

# Comparison of Methods for Detecting and Deisotoping Weak, High Charge Signals in Data

M.R.Alecio<sup>1</sup>, A.G.Ferrige<sup>1</sup>, L.Pannell<sup>2</sup>, R.S.Ray<sup>1</sup>

<sup>1</sup> Positive Probability Ltd, Isleham, U.K., CB7 5RX;

<sup>2</sup> Mitchell Cancer Institute, University of South Alabama, Mobile AL 36688

## Overview

Most methods that are used for deisotoping have several deficiencies. For example, almost all *Algebraic Methods* over-fit the data and so generate artefacts, particularly where charges overlap. Additionally these methods do not generally estimate the errors in the results. Similarly *Data Reconstruction Methods* that constrain the result to the same intensity as the data can produce serious artefacts. Error assessments in some of these methods are highly dependent on a user-supplied random number seed and can therefore be misleading. In this work we compare the abilities of the algebra-based Mascot Distiller deisotoping program and the MaxEnt3 and *ReSpect*<sup>TM</sup> data reconstruction methodologies employed respectively by Waters and Positive Probability Ltd (PPL) to produce reliable deisotoped mass spectra.

## Introduction

The MS identification of macromolecules often depends on the detection of diagnostic fragments for database search. For example, ESIMS is commonly used to identify peptides in protein digests. However, the spectra are generally highly complex and difficult to interpret. Furthermore, conventional deisotoping methods are prone to error and artefacts so that high charge states, particularly when they are weak, may be assigned the wrong charge or may not be identified at all. In addition, thresholding the data risks losing weak but significant ions. Consequently the ability to incorporate often highly diagnostic large peptide fragments in the searching process may be compromised. These complications make it difficult to choose the best method for processing complex spectra. In this technical work we do not attempt to identify proteins from the LCMS of digests. Instead, we explore and compare 3 available methods for processing spectra and report in particular on the detection of weak, high charge ions and the presence of artefacts in the results.

## Methods

**a) Algebraic Methods:** These methods work by performing calculations directly on the data. These methods make use of a model that defines the way the isotopes vary in intensity with  $m/z$  and  $Z$ , based on an empirical formula that represents the type of sample under investigation. The methods generally solve the problem using simultaneous equations but they cannot take noise and measurement errors into account. Therefore, any discrepancy between the observed isotope intensity distributions and the theoretical patterns, even if due to noise, must be accounted for by incorporating additional peaks in the result. This leads to serious over-fitting of the data and the generation of artefacts, with potentially confusing results. In addition, Distiller identifies signals in the data using a centroiding method, which has well-known reliability and other problems. The information that can be retrieved in this way is limited by the instrument resolution and isotope clusters can be resolved up to around  $Z=10$  on a ToF. Although Distiller's maximum limit of  $Z=8$  as its maximum applies to most data, it does mean that higher charge states that are resolved will not be interpreted correctly.

**b) Data Reconstruction Methods:** These methods are less prone to noise because instead of performing calculations on the data, they adjust a model of the measurement until it matches the data within the noise. Therefore, they do not of necessity over-interpret data and in principle only ions or masses for which there is evidence are reported. They are most efficient when a spectrum deconvolution is part of the process, since this enables higher charge states to be analysed. However, they may have shortcomings. For example, a method may require the result to have the same intensity as the data. This means that noise and any irrelevant peaks must contribute intensity to the result, leading to artefacts. Another problem is that the mathematically most plausible result is not necessarily the most realistic. This can arise if the method does not support an adequate model of the experimental technique. For example, even charge states may be interpreted as having only half their true charge, generating harmonic masses at half the true mass. Additional charges that are not present may erroneously be reconstructed, along with the insertion of additional isotopes to generate high mass harmonics. Finally it is possible for a higher charge state to be assigned the wrong charge so that its mass is calculated incorrectly.

## Experimental

A tryptic digest of human urine was analysed by LCMS using a Waters QToF instrument. The resulting data contained a large number of high mass peptides. Two typical example ESIMS spectra for further study were obtained by co-adding 9 consecutive scans in which high mass peptides were present. The resulting spectra were complex, containing strong and weak signals from peptides and glycopeptides. Charge states ranged from Z1-Z11 and there were numerous overlaps between different charges.

The spectra were analysed independently in 2 or 3 laboratories using commercial methods including Mascot Distiller, MaxEnt3 (Waters) and ReSpect (PPL), using input parameters set according to the respective manuals. The maximum charge for MaxEnt3 and ReSpect was set to Z=12 since charges of Z=11 were observed in the data. For Mascot Distiller the maximum charge was Z=8. To simplify the comparison of high masses in the results, the co-adds were deisotoped to mass and where outputs were M+1 these were converted to M (zero-charge mass).

## Results

Figure 1 shows the test LCMS data. The run time was 100 min. and the m/z 180-2000. The central scan numbers for the co-adds of 9 scans are shown. Figure 2 shows the co-added spectra that were investigated further. The expansion (Figure 3) illustrates the complexity of the data for the 1678 co-add. The S/N is generally high but is low for the weak ions. Charges range from Z=1-10 and charge overlap is evident. The charge states present are indicted. These were confirmed by finding other charges for the same masses.

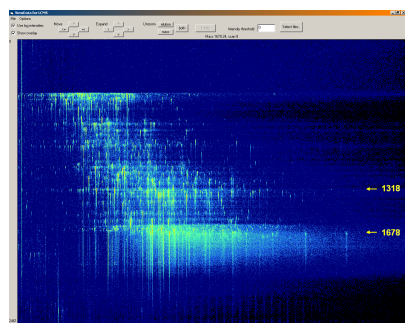


Figure 1. LCMS data

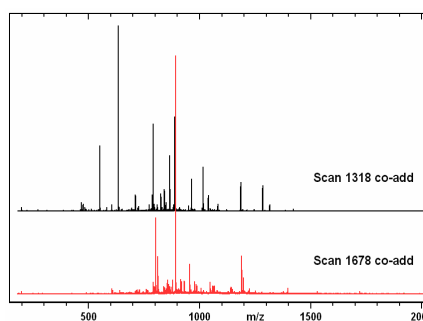


Figure 2. Co-added spectra

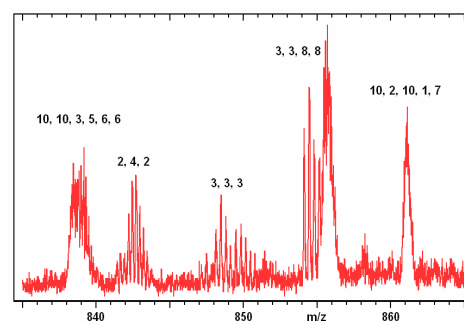


Figure 3. Part of 1678 co-add

The zero-charge results for the two co-adds are shown in Figures 4 & 5. For presentation purposes the intensities of signals reported by Distiller were multiplied by 5 to make them comparable with the results of the other programs.

Result quality was assessed from the data to establish the evidence in support of each of the masses identified by the three programs.

We combined the 10 most intense high masses reported by each program into a single table of candidate masses for each co-add (Tables 1 & 2). Weak satellite peaks at  $\pm 1$  Da, which arise because the peptide empirical formula is a compromise, were ignored. The ReSpect results were filtered to 99% confidence.

*Note: For Distiller, masses were found in the 1318 co-add at 3848.12 & 3848.13, intensities 3321 & 3246. Similarly, masses were found at 8369.52 & 8369.60 for 1678 co-add, intensities 13176 & 9030. These have been combined in the tables.*

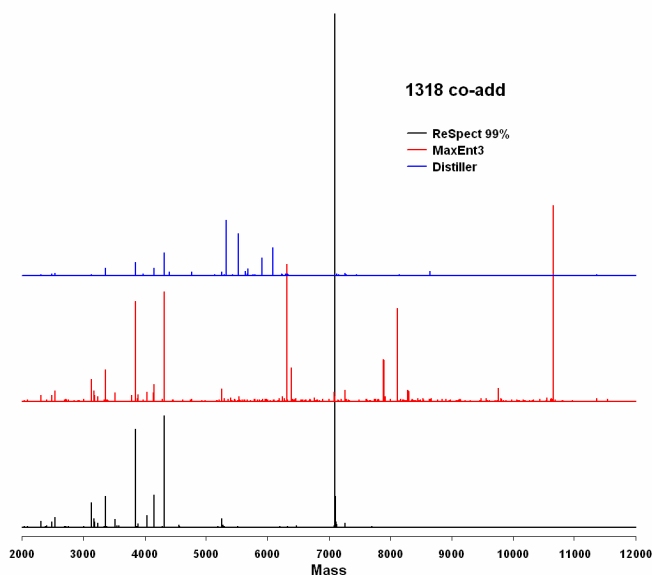


Figure 4. Zero-charge results for 1318 co-add

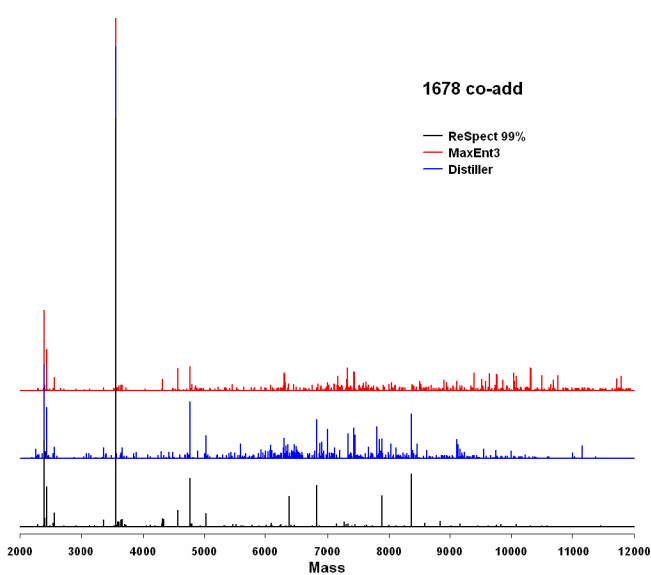


Figure 5. Zero-charge results for 1678 co-add

Mass	Intensities		
	MaxEnt3	Distiller	ReSpect
2542.37			4573
3139.69	5376		5990
3175.65			4352
3361.69	7692	1840	7532
3848.12	24070	6567	23495
4038.05			3033
4151.30		1853	7979
4322.26	26320	5513	26851
5256.58			4246
5326.27		13226	
5523.55		10153	
5651.57		963	
5681.32		1691	
5918.16		4261	
6087.16		6606	
6312.51	32810		
6391.39	8002		
7101.50			122395
7890.48	10190		
8115.76	22300		
10651.16	46790		
10656.19	13080		

Mass	Intensities		
	ReSpect	Distiller	MaxEnt3
3563.03	120012	121376	113500
4322.25	2529		
4349.54	2333		
4566.61			6599
4567.65	4849		
4769.74			7254
4770.72	14338	16834	
5023.72		6809	
5025.93	3848		
6313.26			5184
6388.52	9081		
6832.91		11404	
6833.69	12260		
7014.71		8753	
7323.40			6707
7335.22		7411	
7439.57		9155	5629
7453.33		7127	
7809.93		9614	
7891.25	9182		
8369.52	15695	22206	
9395.71			5323
9644.28			5082
10038.14			5274
10314.97			6736

## Analysis of Results

In order to test the validity of the reported candidate masses, we first calculated the  $m/z$  values of all their potential charge states up to  $Z = 11$  (Tables 3 & 6) that could be present in the data. We then manually searched the co-adds for evidence of these ions to support the reported masses.

In some cases a predicted isotope pattern was found to be incorrectly aligned with the data, being displaced by one isotopic interval, which produced a 1 Da error in the full mass. There were also cases where predicted ions overlaid a cluster in the data but the assigned charge was double the true value, giving rise to 'harmonic' masses of approximately twice the true mass in the results. There were also a few cases where MaxEnt3 introduced small errors in the reconstructed masses even though ions and their charge state were correctly identified. For many Distiller and MaxEnt3 masses there was no evidence at all.

### 1318 co-add

The possible charge states (up to Z=11) that may theoretically be present in the data are shown in Table 3. Ions that are not actually present are highlighted in grey, while those that are clearly present are in green, and weak ions are orange. The pink highlight indicates masses for which there is **no evidence in the data**. These ions must therefore be considered to be artefacts. Significantly, the most intense mass reconstructed by ReSpect (M=7101) is absent in the MaxEnt3 and Distiller results. However, there is no evidence in the data for the three most intense masses found by Distiller or for the most intense mass found by MaxEnt3.

**Table 3. Charges in Data for Found Masses**

Mass	Intensities			Note	Possible charges								
	MaxEnt3	Distiller	ReSpect		11	10	9	8	7	6	5	4	3
2542.37			4573	3	232.13	255.25	283.49	318.80	364.20	424.74	509.48	636.60	848.47
3139.69	5376		5990	2	286.43	314.98	349.86	393.47	449.53	524.29	628.94	785.93	1047.57
3175.65			4352	1	289.70	318.57	353.86	397.96	454.67	530.28	636.14	794.92	1059.56
3361.69	7692	1840	7532		306.62	337.18	374.53	421.22	481.25	561.29	673.35	841.43	1121.57
3848.12	24070	6567	23495	8	350.84	385.82	428.58	482.02	550.74	642.36	770.63	963.04	1283.72
4038.05			3033	6	368.10	404.81	449.68	505.76	577.87	674.02	808.62	1010.52	1347.02
4151.30		1853	7979	6	378.40	416.14	462.26	519.92	594.05	692.89	831.27	1038.83	1384.77
4322.26	26320	5513	26851		393.94	433.23	481.26	541.29	618.47	721.38	865.46	1081.57	1441.76
5256.58			4246	3	478.88	526.67	585.07	658.08	751.95	877.11	1052.32	1315.15	1753.20
5326.27		13226		6	485.21	533.64	592.82	666.79	761.90	888.72	1066.26	1332.58	1776.43
5523.55		10153		4	503.15	553.36	614.74	691.45	790.09	921.60	1105.72	1381.90	1842.19
5651.57		963		7	514.79	566.16	628.96	707.45	808.37	942.94	1131.32	1413.90	1884.86
5681.32		1691		4	517.49	569.14	632.27	711.17	812.62	947.89	1137.27	1421.34	1894.78
5918.16		4261		4	539.02	592.82	658.58	740.78	846.46	987.37	1184.64	1480.55	1973.73
6087.16		6606		4	554.39	609.72	677.36	761.90	870.60	1015.53	1218.44	1522.80	
6312.51	32810			4	574.87	632.26	702.40	790.07	902.80	1053.09	1263.51	1579.14	
6391.39	8002			4	582.04	640.15	711.16	799.93	914.06	1066.24	1279.29	1598.85	
7101.50			122395		646.60	711.16	790.06	888.69	1015.51	1184.59	1421.31	1776.38	
7890.48	10190			4	718.32	790.06	877.73	987.32	1128.22	1316.09	1579.10	1973.63	
8115.76	22300			4	738.80	812.58	902.76	1015.48	1160.40	1353.63	1624.16		
10651.16	46790			5	969.29	1066.12	1184.47	1332.40	1522.60	1776.20			
10656.19	13080			5	969.75	1066.63	1185.03	1333.03	1523.32	1777.04			

The mass incorrectly reported by Distiller at 5523 could arise from signals in the data at m/z 790. Figure 6, shows the data (red trace) and deconvolved peak positions (black spikes, deconvolved in ReSpect). This ion evidently has Z=9 which exceeds the maximum charge allowed by Distiller. Furthermore, the C12 isotope, which is unresolved, is unlikely to be detected by centroiding and the remaining pattern may be a better fit for a lower charge. Distiller has therefore interpreted this signal as Z=7 and reports an incorrect mass of 5523 instead of 7101.

Comparing the predicted m/z values (Table 3) with those found in the data, we find several instances of charge misassignment in the Distiller and MaxEnt3 results (Table 4). Such misassignments account for most of the masses found by MaxEnt3 and Distiller for which there is no evidence. However, 4 of the Distiller masses have true charges of 8 and below that are still interpreted incorrectly by the program. These could arise if centroiding errors are excessive.

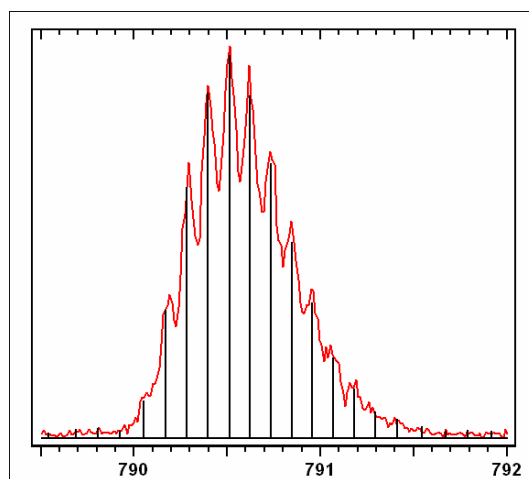


Figure 6. Z9 interpreted as Z7 by Distiller

**Table 4. Charge Misassignments**

Mass	True Z	MaxEnt3 Z	Distiller Z
5326.27	8		6
5523.55	9		7
5681.32	10, 5		8, 4
5918.16	6		5
6087.16	6		5
6312.51	9	8	
6391.39	10	9	
7890.48	9	10	
8115.76	7	8	
10651.16	6	9	
10656.19	6	9	

The remaining MaxEnt3 masses for which there is no evidence appear to be harmonic masses (highlighted in red in Table 4). These could arise because MaxEnt3 generates the most probable result in a mathematical sense, and may thus add unobserved charge states and isotope peaks to improve the fit to the data.

The ReSpect results appear to be free of charge misassignment and harmonic artefacts, and all the masses reported in Table 3 are supported by positive evidence in the data. For example, the predicted isotope patterns for the different charge states of the ReSpect mass at 7101 are shown in Figure 7.

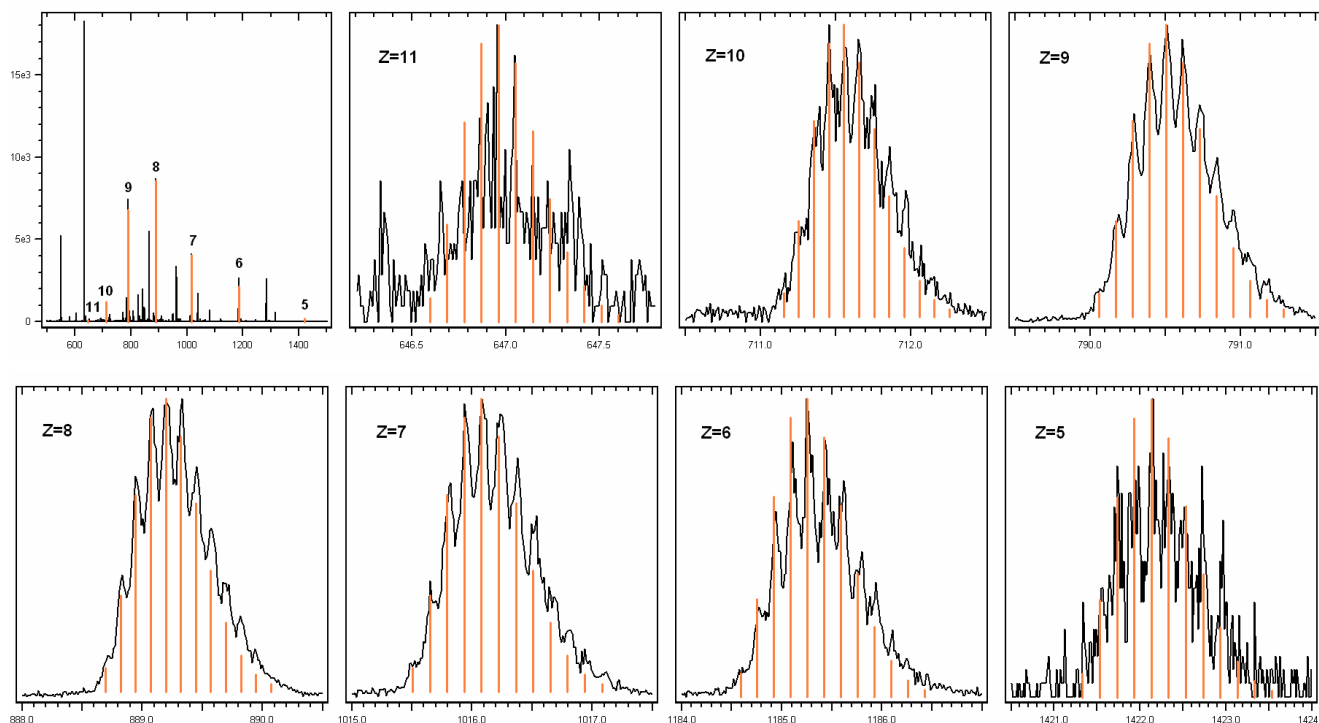


Figure 7. Evidence for the mass at 7101 reconstructed by ReSpect

The ReSpect mass at 7101 is totally absent from the MaxEnt3 and Distiller results, yet it is the most intense ReSpect mass and is strongly supported by the data.

The main reason why both Distiller and MaxEnt3 miss such an intense series of unambiguous charges (Figure 7) is that many of the charges are misassigned (Table 5). For example,  $m/z$  888.72 must have been assigned  $Z=6$  to obtain the reported mass when the correct charge is  $Z=8$ . Significantly, such charge misassignments account for all high masses reported by MaxEnt3 and all but one high mass reported by Distiller.

**Table 5. Charge Misassignments for  $M=7101$  by Distiller and MaxEnt3**

Mass	MaxEnt3	Distiller	ReSpect	Note	11	10	9	8	7	6	5	4	3
5326.27		13226		6	485.21	533.64	592.82	666.79	761.9	888.72	1066.26	1332.58	1776.43
5523.55		10153		4	503.15	553.36	614.74	691.45	790.09	921.6	1105.72	1381.9	1842.19
5651.57		963		7	514.79	566.16	628.96	707.45	808.37	942.94	1131.32	1413.9	1884.86
5681.32		1691		4	517.49	569.14	632.27	711.17	812.62	947.89	1137.27	1421.34	1894.78
5918.16		4261		4	539.02	592.82	658.58	740.78	846.46	987.37	1184.64	1480.55	1973.73
6087.16		6606		4	554.39	609.72	677.36	761.9	870.6	1015.53	1218.44	1522.8	
6312.51	32810			4	574.87	632.26	702.4	790.07	902.8	1053.09	1263.51	1579.14	
6391.39	8002			4	582.04	640.15	711.16	799.93	914.06	1066.24	1279.29	1598.85	
7101.5			122395		646.6	711.16	790.06	888.69	1015.51	1184.59	1421.31	1776.38	
7890.48	10190			4	718.32	790.06	877.73	987.32	1128.22	1316.09	1579.1	1973.63	
8115.76	22300			4	738.8	812.58	902.76	1015.48	1160.4	1353.63	1624.16		
10651.16	46790			5	969.29	1066.12	1184.47	1332.4	1522.6	1776.2			
10656.19	13080			5	969.75	1066.63	1185.03	1333.03	1523.32	1777.04			

The  $m/z$  values present in the data but assigned the wrong charge by Distiller and MaxEnt3 are highlighted in yellow and orange, respectively. Values supported by the data are highlighted in green.

### 1678 co-add

Table 6 shows all possible charges that may be present in the data window. Table details are the same as for Table 4. Once again, there is no evidence in the data for the majority of the results reported by Distiller (7 masses) or MaxEnt3 (9 masses). In fact, there is only evidence for one MaxEnt3 mass above 3000 Da.

A comparison of MaxEnt3 and Distiller masses with the predicted  $m/z$  values again shows numerous charge misassignments (Table 7), accounting for most of the MaxEnt3 and Distiller masses as before. Three ions are within Distiller's maximum charge but were all identified incorrectly, possibly because of poor S/N.

Mass	MaxEnt3	Distiller	ReSpect	Note	11	10	9	8	7	6	5	4	3
3563.03	113500	121376	120012		324.92	357.31	396.90	446.39	510.01	594.85	713.61	891.76	1188.68
4322.25			2529	1	393.94	433.23	481.26	541.29	618.47	721.38	865.46	1081.57	1441.76
4349.54			2333		396.42	435.96	484.29	544.70	622.37	725.93	870.92	1088.39	1450.85
4566.61	6599				416.15	457.67	508.41	571.83	653.38	762.11	914.33	1142.66	1523.21
4567.65			4849		416.25	457.77	508.52	571.96	653.53	762.28	914.54	1142.92	1523.56
4769.74	7254				434.62	477.98	530.98	597.23	682.40	795.96	954.96	1193.44	1590.92
4770.72		16834	14338		434.71	478.08	531.09	597.35	682.54	796.13	955.15	1193.69	1591.25
5023.72		6809		7	457.71	503.38	559.20	628.97	718.68	838.29	1005.75	1256.94	1675.58
5025.93			3848		457.91	503.60	559.44	629.25	719.00	838.66	1006.19	1257.49	1676.32
6313.26	5184			4	574.94	632.33	702.48	790.17	902.90	1053.22	1263.66	1579.32	
6388.52			9081		581.78	639.86	710.84	799.57	913.65	1065.76	1278.71	1598.14	
6832.91		11404			622.18	684.30	760.22	855.12	977.14	1139.83	1367.59	1709.23	
6833.69			12260		622.25	684.38	760.31	855.22	977.25	1139.96	1367.75	1709.43	
7014.71		8753		4	638.71	702.48	780.42	877.85	1003.11	1170.13	1403.95	1754.69	
7323.40	6707			4	666.77	733.35	814.72	916.43	1047.21	1221.58	1465.69	1831.86	
7335.22		7411		4	667.85	734.53	816.03	917.91	1048.90	1223.54	1468.05	1834.81	
7439.57		5629	9155	4	677.33	744.97	827.63	930.95	1063.80	1240.94	1488.92	1860.90	
7453.33		7127		4	678.58	746.34	829.16	932.67	1065.77	1243.23	1491.67	1864.34	
7809.93		9614		4	711.00	782.00	868.78	977.25	1116.71	1302.66	1562.99	1953.49	
7891.25			9182	2	718.39	790.13	877.81	987.41	1128.33	1316.22	1579.26	1973.82	
8369.52		22206	15695		761.87	837.96	930.95	1047.20	1196.65	1395.93	1674.91		
9395.71	5323			4	855.16	940.58	1044.98	1175.47	1343.25	1566.96	1880.15		
9644.28	5082			4	877.76	965.44	1072.59	1206.54	1378.76	1608.39	1929.86		
10038.14	5274			7	913.57	1004.82	1116.36	1255.78	1435.03	1674.03			
10314.97	6736			4	938.73	1032.50	1147.12	1290.38	1474.57	1720.17			

Mass	True Z	MaxEnt3 Z	Distiller Z
6313.26	10	8	
7014.71	9		8
7323.40	8	7	
7335.22	3		6
7439.57	9		8
7453.33	6		7
7809.93	7		8
9395.71	8	11	
9644.28	9	11	
10314.97	8	9	

In this study we find that the ReSpect method identifies charges correctly even when there is serious overlap between isotope clusters of different charge. Furthermore there is always positive evidence in the data for all reconstructed masses with the possible exception of those that are particularly weak and at the selected confidence limit.

However, there are several puzzling results that are reported by both MaxEnt3 and Distiller, that we now describe. Without knowledge of full methodological details, it is only possible to speculate on the reasons for these frequent and unpredictable peculiarities that could lead to a serious misinterpretation of the results.

**Puzzle 1:**

MaxEnt3 and ReSpect reconstruct respective masses of 4769.74 and 4770.72. MaxEnt3 frequently reports masses low or high by 1 Da even when it has correctly reconstructed the charge state.

In this example there is a strong Z=5 charge with the C12 at m/z 955.15 (Figure 9). This is correctly identified by ReSpect but MaxEnt3 is low by the isotope separation, equivalent to 1 Da at the full mass.

This common behaviour of MaxEnt3 is currently inexplicable.

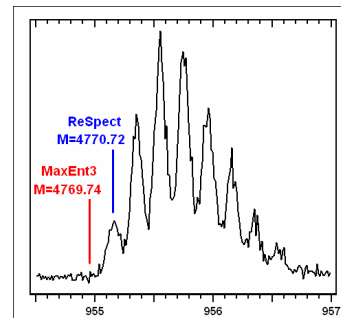


Figure 9. C12 positions for MaxEnt3 & ReSpect

**Puzzle 2:**

In this example, Distiller finds a mass of 7335.22. The only ion that can account for this is a Z=3 at m/z 1223.54 but it would have to be Z=6 to give the found mass. The data are shown in Figure 10. Although the S/N for this ion is not particularly low, most standard peak picking methods are likely to generate rather more centroids than the desired number and some peaks may be split into more than one centroid. Depending on the m/z error limits allowed, the program may or may not combine centroids during its algebraic analysis. If each centroid is separately accounted for, then fitting for Z=6 may provide the best solution, leading to a harmonic mass and a total absence of the true mass.

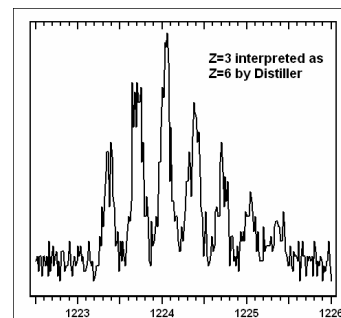


Figure 10. Wrong charge assignment by Distiller

### Puzzle 3:

In this case MaxEnt3 has misassigned a charge state. The data for the relevant charge are shown in Figure 11. Here, there are several overlapping charges. The Z=8 charge is part of a series from Z=9 down to Z=5 for a mass of 6833.69, which are all very clear in the data. MaxEnt3 fails to identify the correct mass altogether. The only ion that can account for its reconstructed mass of 9395.71 would appear to be the overlapped charges shown in Figure 11. The positions of the isotopes for the MaxEnt3 mass are shown in the figure but only Z=11 can account for the mass.

Distiller does correctly identify the charges but generates a mass 0.8 Da too low at 6832.91, possibly because of centroiding errors on lower S/N charges).

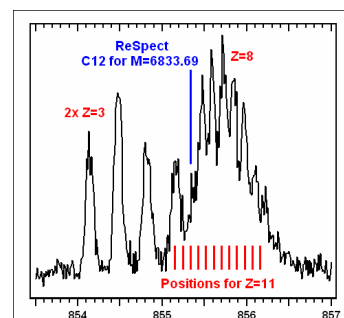


Figure 11. Wrong charge assigned by MaxEnt3

**Puzzle 4:** The strongest mass above 4000 observed in the Distiller and ReSpect results is at 8369.52. However, it is totally absent in the MaxEnt3 result. Charges are present from Z=10 to Z=7. Although the maximum allowed Distiller charge is 8, it has correctly identified the mass from the Z=8 and Z=7 charges. All charges are clear and unambiguous in the data even though those above Z=8 are poorly resolved. Figure 12 shows the Z=8 ion and the predicted isotope positions.

It is indeed surprising that MaxEnt3 should fail to identify a mass from such an obvious charge.

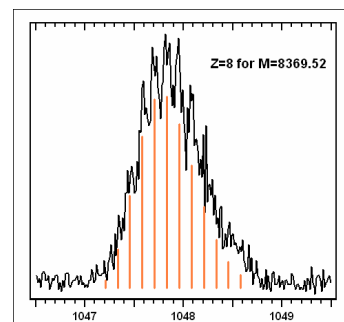


Figure 12. Mass missed by MaxEnt3

## Discussion

All three methods appear to work well for the lower masses. Thus, in the regions M=2000-4000 of Figures 4 & 5, the results appear very similar and a detailed comparison shows that the masses reported by the 3 methods are virtually identical. In this study we have found problems with the higher masses and charges. For example both MaxEnt3 and Mascot Distiller can fail to identify high masses for which there is overwhelming evidence. We have also found that these methods can generate intense masses for which there is no evidence. However, we found positive evidence in the data to support results obtained using the ReSpect algorithm, and that the method generated no strong artefacts. We have also reported on potential reasons why MaxEnt3 and Distiller may miss obvious masses and also produce artefact masses.

**Distiller:** It is surprising is that Distiller can miss some high masses for which there is overwhelming evidence. Based on the data presented here, there is also a tendency for Distiller to generate two almost identical masses instead of a single mass (Tables 1 & 3). One can only presume that this arises from centroiding errors.

**MaxEnt3:** To our knowledge, MaxEnt3 does not take into account any variation in peak width in the data and that a single model is used. The model is therefore a compromise if the peak width varies. For low charges this should not present a serious problem and there is evidence that lower masses are correctly reported. However, high charge states may be incorrectly resolved in the computation so that they appear to have a different charge.

**ReSpect:** ReSpect is a probabilistic data reconstruction method which calculates the deisotoped result into two steps. First the spectrum is deconvolved to give a peak table uncompromised by any variations in noise or peak width. In the deisotoping step a model of isotope intensity distribution which is continuously variable with mass is used for the fitting to the peak table. Unlike the other methods, features that do not fit the constraints are ignored as noise or error, and the intensity in the result may be less than that in the data.

## Conclusions

We have examined the results of deisotoping LCMS data containing high charges, many of which have a low S/N, using MaxEnt3, Mascot Distiller and ReSpect. We have found that:

- The vast majority of high masses reconstructed by MaxEnt3 and Distiller are artefacts arising from misassigning charges. However, there is clear evidence in the data for masses reconstructed by ReSpect. All methods work well for low charges and masses.
- Unlike ReSpect, MaxEnt3 and Distiller fail to report some intense masses for which there is overwhelming evidence. This is probably due to centroid errors and the maximum charge allowed by Distiller and misassigned charges for MaxEnt3. Forcing the result to have the same intensity as the data also introduces substantial noise in MaxEnt3 results.
- For high masses Distiller has a tendency to generate more than one similar mass when a single mass might be expected.

Little use is currently made of high mass peptides for database searching even though many of these can be highly diagnostic. We believe this is because the deisotoping methods commonly used are simply inadequate for the job, sometimes failing to report even clear signals and generating numerous misleading artefactual results, as demonstrated in this comparative study. However, in this work we find that ReSpect reliably reports masses for which there is evidence in the data, even at high charge and low S/N.

### **Acknowledgement**

The authors are grateful to Peter Leopold, BioAnalyte Inc., for his comments and assistance in analysing some of the data.