Waters Alliance™ Systems for LC/MS ESI/APCI Applications

MaxEnt:: An Advanced Maximum Entropy Based Tool for Disentangling the Electrospray Mass Spectra from Biopolymer Mixtures

Highlights: The Electrospray (ESI) mass spectrum of a biopolymer (e.g. protein, glycoprotein, or oligonucleotide) is composed of a series of multiply-charged ions on the mass-to-charge ratio (m/z) scale. Typically, a 20 kDa protein generates a series containing some 10-20 multiply-charged molecular ion peaks and each protein in the mixture will give rise to its own characteristic series of such peaks. Simplification of the raw data is thus mandatory before interpretation of the complex spectrum can be attempted. If each series from a given component of the mixture could be condensed into a single peak on a true molecular mass scale, the amount of data to be considered would be significantly reduced and the process of interpretation could begin. In 1991, an advanced maximum entropy (MaxEnt) based procedure developed by J. Skilling of MaxEnt Solutions Ltd, Cambridge, UK (1, 3) was incorporated into the Micromass MassLynx software package, and has been employed in this form to solve many problems in biochemistry. Here, the benefits of MaxEnt over alternative methods of processing ESI mass spectra, particularly from biopolymer mixtures, are summarized.

MaxEnt has the following advantages over alternative methods of processing multiply-charged ESI data:

1. MaxEnt automatically finds the molecular masses of the components present in a mixture:
With a completely unknown sample, a first MaxEnt analysis or zero-charge survey is made over a deliberately chosen wide molecular mass range in order to localize the mass range or ranges for a second definitive analysis. Once the approximate masses of the components have been found from the survey, a second definitive run is made over a narrower output mass range (or ranges) using a finer mass scale (0.5-1 Da channel) in order to obtain fully accurate molecular mass values. An example of the latter is illustrated in Fig.1. Here, the original spectrum from the mixture of the proteins and glycoproteins obtained by denaturing the extracellular hemoglobin from the earthworm, Lumbricus terrestris (26, 32) is shown in Fig. 1A. The native Hb has a molecular weight of 3.5 million and is composed of some 15 different proteins and glycoproteins ranging in mass from 16 to 53 kDa. The MaxEnt deconvoluted spectrum shown in Fig. 1B, where each component in the mixture is represented as a single peak on a true molecular mass scale, is much easier to interpret than the original data.

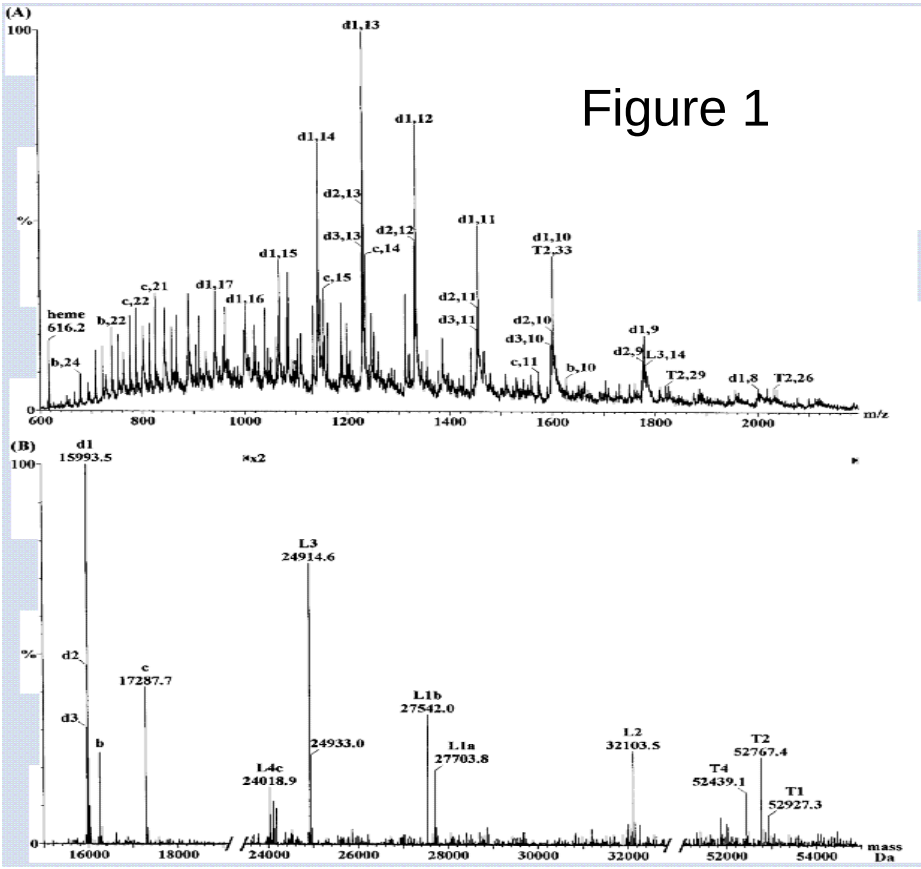

Figure 1. ESI mass spectra of the globins and subunits from the hemoglobin of the earthworm *Lumbricus terrestris* analyzed in denaturing solvent. 1A is the original m/z spectrum and 1B is the MaxEnt processed spectrum on a true molecular mass scale. "d" are monomer globins and "T" are glycosylated disulfide bonded trimers. Linker subunits (L), of which L1 is also glycosylated, are necessary for assembly of the various proteins into the native hemoglobin of molecular mass 3.5 million. In Fig. 1A, the figures after the commas indicate the number of charges on the ions, e.g. d1,13 means component d1 with 13 positive charges. (Copyright 1996 American Chemical Society (32)).

2. MaxEnt enhances the resolution i.e. improves the ability to separate and accurately measure the molecular masses of otherwise unresolved components:

Resolution enhancement in the analysis of hemoglobins by ESI mass spectrometry is due to a basic limit to peak width, and hence resolution, which is determined by the isotopic distribution of the elements in the protein molecule. For hemoglobin chains (15-16 kDa), this basic width at half peak height is 8 Da. With state-of-the-art quadrupole analyzers, the instrumental contribution to peak width is small for the hemoglobin chains, and increases the overall width to a little less than 10 Da, thus giving rise to the practical limit for resolving two hemoglobin chains, without deconvolution, of 12 Da. MaxEnt extends the practical limit for resolving, and accurately measuring the masses of two hemoglobin chains from 12 to 6 Da. Fig. 2 shows an example of resolution enhancement by MaxEnt from the analysis of a heterozygote for the a-chain variant Le Lamentin, [a 20 (HisÝGln)]. Here, the variant a-chain (Mr=15117.4) is not resolved from the normal a-chain (Mr=15126.4) in the original data since the mass difference is only 9 Da (Fig. 2A and inset). After deconvolution by MaxEnt (Fig. 2B), the two a-chains are clearly resolved allowing their mass difference to be accurately determined.

3. MaxEnt improves the signal-to-noise ratio:

MaxEnt has the power to extract useful zero-charge spectra from noisy multiply-charged m/z data. MaxEnt is not restricted to processing biopolymer data but can be used to process any data containing two or more ions from the same molecule, provided they carry consecutive numbers of charges. Fig. 3 shows the result of employing MaxEnt to produce the isotopically resolved zero-charge spectrum of a poly-sulphonated compound (Mr=1815) from partly resolved multiply-charged data. The original spectrum (Fig. 3A and inset) shows predominantly an ion with 6 negative charges together with less intense ions having 5 and 4 charges. The MaxEnt processed spectrum (Fig. 3B and insets) shows the zero-charge spectrum with the isotope peaks almost fully resolved. Moreover, the isotope pattern closely resembles the expected isotope pattern for this molecule of elemental composition C 61 H 40 N 19 O 26 S 8 Cl 3 .

Figure 2

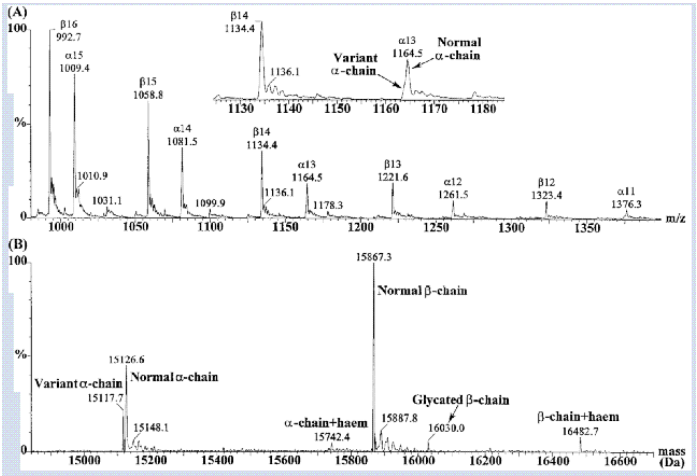

Figure 3

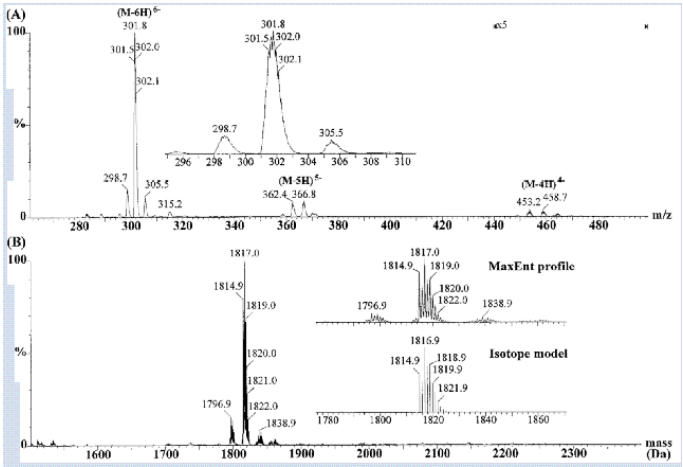

Figure 2. Data from the blood of a heterozygote for the variant hemoglobin Le Lamentin, [a20 (His ÝGln)]. In the original m/z spectrum (Fig. 2A and inset), the variant and normal a-chains are not resolved, whereas after processing by MaxEnt (Fig. 2B), they are clearly resolved to reveal their 9 Da mass difference. These data were produced from whole blood simply diluted 500 fold in 1:1 water:acetonitrile containing 0.2% formic acid.

Figure 3. Electrospray data from a poly-sulphonated compound of molecular mass 1815. A, the original m/z spectrum and B, the zero charge spectrum produced by MaxEnt. Note that in the original data (Fig. 3A inset), the peak with six negative charges is only partly isotopically resolved, whereas in the MaxEnt spectrum, the isotope peaks are clearly resolved (Fig. 3B inset). Furthermore, the experimental isotope pattern closely resembles the isotope pattern expected for this molecule.