MaxEnt: An Advanced Maximum Entropy Based Tool for Disentangling the Electrospray Mass Spectra from Biopolymer Mixtures. A Brief Description and Bibliography.

Authors: B.N. Green and R.W.A. Oliver

Introduction
The electrospray mass spectrum from 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. On this scale, the positions of the multiply-charged positive ions from a single protein or glycoprotein of molecular mass Mᵢ are given by m/z = (Mᵢ+nH)/n, where H is the mass of the proton and n is an integer in a series of consecutive integers. Typically, a 20 kDa protein generates a series containing some 10-20 multiply-charged molecular ion peaks. It follows that the electrospray mass spectrum from a mixture of proteins can be extremely complex, since 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.

Various methods have been described in the literature for carrying out this deconvolution procedure and their drawbacks discussed. In 1991, an advanced maximum entropy (MaxEnt) based procedure was developed by J. Skilling of MaxEnt Solutions Ltd, Cambridge, UK. Subsequently, this MemSys5 maximum entropy algorithm 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 electrospray mass spectra, particularly from biopolymer mixtures, are first summarised. This summary is followed by a bibliography of over 50 papers in which MaxEnt has been employed to deconvolute the electrospray mass spectra from a wide range of biological samples.
The Benefits of MaxEnt

MaxEnt has the following distinct advantages over alternative methods of processing multiply-charged electrospray 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 localise the mass range or ranges for a second definitive analysis. MaxEnt is computationally intensive and the processing time increases with the number of data points (channels) in the zero-charge (output) spectrum. Consequently, when making a wide mass range survey, it is usual to employ a relatively coarse mass scale (5-10 Da/channel) to keep the processing time to within 5 minutes or so. However, when this is done, the molecular weights are not calculated with full accuracy. Therefore, 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 haemoglobin from the earthworm, *Lumbricus terrestris*\(^\text{26, 32}\) 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. These data were acquired on a Micromass Quattro II.

2. MaxEnt enhances the resolution i.e.

improves the ability to separate and accurately measure the molecular masses of otherwise unresolved components.

The arguments for needing resolution enhancement in the analysis of haemoglobins by electrospray mass spectrometry have been described in an excellent review by Shackleton *et al.*\(^\text{8}\). In essence, there is a basic limit to peak width, and hence resolution, which is determined by the isotopic distribution of the elements in the protein molecule. For the haemoglobin chains (15-16 kDa), this basic width at half peak height is 8 Da. With state-of-the-art quadrupole analysers, the instrumental contribution to peak width is small for the haemoglobin chains, and increases the overall width to a little less than 10 Da, thus giving rise to the

![Figure 1. Electrospray mass spectra of the globins and subunits from the haemoglobin of the earthworm *Lumbricus terrestris* analysed in denaturing solvent. A, the original m/z spectrum and B, the MaxEnt processed spectrum on a true molecular mass scale. d are monomer globins and T glycosylated disulphide bonded trimers. Linker subunits (L), of which L1 is also glycosylated, are necessary for assembly of the various proteins into the native haemoglobin 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)).](image-url)
practical limit for resolving two haemoglobin chains, without deconvolution, of 12 Da. The Shackleton review (pp 148-151) shows how MaxEnt extends the practical limit for resolving and accurately measuring the masses of two haemoglobin chains from 12 to 6 Da.

Fig. 2 shows an example of resolution enhancement by MaxEnt from the analysis of a heterozygote for the α-chain variant Le Lamentin, [α20 (His→Gln)]. Here, the variant α-chain (Mr=15117.4) is not resolved from the normal α-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 α-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.

Although the majority of publications in the Bibliography involve the analysis of proteins, it should be stressed that 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 (M_r=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. Data from the blood of a heterozygote for the variant haemoglobin Le Lamentin, [α20 (His→Gln)]. In the original m/z spectrum (Fig. 2A and inset), the variant and normal α-chains are not resolved, whereas after processing by MaxEnt (Fig. 2B), they are clearly resolved to reveal their 9 Da mass difference. These Quattro II 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.
The following is a Bibliography of published papers in which MaxEnt has been used to deconvolute electrospray ionisation mass spectra.


32. Giant hexagonal bilayer hemoglobins. Lamy, J.N., Green, B.N., Toulmond, A., Wall, J.S., Weber, R.E., Vinogradov, S.N. Fig. 1 reprinted with permission from Chem. Rev. 96, 3113-3124 (1996). Copyright 1996 American Chemical Society


Discussion of Bibliography

The wide range and exact nature of the biological mixtures analysed in which MaxEnt procedures have been employed in order to deconvolute their electrospray mass spectra can be readily ascertained by a study of the titles of the papers given in the bibliography. If this information was not given by the original authors of the paper then it has been added to the title in square brackets. At the present time, the most common mixtures studied by this method are those of various haemoglobins, possibly because, as Shackleton and Witkowska comment in paper 45, these proteins afford “dream molecules” for the mass spectrometrist. A study of the adjacent Index of Journals shows that MaxEnt applications have been published in a very wide range of journals, some 27 in total, with the two most popular journals being the two most important biochemical journals, namely the Journal of Biological Chemistry and the Biochemical Journal.

Finally, a study of the country of origin of the various papers forming the bibliography showed that at present MaxEnt is most widely used in the UK (39 papers) followed by American laboratories (18 papers). This study also showed that papers have been published from laboratories in Australia (3 papers), Belgium (2 papers), Denmark (3 papers), France (4 papers), Japan (3 papers), Taiwan (4 papers) and Spain (1 paper) and hence it may be safely concluded that the MaxEnt procedures have been adopted worldwide.

Acknowledgements.

The authors would like to thank Dr T. M. Reynolds, Clinical Chemistry Dept., Burton Hospital, Burton upon Trent, Staffordshire for permission to use the Hb Le Lamentin data and BASF plc., P.O. Box 4, Hexagon House, Blackley, Manchester for permission to use the poly-sulphonated compound data.

INDEX OF JOURNALS

<table>
<thead>
<tr>
<th>Journal</th>
<th>Reference Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical Biochemistry</td>
<td>55, 56.</td>
</tr>
<tr>
<td>Analytical Chemistry</td>
<td>45, 47.</td>
</tr>
<tr>
<td>American Laboratory</td>
<td>2.</td>
</tr>
<tr>
<td>Biochemical Journal</td>
<td>10, 18, 19, 20, 33, 44.</td>
</tr>
<tr>
<td>Biochemical Society Transactions</td>
<td>35.</td>
</tr>
<tr>
<td>Biochemistry</td>
<td>9, 25, 38, 43.</td>
</tr>
<tr>
<td>Biochimica et Biophysica Acta</td>
<td>30, 51.</td>
</tr>
<tr>
<td>Bioconjugate Chemistry</td>
<td>48.</td>
</tr>
<tr>
<td>Biological Mass Spectrometry</td>
<td>13.</td>
</tr>
<tr>
<td>Blood</td>
<td>46.</td>
</tr>
<tr>
<td>Chemical Communications</td>
<td>36.</td>
</tr>
<tr>
<td>Chemical Research and Toxicology</td>
<td>21.</td>
</tr>
<tr>
<td>Chemical Reviews</td>
<td>32.</td>
</tr>
<tr>
<td>European Journal of Biochemistry</td>
<td>54.</td>
</tr>
<tr>
<td>European Mass Spectrometry</td>
<td>17, 28.</td>
</tr>
<tr>
<td>International Journal Biochemistry</td>
<td>49.</td>
</tr>
<tr>
<td>and Cell Biology</td>
<td></td>
</tr>
<tr>
<td>Journal of the American Chemical Society</td>
<td>41.</td>
</tr>
<tr>
<td>Journal of Bacteriology</td>
<td>22.</td>
</tr>
<tr>
<td>Journal of Biological Chemistry</td>
<td>14, 29, 37, 39, 40, 42.</td>
</tr>
<tr>
<td>Journal of Mass Spectrometry</td>
<td>16.</td>
</tr>
<tr>
<td>Mayo Proceedings</td>
<td>53.</td>
</tr>
<tr>
<td>Planta</td>
<td>50.</td>
</tr>
<tr>
<td>Protein Expression and Purification</td>
<td>34.</td>
</tr>
<tr>
<td>Protein Science</td>
<td>27, 52.</td>
</tr>
<tr>
<td>Rapid Communications in Mass Spectrometry</td>
<td>1, 3, 4, 5, 6.</td>
</tr>
</tbody>
</table>
Certificate No 951387

TECHNICAL NOTE NO. TN202/ICA Version 1

The product information in this document was correct at the time of printing, however specifications are subject to alteration without notice due to the Micromass UK Limited philosophy of continual product innovation and development and any information contained herein should not be construed as a contractual commitment by Micromass UK Limited. All orders are accepted subject to our current Conditions of Sale, copies of which are available on request.