances vs. concentration from these same mass spectra has a slope of 14.2 ($r^2 = 0.96$). Quantitation using this cluster method therefore provides a significantly more accurate measure of the solution-phase concentration than the relative abundances of the protonated molecular ions over a wide range of solution compositions. Because the composition of the large clusters is similar, effects of preferential ionization should be minimal.

Effects of Analyte Identity

To investigate the generality of this method, analogous experiments were performed on four additional peptide mixtures. Solutions containing various molar fractions of either ME or KLE with LE, and either BK1-8 or BK2-9 with BK were prepared and three replicate mass spectra at the three hexapole DC offset values were obtained. These peptides were chosen to investigate effects of surface activity and basicity. As was observed for A₃LE in LE, the molar fraction obtained from the cluster ion intensities varies with cluster size for smaller clusters. In each case, the protonated molecular ions and clusters with $n \geq 10$ result in measured molar fractions that are significantly different from those expected based on solute concentration. Data for 1% molar fraction BK1-8 in BK are shown in Figure 7. The molar fraction value obtained from $n = 4$ and 8 is 5.2% and 4.3%, respectively, and are significantly higher than the solution value of 1%. Similar results were obtained for the smaller clusters formed from solutions containing the other peptides, suggesting that preferential incorporation of some peptides into the smaller clusters occurs.

The molar fraction calculated for the larger cluster ions approaches solution values for all the mixtures, indicating that effects of cluster specificity and ionization effects are minor for these larger clusters. Clusters formed from the 1% solution of BK1-8 in BK with $n \geq 15$ are relatively constant with an average value of 1.4 ± 0.3 . Similar results are obtained for the other solutions, but the cluster size where the value becomes constant depends on the analyte identities. Clusters sizes of $n \geq 12$, 20 and 15 were used to obtain the molar fractions for solutions containing ME with LE, KLE with LE and BK2-9 with BK, respectively. The lower cluster size for ME with LE is likely because substitution of a methionine for leucine should have a relatively small effect on cluster specificity or ionization efficiency. The larger cluster size for KLE in LE is likely due to a more significant change to the physical properties of these peptides. The % molar fractions determined from the cluster data, as well as the values determined from the protonated molecular ion intensities, are plotted against the solution % molar fractions in Figure 8a–d for these solutions.

For all four of these mixtures, the data for the protonated molecular ions are not well fit by a line. The slope of these data for the protonated molecular ions from solutions containing BK1-8/BK and BK2-9/BK are 13.7 ($r^2 = 0.92$) and 0.2 ($r^2 = 0.64$), respectively, and for solutions containing ME with LE and KLE with LE the slopes are 1.11 ($r^2 = 0.88$) and 1.18 ($r^2 = 0.88$), respectively. These results show that the relative protonated molecular ion abundances can either be significantly higher or lower than their corresponding solution-

phase concentration. However, molar fractions obtained from the cluster data are linear, and the slope of best fit lines are 0.95 ($r^2 > 0.99$), 1.19 ($r^2 = 0.98$), 1.08 ($r^2 > 0.99$) and 1.10 ($r^2 > 0.99$) for solutions containing ME, BK2-9, BK1-8 and KLE, respectively. Thus, the molar fractions obtained from the cluster ion data are within ~10% of the solution-phase concentrations over a four-order of magnitude range in solution molar fraction for four of the five peptide mixtures investigated, and within 20% over a three-order of magnitude range for BK2-9 in BK. These results indicate that this cluster method should be a general and robust technique to obtain reasonably accurate quantitative information about peptide concentration in solution without the need for standards or knowledge of molecular identity.

Conclusions

Quantitation of peptides in solution using MS is complicated by differences in the physical properties of the molecules, which can result in preferential ionization of some components in a mixture as well as instrumental effects including m/z dependent ion transmission and detection efficiency. These effects can be reduced by using the abundances of large, nonspecific clusters from which accurate measures of the solution-phase concentrations can be obtained. For five different peptide mixtures, the relative solution-phase concentrations obtained from the abundances of large clusters are accurate to within ~20%. In contrast, the relative abundances of the protonated molecular ions can differ from the relative solution-phase concentrations by more than an order of magnitude.

Although the accuracy of this method is lower than that obtainable by using internal standards, this method has several important advantages: 1) it does not require either an internal nor an external standard which greatly reduces the time and effort necessary to obtain quantitative information, 2) the components do not need to be identified in order to measure their relative solution-phase concentration and 3) effects of instrument or detector mass bias are significantly reduced which is of key importance when the charge states or masses of peptides differ significantly. This method requires relatively high solute concentrations to produce large clusters, but only a few picomoles of sample are consumed using nanoESI at typical flow rates of ~5 nL/min.

The accuracy of this method should be improved by obtaining data from even larger clusters, although this also increases the instrument performance necessary to resolve the isotopic distributions for charge assignment and for separating overlapping cluster distributions for more complicated mixtures. Systematic error due to effects of the asymmetry of the cluster distributions is reduced at higher S/N. Both of these improvements would increase the resolution and dynamic range of the method and could make possible simultaneous quantitation of multiple impurities in a single mixture. This method should be equally applicable to smaller molecules, such as pharmaceuticals, for which larger clusters could be produced and analyzed to provide rapid quantitative information about possible impurities without having to use standards or even knowing the identity of the impurities. Results from the large peptide clusters also provide compelling support for the use of ion abundances in MS to obtain accurate measures of protein modifications or adducts where relatively minor chemical changes to a large protein should have only minor effects on ionization.

Acknowledgments

The authors are grateful to Dr. Gregory T. Blakney for assistance in the development of the Predator data station and to generous financial support from the National Institutes of Health (R01GM064712-08).

References

1. Görög S. *Anal Bioanal Chem.* 2003; 377:852–862. [PubMed: 12937879]
2. Görög S. *J Pharm Biomed Anal.* 2008;247–253. [PubMed: 18082992]
3. Qiu F, Norwood DL. *J Liq Chromatogr Relat Technol.* 2007; 30:877–935.
4. Freitas MA, Sklenar AR, Parthun MR. *J Cell Biochem.* 2004; 92:691–700. [PubMed: 15211567]
5. Jensen ON. *Nat Rev Mol Cell Biol.* 2006; 7:391–403. [PubMed: 16723975]
6. Hanash S. *Nature.* 2003; 422:226–232. [PubMed: 12634796]
7. Zhou H, Watts JD, Aebersold R. *Nat Biotechnol.* 2001; 19:375–378. [PubMed: 11283598]
8. Kelleher NL, Hicks LM. *Curr Opin Chem Biol.* 2005; 9:424–430. [PubMed: 16129650]
9. Aebersold R, Mann M. *Nature.* 2003; 422:198–207. [PubMed: 12634793]
10. Gygi SP, Rist B, Gerber SA, Turek F, Gelb MH, Aebersold R. *Nat Biotechnol.* 1999; 17:994–999. [PubMed: 10504701]
11. Olah TV, McLoughlin DA, Gilbert JD. *Rapid Commun Mass Spectrom.* 1997; 11:17–23. [PubMed: 9050259]
12. Wiener MC, Sachs JR, Deyanova EG, Yates NA. *Anal Chem.* 2004; 76:6085–6096. [PubMed: 15481957]
13. Liu H, Sadygov RG, Yates JR III. *Anal Chem.* 2004; 76:4193–4201. [PubMed: 15253663]
14. Moresco JJ, Dong MQ, Yates JR III. *Am J Clin Nutr.* 2008; 88:597–604. [PubMed: 18779273]
15. Petyuk VA, Qian W-J, Hinault C, Gritsenko MA, Singhal M, Monroe ME, Camp DG II, Kulkarni RN, Smith RD. *J Proteome Res.* 2008; 7:3114–3126. [PubMed: 18570455]
16. Metz TO, Jacobs JM, Gritsenko MA, Fontès G, Qian WJ, Camp DG, Poitout V, Smith RD. *J Proteome Res.* 2006; 5:3345–3354. [PubMed: 17137336]
17. Petyuk VA, Qian W-J, Chin MH, Wang H, Livesay EA, Monroe ME, Adkins JN, Jaitly N, Anderson DJ, Camp DG II, Smith DJ, Smith RD. *Genome Res.* 2007; 17:328–336. [PubMed: 17255552]
18. Strittmatter EF, Ferguson PL, Tang K, Smith RD. *J Am Soc Mass Spectrom.* 2003; 14:980–991. [PubMed: 12954166]
19. Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y. *Nat Biotechnol.* 2005; 23:617–621. [PubMed: 15834404]
20. Patrie SM, Ferguson JT, Robinson DE, Whipple D, Rother M, Metcalf WW, Kelleher NL. *Mol Cell Proteomics.* 2006; 5:14–25. [PubMed: 16236702]
21. Parks BA, Jiang L, Thomas PM, Wenger CD, Roth MJ, Boyne MT II, Burke PV, Kwast KE, Kelleher NL. *Anal Chem.* 2007; 79:7984–7991. [PubMed: 17915963]
22. Forbes AJ, Patrie SM, Taylor GK, Kim YB, Jiang L, Kelleher NL. *Proc Natl Acad Sci USA.* 2004; 101:2678–2683. [PubMed: 14976258]
23. Dong MQ, Venable JD, Au N, Xu T, Park SK, Cociorva D, Johnson JR, Dillin A, Yates JR III. *Science.* 2007; 317:660–663. [PubMed: 17673661]
24. Liao L, McClatchy DB, Park SK, Xu T, Lu B, Yates JR III. *J Proteome Res.* 2008; 7:4743–4755. [PubMed: 18823140]
25. Denny P, Hagen FK, Hardt M, Liao L, Yan W, Arellano M, Bassilian S, Bedi GS, Boontheung P, Cociorva D, Delahunty CM, Denny T, Dunsmore J, Faull KF, Gilligan J, Gonzalez-Begne M, Halgand F, Hall SC, Han X, Henson B, Hewel J, Hu S, Jeffrey S, Jiang J, Loo JA, Loo RRO, Malamud D, Melvin JE, Miroshnychenko O, Navazesh M, Niles R, Park SK, Prakobphol A, Ramachandran P, Richert M, Robinson S, Sondej M, Souda P, Sullivan MA, Takashima J, Than S, Wang J, Whitelegge JP, Witkowska HE, Wolinsky L, Xie Y, Xu T, Yu W, Ytterberg J, Wong DT, Yates JR III, Fisher SJ. *J Proteome Res.* 2008; 7:1994–2006. [PubMed: 18361515]

26. Zhai B, Villén J, Beausoleil SA, Mintseris J, Gygi SP. *J Proteome Res.* 2008; 7:1675–1682. [PubMed: 18327897]
27. Swaney DL, McAlister GC, Coon JJ. *Nat Methods.* 2008; 5:959–964. [PubMed: 18931669]
28. Savitski MM, Nielsen ML, Kjeldsen F, Zubarev RA. *J Proteome Res.* 2005; 4:2348–2354. [PubMed: 16335984]
29. Krijgsveld J, Ketting RF, Mahmoudi T, Johansen J, Artal-Sanz M, Verrijzer CP, Plasterk RHA, Heck AJR. *Nat Biotechnol.* 2003; 21:927–931. [PubMed: 12858183]
30. Fenn JB. *J Am Soc Mass Spectrom.* 1993; 4:524–535. [PubMed: 24227639]
31. Cohen SL, Chait BT. *Anal Chem.* 1996; 68:31–37. [PubMed: 8779435]
32. Kebarle P. *J Mass Spectrom.* 2000; 35:804–817. [PubMed: 10934434]
33. Smith RD, Shen Y, Tang K. *Acc Chem Res.* 2004; 37:269–278. [PubMed: 15096064]
34. King R, Bonfiglio R, Fernandez-Metzler C, Miller-Stein C, Olah T. *J Am Soc Mass Spectrom.* 2000; 11:942–950. [PubMed: 11073257]
35. Belov ME, Nikolaev EN, Harkewicz R, Masselon CD, Alving K, Smith RD. *Int J Mass Spectrom.* 2001; 208:205–225.
36. Page JS, Kelly RT, Tang K, Smith RD. *J Am Soc Mass Spectrom.* 2007; 18:1582–1590. [PubMed: 17627841]
37. Axelsson J, Scrivener E, Haddleton DM, Derrick PJ. *Macromolecules.* 1996; 29:8875–8882.
38. Marshall AG, Hendrickson CL, Jackson GS. *Mass Spectrom Rev.* 1998; 17:1–35. [PubMed: 9768511]
39. Tang K, Allman SL, Jones RB, Chen CH. *Anal Chem.* 1993; 65:2164–2166.
40. Cheng J, Vecitis CD, Hoffmann MR, Colussi AJ. *J Phys Chem B.* 2006; 110:25598–25602. [PubMed: 17181193]
41. Cech NB, Enke CG. *Anal Chem.* 2000; 72:2717–2723. [PubMed: 10905298]
42. Cech NB, Enke CG. *Mass Spectrom Rev.* 2001; 20:362–387. [PubMed: 11997944]
43. Kuprowski MC, Boys BL, Konermann L. *J Am Soc Mass Spectrom.* 2007; 18:1279–1285. [PubMed: 17500007]
44. Iavarone AT, Jurchen JC, Williams ER. *J Am Soc Mass Spectrom.* 2000; 11:976–985. [PubMed: 11073261]
45. Iavarone AT, Jurchen JC, Williams ER. *Anal Chem.* 2001; 73:1455–1460. [PubMed: 11321294]
46. Iavarone AT, Williams ER. *Int J Mass Spectrom.* 2002; 219:63–72.
47. Kebarle P, Peschke M. *Anal Chim Acta.* 2000; 406:11–35.
48. Loo JA, Loo RRO, Udseth HR, Edmonds CG, Smith RD. *Rapid Commun Mass Spectrom.* 1991; 5:101–105. [PubMed: 1666527]
49. Konermann L, Douglas DJ. *Biochemistry.* 1997; 36:12296–12302. [PubMed: 9315869]
50. Dobo A, Kaltashov IA. *Anal Chem.* 2001; 73:4763–4773. [PubMed: 11681449]
51. Tang L, Kebarle P. *Anal Chem.* 1993; 65:3654–3668.
52. Iavarone AT, Williams ER. *J Am Chem Soc.* 2003; 125:2319–2327. [PubMed: 12590562]
53. Pan P, McLuckey SA. *Anal Chem.* 2003; 75:1491–1499. [PubMed: 12659214]
54. Han DK, Eng J, Zhou H, Aebersold R. *Nat Biotechnol.* 2001; 19:946–951. [PubMed: 11581660]
55. Gygi SP, Rist B, Griffin TJ, Eng J, Aebersold R. *J Proteome Res.* 2002; 1:47–54. [PubMed: 12643526]
56. DeSouza L, Diehl G, Rodrigues MJ, Guo J, Romaschin AD, Colgan TJ, Siu KWM. *J Proteome Res.* 2005; 4:377–386. [PubMed: 15822913]
57. Hall MP, Ashrafi S, Obegi I, Petesch R, Peterson JN, Schneider LV. *J Mass Spectrom.* 2003; 38:809–816. [PubMed: 12938101]
58. Rodgers RP, Blumer EN, Hendrickson CL, Marshall AG. *J Am Soc Mass Spectrom.* 2000; 11:835–840. [PubMed: 11014445]
59. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. *Proc Natl Acad Sci USA.* 1999; 96:6591–6596. [PubMed: 10359756]

60. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. *Mol Cell Proteomics*. 2002; 1:376–386. [PubMed: 12118079]
61. Ong SE, Kratchmarova I, Mann M. *J Proteome Res*. 2003; 2:173–181. [PubMed: 12716131]
62. Ong SE, Mann M. *Nat Chem Biol*. 2005; 1:252–262. [PubMed: 16408053]
63. Ji J, Chakraborty A, Geng M, Zhang X, Amini A, Bina M, Regnier F. *J Chromatogr B*. 2000; 745:197–210.
64. Havliš J, Shevchenko A. *Anal Chem*. 2004; 76:3029–3036. [PubMed: 15167779]
65. Hernandez H, Niehauser S, Boltz SA, Gawandi V, Phillips RS, Amster IJ. *Anal Chem*. 2006; 78:3417–3423. [PubMed: 16689545]
66. Yi EC, Li XJ, Cooke K, Lee H, Raught B, Page A, Aneliunas V, Hieter P, Goodlett DR, Aebersold R. *Proteomics*. 2005; 5:380–387. [PubMed: 15648049]
67. Seipert RR, Barboza M, Ninonuevo MR, LoCascio RG, Mills DA, Freeman SL, German JB, Lebrilla CB. *Anal Chem*. 2008; 80:159–165. [PubMed: 18044961]
68. Zhang D, Tao WA, Cooks RG. *Int J Mass Spectrom*. 2001; 204:159–169.
69. Augusti DV, Carazza F, Augusti R, Tao WA, Cooks RG. *Anal Chem*. 2002; 74:3458–3462. [PubMed: 12139054]
70. Augusti DV, Augusti R, Carazza F, Cooks RG. *Chem Commun*. 2002:2242–2243.
71. Dwivedi P, Wu C, Matz LM, Clowers BH, Siems WF, Hill HH Jr. *Anal Chem*. 2006; 78:8200–8206. [PubMed: 17165808]
72. Grigorean G, Ramirez J, Ahn SH, Lebrilla CB. *Anal Chem*. 2000; 72:4275–4281. [PubMed: 11008760]
73. Ahn S, Ramirez J, Grigorean G, Lebrilla CB. *J Am Soc Mass Spectrom*. 2001; 12:278–287. [PubMed: 11281603]
74. Guo J, Wu J, Siuzdak G, Finn MG. *Angew Chem-Int Edit*. 1999; 38:1755–1758.
75. Mie A, Jörlén-Karlsson M, Axelsson BO, Ray A, Reimann CT. *Anal Chem*. 2007; 79:2850–2858. [PubMed: 17326611]
76. Jurchen JC, Williams ER. *J Am Chem Soc*. 2003; 125:2817–2826. [PubMed: 12603172]
77. Robinson EW, Williams ER. *J Am Soc Mass Spectrom*. 2005; 16:1427–1437. [PubMed: 16023362]
78. Jurchen JC, Garcia DE, Williams ER. *J Am Soc Mass Spectrom*. 2003; 14:1373–1386. [PubMed: 14652186]
79. Baker ES, Bernstein SL, Bowers MT. *J Am Soc Mass Spectrom*. 2005; 16:989–997. [PubMed: 15908229]
80. Liu B, Hvelplund P, Nielsen SB. *J Am Soc Mass Spectrom*. 2005; 16:1840–1845. [PubMed: 16185890]

Figure 1.

ESI mass spectra of a solution containing 1% molar fraction of [ala]₃-leucine enkephalin in leucine enkephalin (2.4 mM total peptide concentration) measured at external accumulation hexapole offset potentials of: a) 2.8 b) 3.6 and c) 4.0 V. Regions of the spectra showing molecular clusters are inset.

Figure 2.

a) Normalized abundances of protonated A_3LE (triangles) and LE (squares) from ESI mass spectra as a function of A_3LE molar fraction for solutions with 2.4 mM total peptide concentration. b) Molar fractions of A_3LE calculated from the abundances of the protonated molecular ions from an equimolar solution with LE as a function of total peptide concentration. Error bars correspond to the standard deviation of three replicate mass spectra obtained from three different nanoelectrospray tips at each solution concentration. Dashed lines represent the ideal trend based on solution concentration.

Figure 3.

Expanded region of ESI mass spectrum in Figure 1a showing homogeneous and heterogeneous clusters with 19 constituent molecules corresponding to $[19\text{LE} + 5\text{H}]^{5+}$ and $[18\text{LE} + \text{A}_3\text{LE} + 5\text{H}]^{5+}$. Minor peaks correspond to adducts, such as Na^+ , K^+ and acetate, and other cluster sizes.

Figure 4.

The % molar fraction calculated from the cluster ion intensities assuming statistical incorporation and identical ionization efficiencies obtained from ESI mass spectra of a 1% molar fraction solution of A₃LE in LE (2.4 mM total peptide) as a function of cluster size, n . A dashed line represents the ideal trend based on solution concentration.

Figure 5.

Stickplots representing three theoretical distributions of heterogeneous cluster ions of a given size as a function of increasing cluster size or increasing molar fraction (a–c). Each distribution corresponds to a single cluster size, n , in which each line corresponds to the intensity of a cluster ion that contains some number of impurity molecules, h . The dashed line, N , represents the noise level.

Figure 6.

The % molar fraction of A_3LE in LE obtained from the protonated molecular ions (circles) and from the cluster abundances for $n = 17$ as a function of the solution % molar fraction. The cluster data is linear with a slope of 0.92 whereas the protonated molecular ion data has a slope of 14.2 and is not well fit to a line. Error bars correspond to a standard deviation of three replicate spectra at different DC offset voltages of 2.8, 3.6 and 4.0 V.

Figure 7.

The % molar fraction calculated from cluster ion intensities obtained from ESI mass spectra of a 1% molar fraction BK1-8 with BK (2.4 mM total peptide concentration) as a function of cluster size, n . A dashed line represents the ideal trend based on solute concentration, assuming statistical incorporation of subunits into clusters.

Figure 8.

The % molar fractions calculated from the protonated molecular ions (squares) and cluster ions (triangles) as a function of the % molar fraction in solution for two component mixtures consisting of a) ME with LE, b) BK2-9 with BK, c) BK1-8 with BK and d) KLE with LE. Data from clusters with $n = 12, 15, 15$ and 20 subunits were used for these respective mixtures. Best-fit slopes of 0.95, 1.19, 1.08 and 1.10 are obtained for the cluster data a–d, respectively (see text).

Table 1

Molar Fraction Calculated from Relative Protonated Molecular Ion Intensities from Equimolar solutions

Total Peptide Concentration	Bradykinin		Leucine Enkephalin	
	Bradykinin 1-8	Bradykinin 2-9	[Ala] ₃ Enkephalin	Lysine Enkephalin
2.4 mM	68% ± 7%	55% ± 3%	80% ± 1%	71% ± 1%
12. μM	51% ± 5%	51% ± 3%	66% ± 1%	43% ± 4%
1.2 μM	55% ± 7%	67% ± 5%	50% ± 1%	38% ± 1%