Electrospray-MS Charge Deconvolutions without Compromise – an Enhanced Data Reconstruction Algorithm utilising Variable Peak Modelling

A.Ferrige1, S.Ray1, R.Alecio1, S.Ye2 and K.Waddell2

1 PPL, Isleham, Cambs, UK.; 2 Applied Biosystems, Framingham, MA 01701, USA

Overview

Data reconstruction methods utilising peak models provide the most detailed results for charge deconvolutions. However, their quality will be compromised for high mass proteins unless the change in peak width with m/z is taken into account. The increased information content of zero-charge results is demonstrated for interferon and a large glycoprotein and this work shows the benefits of accounting for varying peak widths.

A. Using a varying peak model as opposed to a constant model provides more reliable peak tables with smaller errors.

B. Subsequent charge deconvolutions provide cleaner, more highly resolved zero-charge results with both more detail and improved mass errors.

Introduction

The peak width increases with m/z for ESI spectra of high mass proteins in both m/z units and sampling intervals. This change can be by up to at least a factor of 3 on quadrupole based systems. Time of Flight data are somewhat less affected due to the decrease in the number of points/Da with increasing m/z and using an average model will frequently still provide excellent results. However, for heterogeneous high mass data there is the risk that where the peak width is narrow compared with the model, close or overlapped peaks at low m/z will not be resolved. At high m/z where the model may be far too narrow there is the risk that single peaks will be split into more than one component, potentially creating anomalies in the charge deconvolved result. The quality of charge deconvolutions is therefore compromised unless peak width variations are taken into account.

In this work the ***ReSpect***™ data reconstruction algorithm was modified to determine the way the peak profile parameters change with m/z and to accommodate the found peak width and shape variations. The benefits of this improved methodology are illustrated for two proteins. Results reported have been compared with data from the use of a single and constant peak model.

Methods

From two or more relatively crude estimates of the peak profile at different points in the data, the ***ReSpect***™ algorithm is used to determine how the four peak parameters – left width, right width, left shape & right shape – that define a peak model change with m/z. To accomplish this, the data are first Fourier transformed to produce a decaying signal. As its starting point, the program computes the most likely position and intensity of the centroids that would be consistent with the data and the user models. The Fourier transform of the predicted centroids is a non-decaying signal. The convolution of this signal with the correct profile will provide the best possible fit to the data. The algorithm performs this task in a few iterations to provide a highly reliable estimate of the way the four peak profile parameters change with m/z. This knowledge is then used to perform a spectrum deconvolution that is not compromised by any peak width variation.

The data used to show the new technique in operation were obtained from a glycoprotein analysed on a QSTAR® Pulsar Hybrid LC/MS/MS system in electrospray time of flight mode and from the protein called interferon analysed on a Finnigan quadrupole instrument.

Results

Results are presented in two parts

Part 1

Part One shows the data from the Interferon protein. The Interferon was cloned from a single cell line that had been degraded (part of a stability trial) by heating in a moist atmosphere. Water is added to the molecule to give differences of ~18. As more water molecules are added, so the conformation changes allowing more to add on.

The raw data in figure 1 shows the problem in trying to interpret this spectrum. The peaks resulting from the addition of the water are not clearly resolved from each other and the peaks tail off into the noise. The question is how many additions of water molecules are present.

Figure 2 demonstrates the issue of an increasing number of data points across the peak as the mass range increases or as the charge state of the peak decreases. The top trace shows the raw data from the peak with z=19, the middle from z=13, the bottom from z=8. The number of data pts and hence the peak width alters by a factor of 2.25.

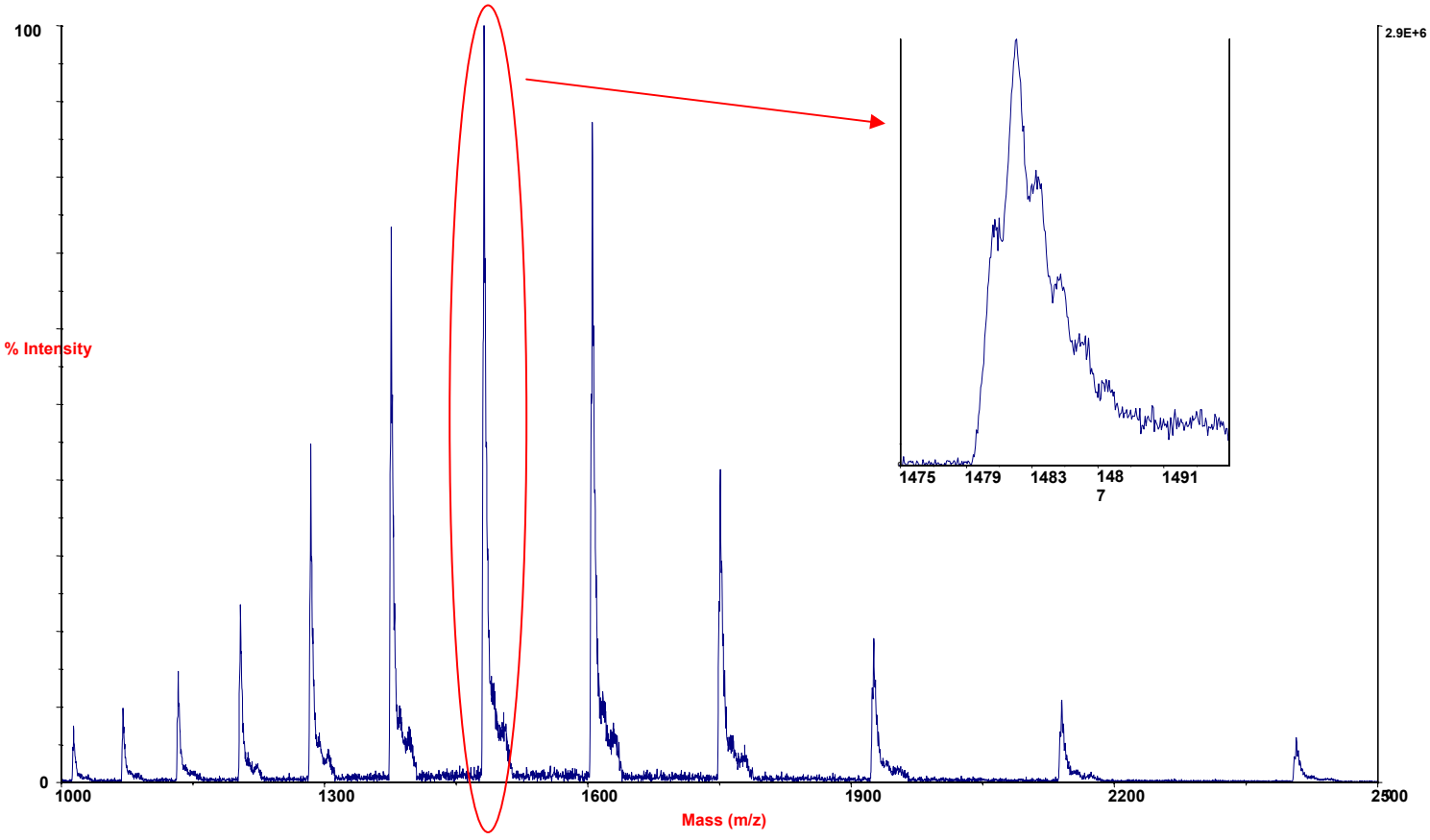


Figure 1: Electrospray MS of Interferon

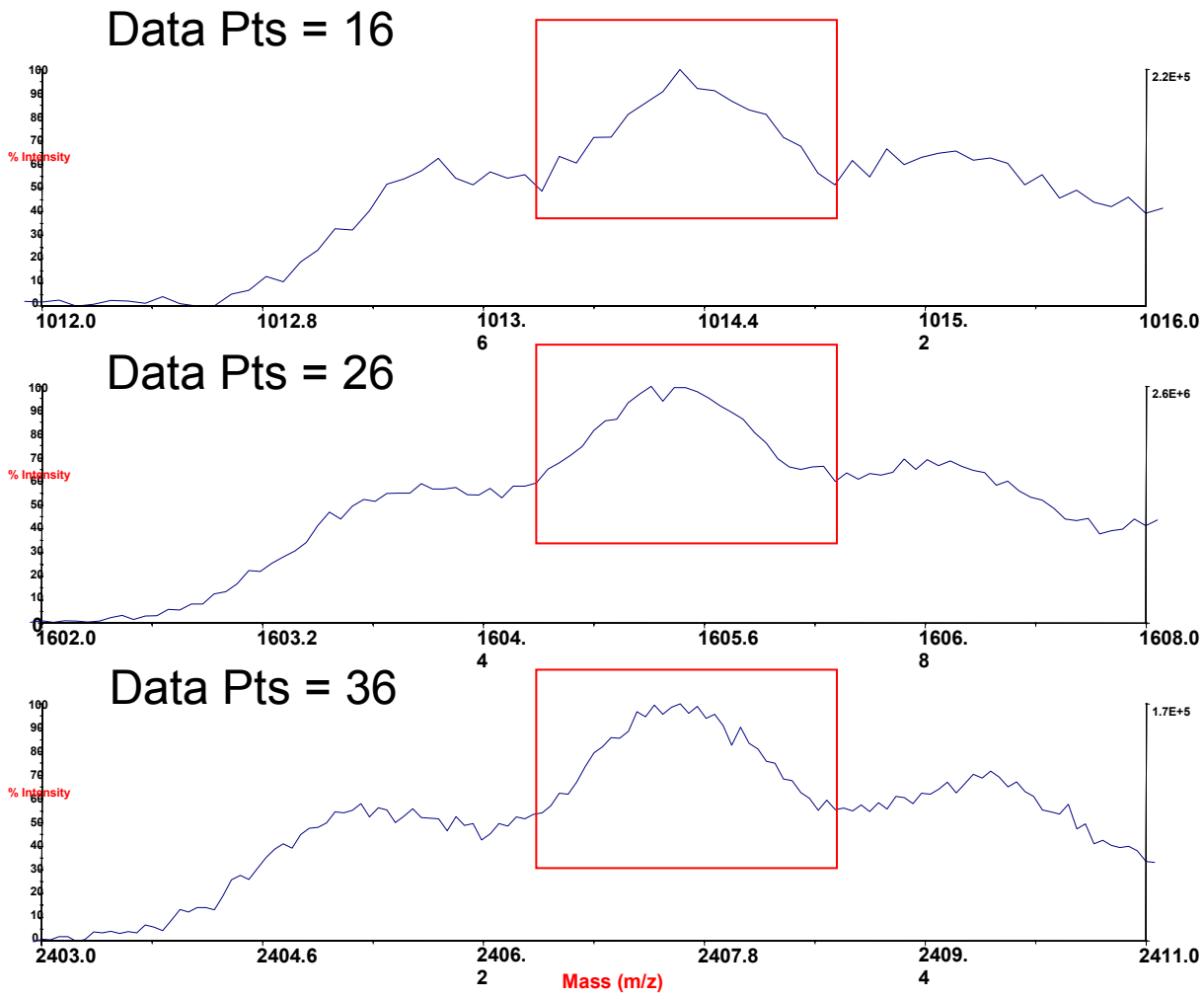


Figure 2: Raw data showing peak width change in data points with m/z

The comparison in the deconvolution of the data can be seen in figure 3. As the peak cluster used for the model for the constant model is that around m/z 1604 (using a similar method to the FT technique mentioned in the Methods), the deconvolved result using the constant peak model is very accurate and provides a slightly sharper and a more confident peak assignment than for the variable model (where the model is derived from a low order polynomial describing the model parameters changing throughout the data). It should be noted that the peak definition in these deconvolved spectra is a representation of the confidence that the program has in the peak being a peak (treat the width as the programs assessment of the error in the peak and the total area as the original peak intensity).

However when the program proceeds to deconvolve the spectra at the two ends of the charge states, the results are very different. At the low mass end or on the peaks at a charge state of 19 (Figure 4) the data shows a marked difference between the use of the variable peak model and the constant model. The constant model is too broad for the true peaks and has lost information in the raw data- the program tries to fit the data to too broad a peak width. In fact the repeat addition of 18 da is lost, which in turn will cause errors in the final calculation of a charge deconvolved spectrum.

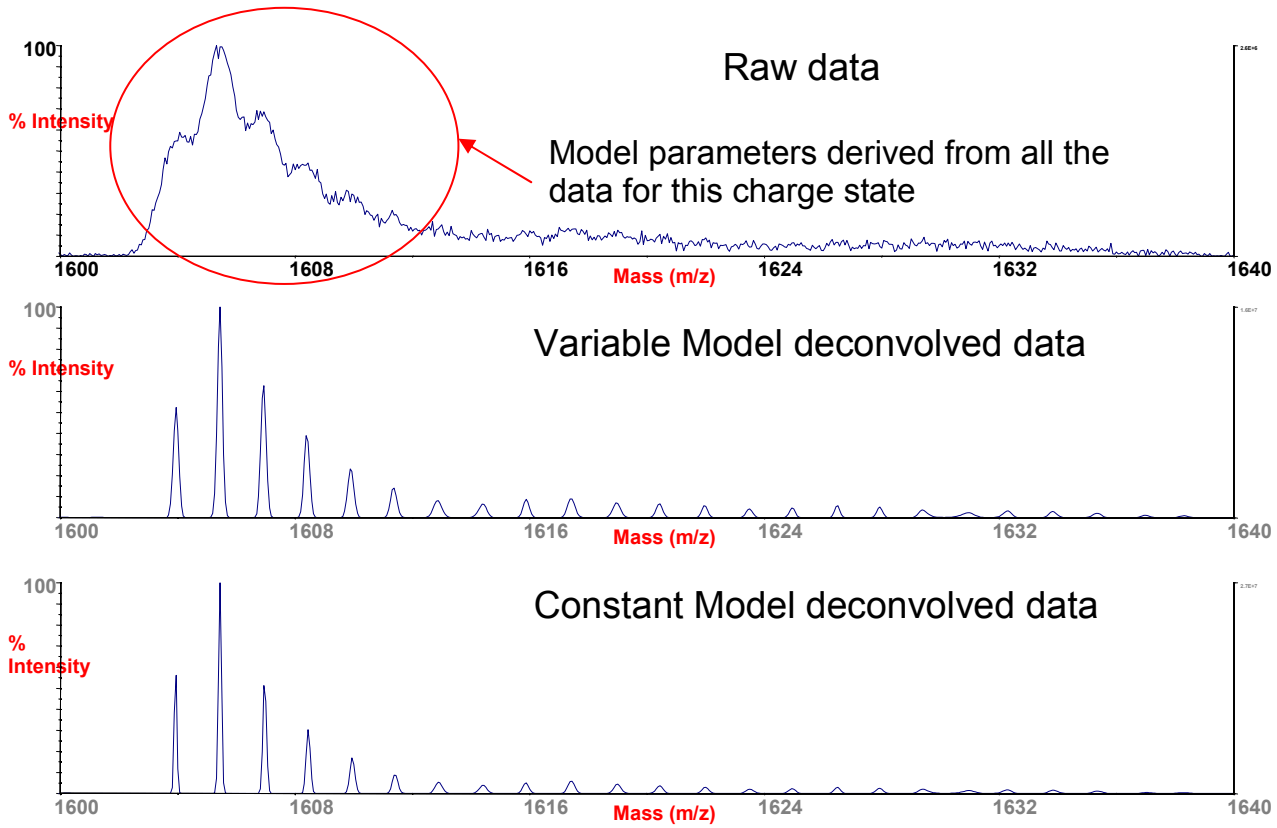


Figure 3: Raw data comparison with deconvolution spectrum using a constant and variable model.

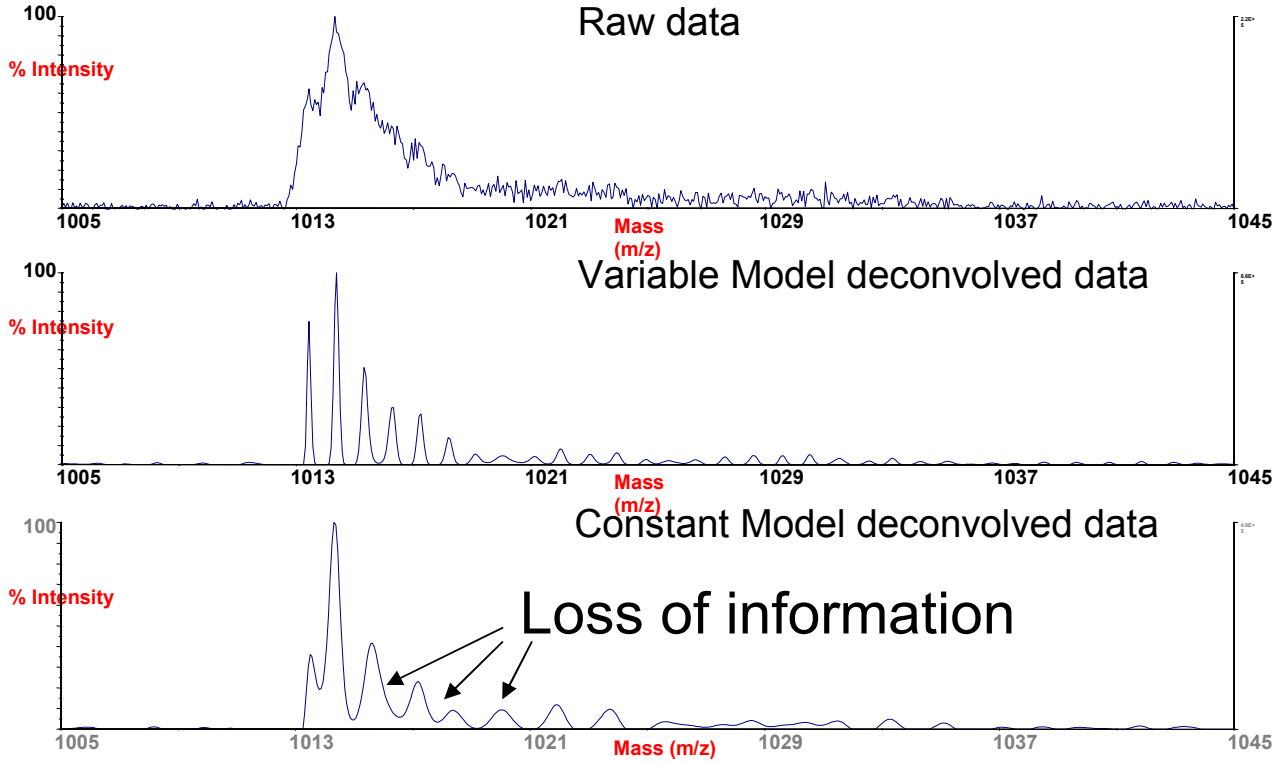


Figure 4: Comparison of Raw data to deconvolved spectra – at the low mass end (charge state z=19)

Figure 5 shows the data at the other end of the mass spectrum on the peaks at charge state z=19. Here the program is trying to fit the data to too narrow a width, with the result that the peak confidence is lower (broader deconvolved result) and some peaks appear to split (trying to find too much in the data).

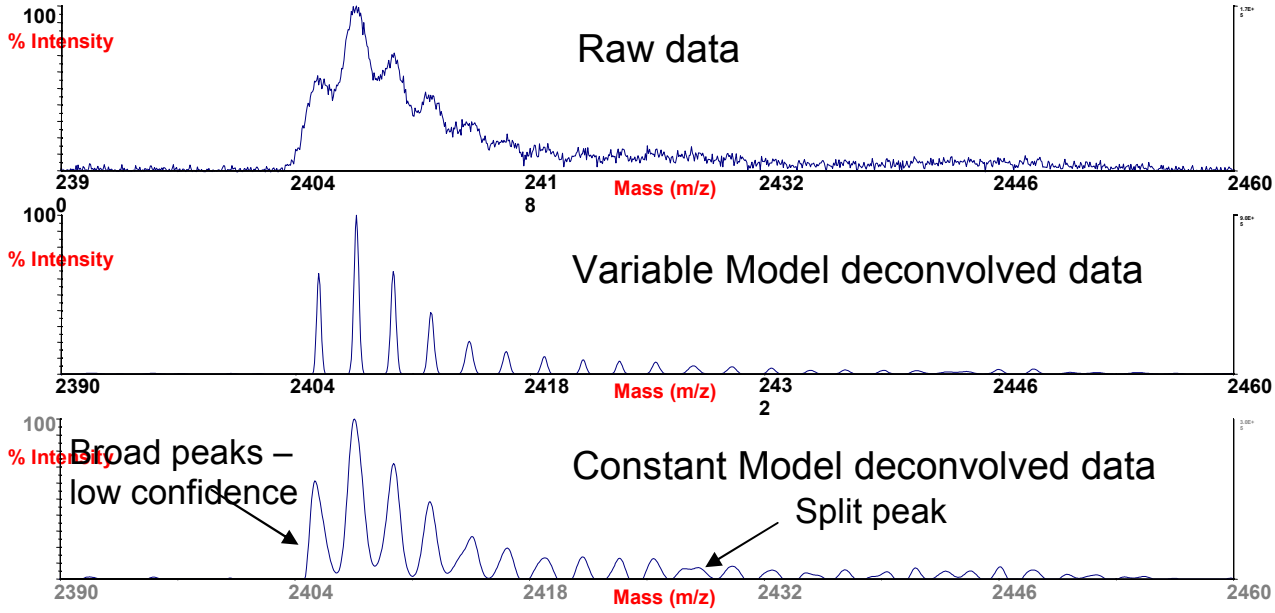


Figure 5: Comparison of Raw data to deconvolved spectra – at the high mass end (charge state z=8)

Figure 6 shows the charge deconvolved results on both the variable and constant model data. As mentioned because the data from using all the charge states provides misleading data at the ends of the charge state envelope, the data for only using 6 charge states is also presented for the constant model. As one can see, for like use of the number of charge states, the constant model yields increased peak errors and missing water molecule additions. Even with the reduced use of just 6 charge states, the peak errors are multiplied. The non use of the other data may affect interpretation of more complex data.

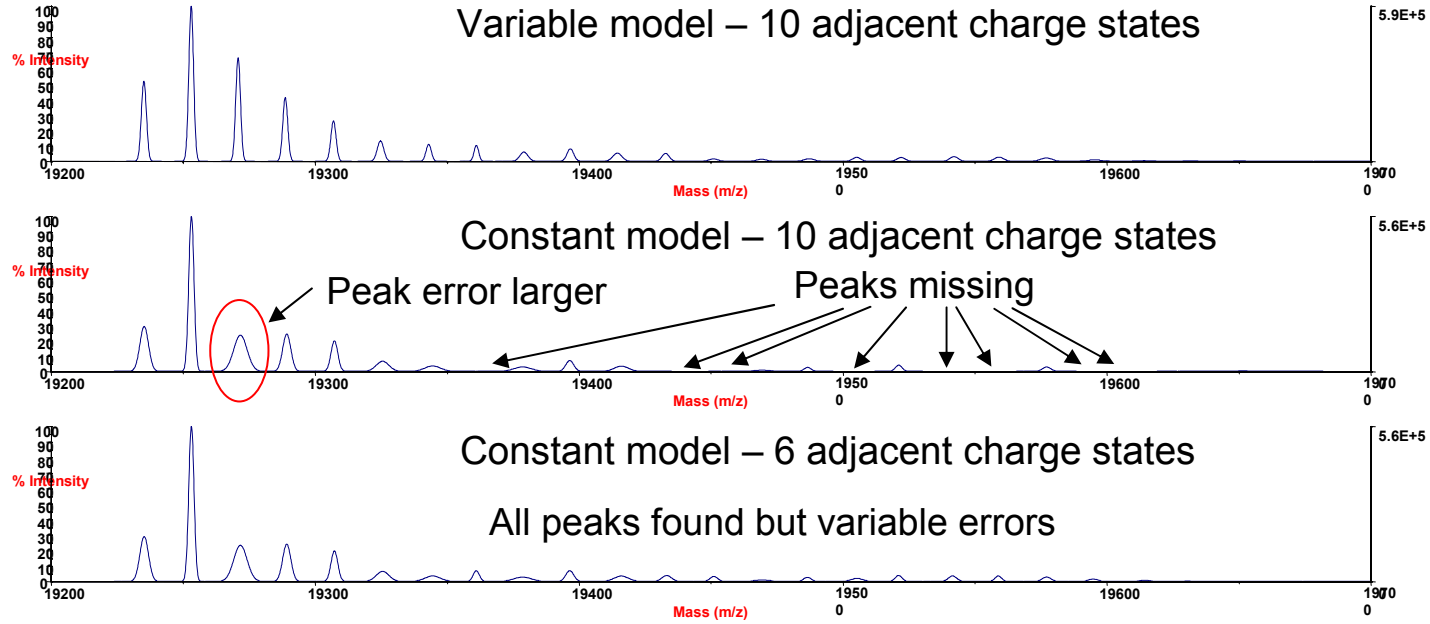
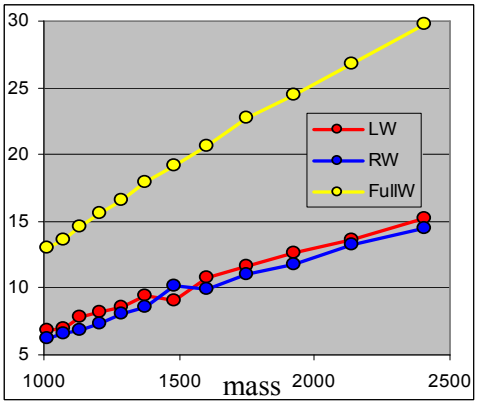
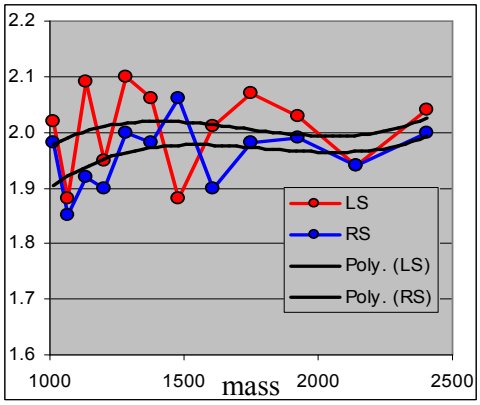


Figure 6: Charge deconvolved results – variable and constant models

Figure 7 shows how the variable model peak widths and shape change across the mass range. It should be of note that with the interferon data the major influence is on the peak width rather than the shape, the former changing by a factor of >2 whereas the peak shape alters by 10% and is insignificant. Note the peak model parameters used for the variable follow the black polynomial trend lines.



Peak width change
LW=Left width, RW=Right width, Full W=Full width



Peak Shape change
LS=Left shape, RS=Right shape

Figure 7: Variable Model – how the peak widths and shape varies across the mass range

Table 1 & 2 show the mass assignments of the peaks at each charge together with the mass differences between adjacent peaks (the water molecule addition). It is clearly seen that the variable model allows the correct identification of peaks even at the extremes of the charge state envelope, whereas the constant model provides weak data interpretation particularly at higher charge states. The Std deviation figure for the mass difference between adjacent peaks is much higher (x 2) using the constant model.

Table 1: Mass Accuracy data for Variable model after charge deconvolution

Interferon: Variable model. Preliminary model measured at all charges to demonstrate principle
Minimum adjacent charges = 10. Mass tolerance = 0.2

Peak	Mass	M Err	Intensity	Evidence (mass, charge...)												Adjacent Dif (Th)	Adjacent Dif (Fnd)	Th-Fnd	Summed Dif (Th)	Summed Dif (Fnd)	Th-Fnd
				19	18	17	16	15	14	13	12	11	10	9	8						
0	19235.1	0.9	143682695	1013.4	1069.7	1132.5	1203.3	1283.4	1375.0	1480.7	1603.9	1749.6	1924.4	2138.2	2405.4						
1	19253.1	0.9	261126509	1014.4	1070.7	1133.6	1204.4	1284.5	1376.2	1482.0	1605.4	1751.3	1926.2	2140.2	2407.6	18.01	18.00	**0.01**	18.01	18.00	**0.01**
2	19270.9	0.9	174659963	1015.3	1071.7	1134.6	1205.4	1285.7	1377.5	1483.4	1606.9	1752.9	1928.1	2142.2	2409.8	18.01	17.80	**0.21**	36.02	35.80	**0.22**
3	19288.7	0.9	112042730	1016.3	1072.7	1135.7	1206.6	1287.0	1378.8	1484.8	1608.4	1754.5	1929.9	2144.1	2412.1	18.01	17.80	**0.21**	54.03	53.60	**0.43**
4	19307.0	1.0	74642891	1017.2	1073.7	1136.7	1207.7	1288.2	1380.1	1486.2	1609.9	1756.2	1931.7	2146.1	2414.4	18.01	18.30	**-0.29**	72.04	71.90	**0.14**
5	19324.8	1.2	47249930	1018.2	1074.7	1137.8	1208.9	1289.4	1381.3	1487.6	1611.4	1757.8	1933.4	2148.2	2416.6	18.01	17.80	**0.21**	90.05	89.70	**0.35**
6	19343.0	1.0	31752868	1019.1	1075.7	1138.9	1209.9	1290.6	1382.6	1489.0	1612.9	1759.4	1935.2	2150.2	2418.9	18.01	18.20	**-0.19**	108.06	107.90	**0.16**
7	19361.1	0.9	26530450	1020.1	1076.6	1139.9	1211.1	1291.8	1383.9	1490.3	1614.4	1761.2	1937.1	2152.2	2421.2	18.01	18.10	**-0.09**	126.07	126.00	**0.07**
8	19379.1	1.5	26804952	1021.1	1077.8	1140.9	1212.3	1293.0	1385.3	1491.8	1615.9	1762.7	1938.8	2154.2	2423.3	18.01	18.00	**0.01**	144.08	144.00	**0.08**
9	19396.7	1.3	29997719	1022.0	1078.6	1142.0	1213.4	1294.2	1386.6	1493.1	1617.4	1764.4	1940.5	2156.2	2425.5	18.01	17.60	**0.41**	162.09	161.60	**0.49**
10	19414.5	1.7	26735025	1023.0	1079.7	1143.1	1214.6	1295.4	1387.8	1494.4	1618.9	1765.9	1942.3	2158.1	2427.7	18.01	17.80	**0.21**	180.11	179.40	**0.71**
11	19432.8	1.4	20665690	1023.9	1080.7	1144.1	1215.7	1296.5	1389.1	1495.9	1620.4	1767.6	1944.2	2160.1	2430.1	18.01	18.30	**-0.29**	198.12	197.70	**0.42**
12	19450.9	1.8	7701164	1024.9	1081.5	1145.1	1216.7	1297.8	1390.5	1497.3	1622.0	1769.3	1945.9	2162.1	2432.4	18.01	18.10	**-0.09**	216.13	215.80	**0.33**
13	19469.2	1.9	7282679	1025.7	1082.6	1146.1	1217.6	1299.1	1391.7	1498.5	1623.5	1770.8	1948.1	2164.2	2434.7	18.01	18.30	**-0.29**	234.14	234.10	**0.04**
14	19487.3	2.0	9286666	1026.6	1083.8	1147.2	1218.9	1300.2	1393.0	1500.4	1624.9	1772.4	1949.7	2166.2	2436.8	18.01	18.10	**-0.09**	252.15	252.20	**-0.05**
15	19505.2	1.5	11373427	1027.6	1084.8	1148.2	1220.1	1301.4	1394.2	1501.5	1626.5	1774.1	1951.5	2168.4	2439.1	18.01	17.90	**0.11**	270.16	270.10	**0.06**
16	19522.1	1.6	11200613	1028.6	1085.6	1149.3	1221.1	1302.6	1395.5	1502.9	1627.9	1775.6	1953.2	2170.2	2441.1	18.01	16.90	**1.11**	288.17	287.00	**1.17**
17	19542.1	1.6	14077415	1029.6	1086.8	1150.4	1222.3	1303.9	1396.8	1504.3	1629.4	1777.6	1955.1	2172.3	2444.1	18.01	20.00	**-1.99**	306.18	307.00	**-0.82**
18	19559.1	1.7	13872314	1030.5	1087.9	1151.5	1223.4	1305.0	1398.1	1505.5	1630.9	1779.1	1956.7	2174.2	2446.0	18.01	17.00	**1.01**	324.19	324.00	**0.19**
19	19577.1	2.2	13876455	1031.6	1088.9	1152.7	1224.4	1306.1	1399.4	1506.8	1632.3	1780.8	1958.6	2176.4	2448.1	18.01	18.00	**0.01**	342.20	342.00	**0.20**
20	19595.4	3.1	7527487	1032.6	1089.9	1154.1	1225.8	1307.2	1400.7	1508.3	1633.8	1782.3	1960.5	2178.2	2450.3	18.01	18.30	**-0.29**	360.21	360.30	**-0.09**
21	19614.2	3.0	4630885	1033.3	1091.0	1155.0	1227.0	1308.3	1402.0	1509.7	1635.3	1783.9	1962.4	2180.6		18.01	18.80	**-0.79**	378.22	379.10	**-0.88**
22	19631.6	2.8	2433966	1034.3	1092.0	1155.7	1228.1	1309.5	1403.3	1511.1	1637.0	1785.8	1963.9	2182.4		18.01	17.40	**0.61**	396.23	396.50	**-0.27**
23	19650.3	4.1	2437566		1093.2	1157.3	1229.2	1310.9	1404.5	1512.6	1638.3	1787.5	1965.6	2184.4		18.01	18.70	**-0.69**	414.24	415.20	**-0.96**
24	*19692.9*	*2.7*	*717224*	*1037.5*	*1095.1*	*1159.2*	*1231.9*	*1314.1*	*1407.7*	*1515.5*	*1641.9*	*1791.3*	*1970.6*								
																	Std Dev.	**0.62**			**0.48**

Table 2: Mass Accuracy data for Constant model after charge deconvolution

Interferon: Constant model. Preliminary model measured at m/z 1604
Minimum adjacent charges = 6. Mass tolerance = 0.2

Peak	Mass	M Err	Intensity	Evidence (mass, charge...)												Adjacent	Adjacent		Summed	Summed	
				19	18	17	16	15	14	13	12	11	10	9	8	Dif (Th)	Dif (Fnd)	**Th-Fnd**	Dif (Th)	Dif (Fnd)	**Th-Fnd**
0	19235.2	1.6	132069697	1013.6	1069.8	1132.5	1203.3	1283.4	1375.0	1480.7	1603.9	1749.6	1924.4	2138.2	2405.4						
1	19253.1	1.0	275086910	1014.3	1070.7	1133.6	1204.4	1284.6	1376.2	1482.1	1605.4	1751.3	1926.2	2140.2	2407.6	18.01	17.90	**0.11**	18.01	17.90	**0.11**
2	19271.7	2.5	165312521	1015.7	1071.8	1134.8	1205.6	1285.8	1377.6	1483.5	1607.0	1752.9	1928.0	2142.2	2409.8	18.01	18.60	**-0.59**	36.02	36.50	**-0.48**
3	19289.2	1.5	103806001		1072.6	1135.9	1206.7	1287.1	1378.9	1484.8	1608.4	1754.5	1929.9	2144.1	2412.1	18.01	17.50	**0.51**	54.03	54.00	**0.03**
4	19307.3	1.3	71075112	1017.2	1073.7	1136.8	1207.8	1288.3	1380.2	1486.2	1609.9	1756.2	1931.7	2146.1	2414.3	18.01	18.10	**-0.09**	72.04	72.10	**-0.06**
5	19325.5	2.4	44154846	1018.4	1074.9	1137.6	1209.0	1289.5	1381.4	1487.7	1611.4	1757.9	1933.4	2148.2	2416.6	18.01	18.20	**-0.19**	90.05	90.30	**-0.25**
6	19344.5	3.0	29631146		1076.0	1139.1	1210.5	1290.6	1382.7	1489.1	1612.9	1759.5	1935.6	2150.3	2418.8	18.01	19.00	**-0.99**	108.06	109.30	**-1.24**
7	19361.0	1.2	22921923					1291.6	1384.0	1490.4	1614.4	1761.2	1937.1	2152.2	2421.2	18.01	16.50	**1.51**	126.07	125.80	**0.27**
8	19378.6	3.3	26816398		1077.1	1140.6	1212.1	1293.0	1385.3	1491.7	1615.9	1762.7	1938.8	2154.5	2423.3	18.01	17.60	**0.41**	144.08	143.40	**0.68**
9	19396.4	1.5	29715445	1021.9	1078.5	1142.0	1213.3	1294.3	1386.7	1493.0	1617.4	1764.4	1940.5	2156.2	2425.4	18.01	17.80	**0.21**	162.09	161.20	**0.89**
10	19416.0	2.7	26652776		1079.9	1143.5	1214.7	1295.4	1388.0	1494.4	1619.0	1765.9	1942.4	2158.2	2428.1	18.01	19.60	**-1.59**	180.11	180.80	**-0.69**
11	19433.2	1.8	19779052				1215.8	1296.6	1389.2	1496.0	1620.4	1767.7	1944.2	2160.1	2430.1	18.01	17.20	**0.81**	198.12	198.00	**0.12**
12	19451.1	1.4	13330552					1297.8	1390.4	1497.4	1622.0	1769.3	1945.9	2162.1	2432.4	18.01	17.90	**0.11**	216.13	215.90	**0.23**
13	19469.4	2.9	8098178	1025.9	1082.7	1146.0	1217.5	1299.2	1391.7	1498.7	1623.5	1770.8	1948.1	2164.2	2434.8	18.01	18.30	**-0.29**	234.14	234.20	**-0.06**
14	19486.6	1.4	10728995		1083.7	1147.2	1218.9	1300.2	1393.0	1499.9	1624.9	1772.4	1949.7	2166.2	2436.8	18.01	17.20	**0.81**	252.15	251.40	**0.75**
15	19505.4	2.2	12450282				1220.4	1301.3	1394.2	1501.4	1626.5	1774.1	1951.5	2168.4	2439.1	18.01	18.80	**-0.79**	270.16	270.20	**-0.04**
16	19521.1	1.3	14157000		1085.1	1149.2	1220.4	1302.5	1395.4	1502.8	1627.9	1775.8	1953.2	2170.2	2441.1	18.01	15.70	**2.31**	288.17	285.90	**2.27**
17	19541.5	1.6	15764500				1222.2	1303.9	1396.7	1504.3	1629.5	1777.5	1955.1	2172.3		18.01	20.40	**-2.39**	306.18	306.30	**-0.12**
18	19558.9	1.3	12485336					1304.9	1398.1	1505.4	1630.9	1779.2	1956.8	2174.1	2446.0	18.01	17.40	**0.61**	324.19	323.70	**0.49**
19	19577.2	1.6	13424905	1031.4	1088.6	1152.8	1224.1	1306.1	1399.5	1506.8	1632.3	1780.8	1958.7	2176.4	2448.0	18.01	18.30	**-0.29**	342.20	342.00	**0.20**
20	19594.8	1.8	8117244				1225.8	1307.6	1400.8	1508.3	1633.8	1782.3	1960.5	2178.2	2450.1	18.01	17.60	**0.41**	360.21	359.60	**0.61**
21	19614.3	2.2	5150820						1402.2	1509.8	1635.4	1783.9	1962.4	2180.6		18.01	19.50	**-1.49**	378.22	379.10	**-0.88**
22	19630.6	2.2	2374908					1309.5	1403.2	1511.1	1637.1	1785.8	1964.0	2182.3	2454.6	18.01	16.30	**1.71**	396.23	395.40	**0.83**
23	19651.5	4.6	2477539		1093.2	1157.2	1229.8	1311.2	1404.4	1512.7	1638.3	1787.6	1965.8	2184.4		18.01	20.90	**-2.89**	414.24	416.30	**-2.06**
24	*19691.4*	*3.2*	*684545*			*1159.1*	*1231.8*	*1314.3*	*1407.6*	*1515.4*	*1641.9*	*1791.3*	*1970.6*								
																	Std Dev.	**1.24**			**0.84**

Part 2

Part two concerned the interpretation of a spectrum of a glycoprotein at mass 60,000. The protein was known to contain up to 4 sites of glycosylation, where each site has a core of 2 GlcNac and 3 Mannose additions. The protein was assessed to include multiple fucose groups and varying Hex and HexNac additions.

The raw electrospray data is shown in figure 8. The complexity of the sample is seen by the way the charge state groupings merge into one another.

The resulting deconvolution and charge deconvolution of the data by use of the constant and variable peak model techniques are shown in figure 9. As with the interferon data the variable model technique obtains more information.

Figures 10 and 11 show portions of the mass range corresponding to the charge state species for z=41 and 24 respectively. In each case the raw spectrum is shown at the top and the constant model deconvolved data in the middle and the variable model data at the bottom. With the z=41 data, the peak at 1469 is fitted to be one peak by the constant model method, whereas it becomes two peaks by the variable method. This is because the constant model is too broad to be able to discern the correct assessment of two peaks at this point. The z=24 data reveals the constant model to be too narrow causing peaks at m/z 2509 to be split into 3 rather than the correct two (as seen using the variable model).

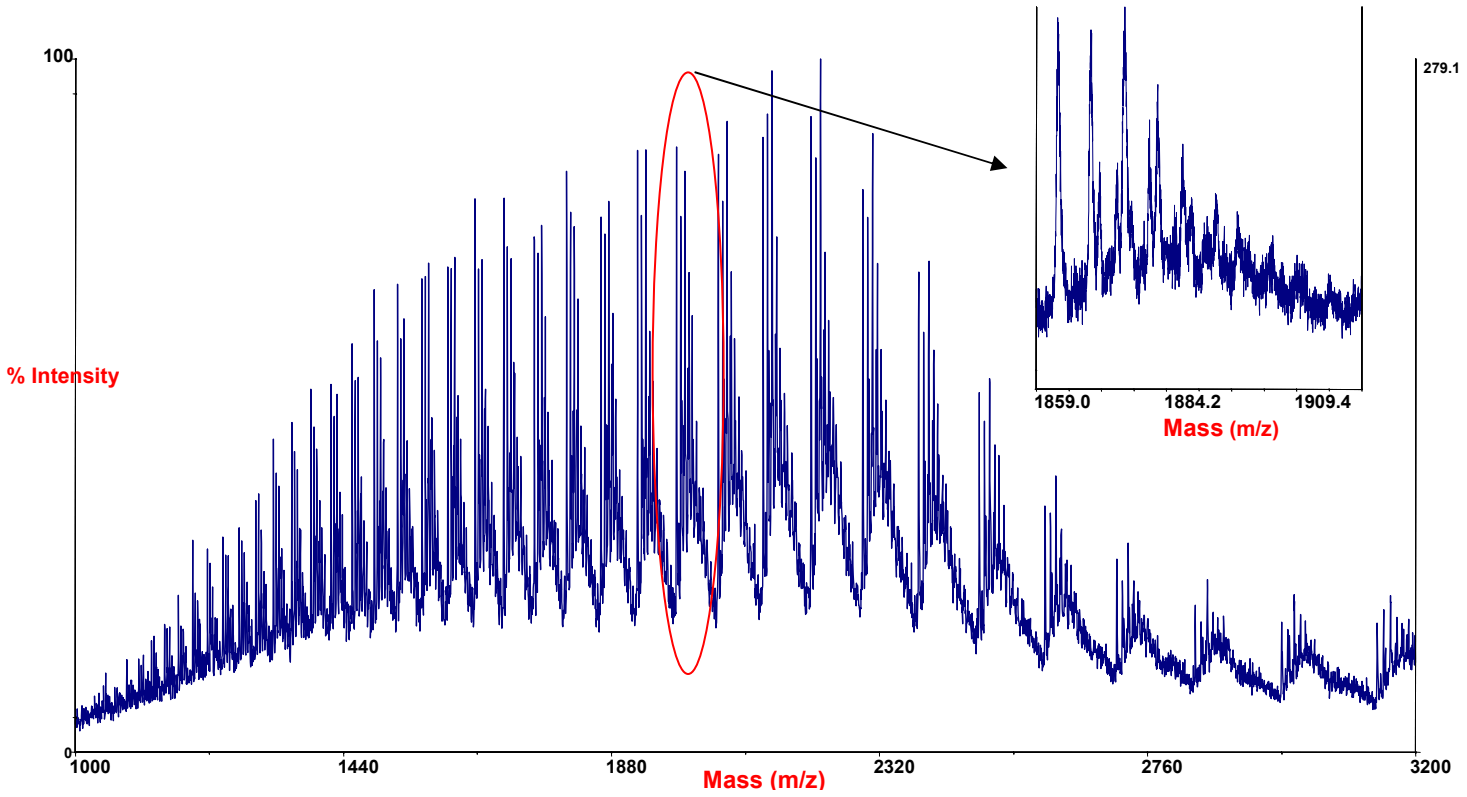


Figure 8: Electrospray MS of Glycoprotein B

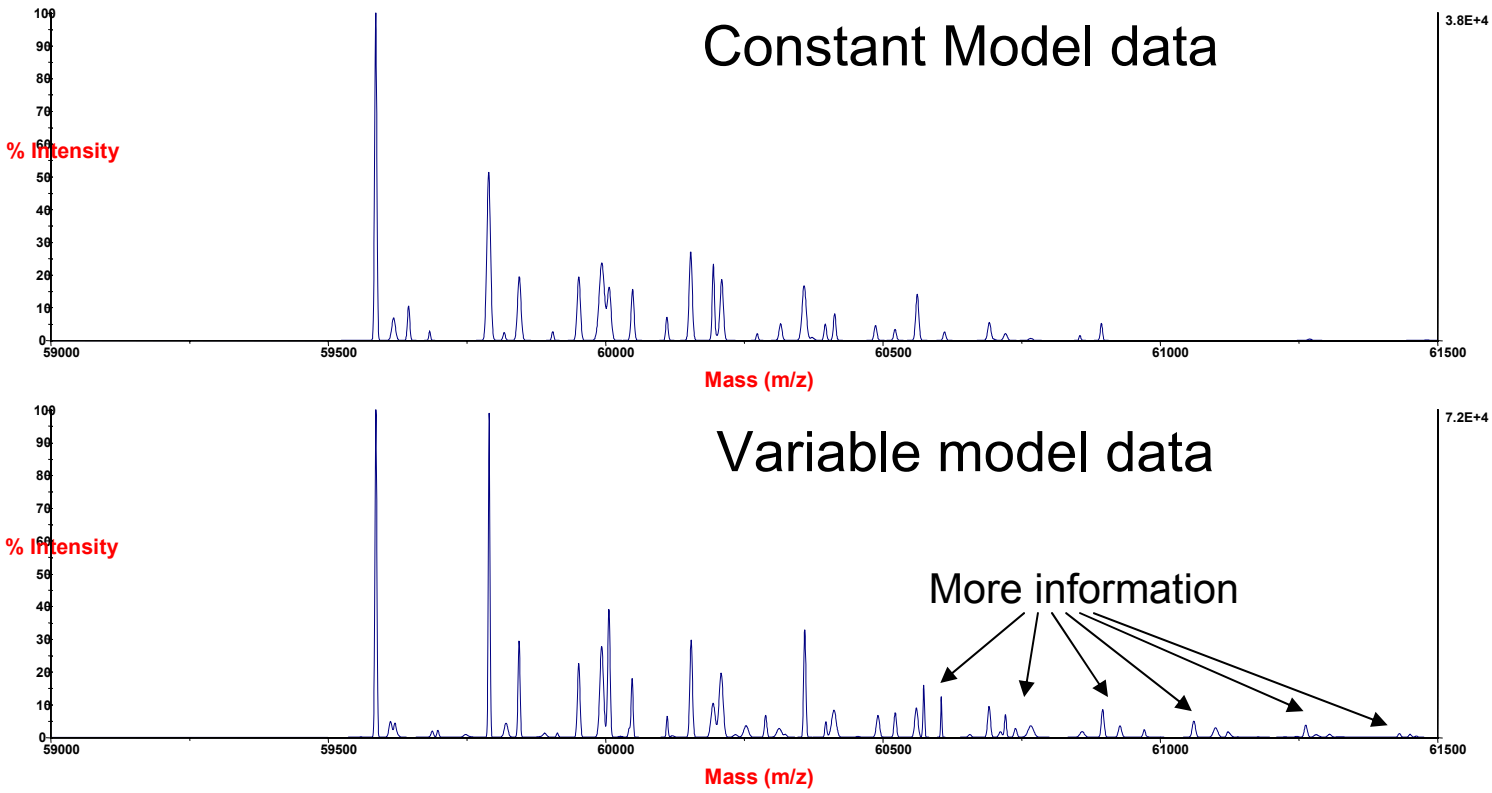


Figure 9: Charge deconvolved data for Glycoprotein B

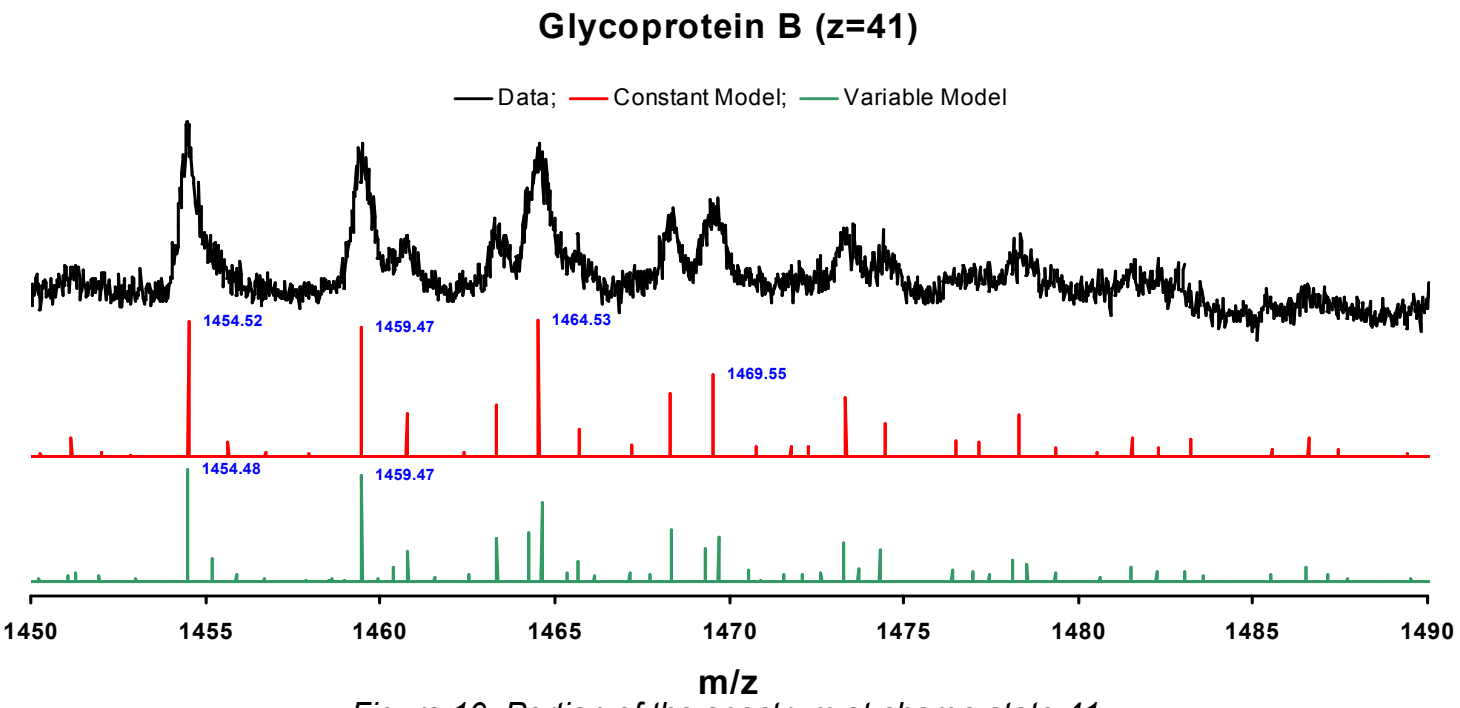


Figure 10: Portion of the spectrum at charge state 41.

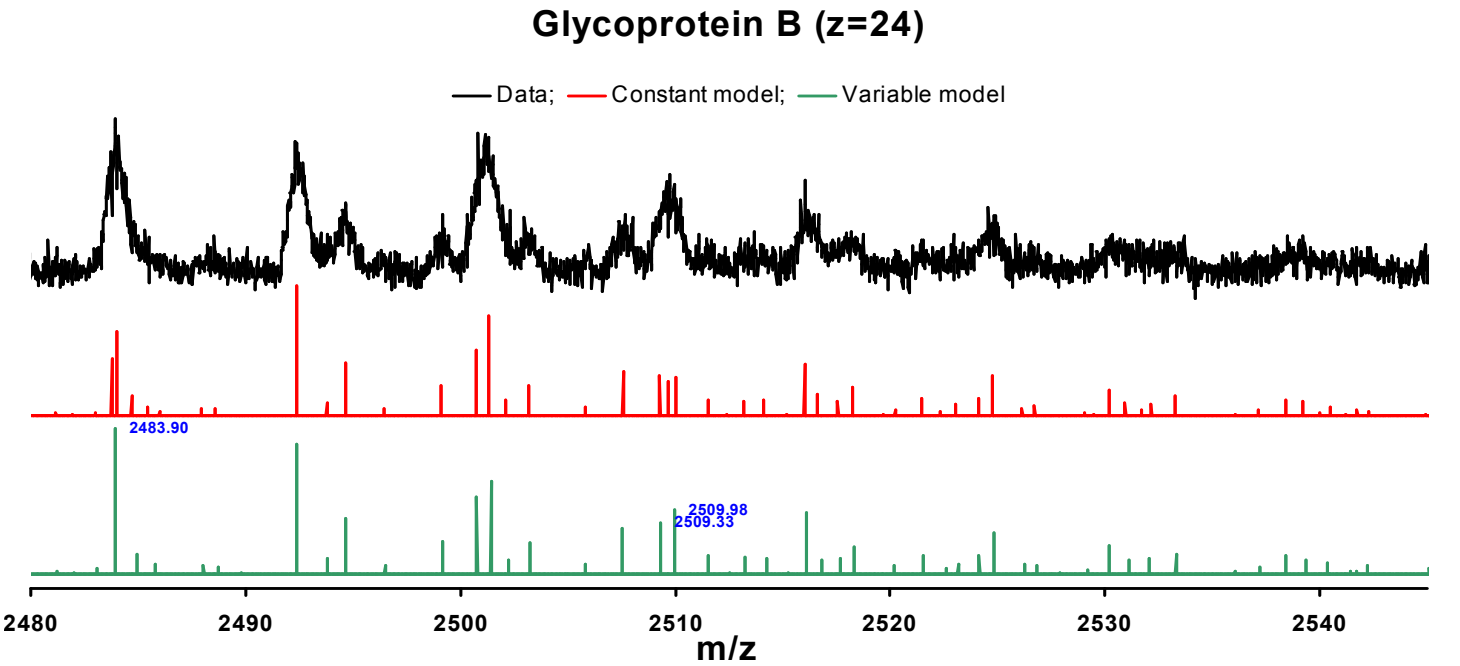


Figure 11: Portion of the spectrum at charge state 24.

Tables 3 and 4 show the mass assignments and calculated errors for each of the peaks for the constant and variable model respectively. Each peak is also assigned to a glycosidic combination. The calculated theoretical mass for each glycoform is calculated and compared with recalibrated found peaks. The recalibration allows any systematic calibration error on the mass spectrometer to be eliminated and allows comparison with the mass error for each peak as determined by the ReSpectTM algorithm. The data clearly show that the variable model technique allows for the identification of many more glycoforms with a resulting improvement in the average Std Deviation over the constant model data.

Table 3: Constant model peak identification & assignment from the charge deconvolved result

Mass Peak Table

Peak #	Found M	M Err	Intensity
1	59587.0	8.6	175
2	59585.9	2.2	82252
3	59618.1	3.8	9876
4	59645.1	2.4	9370
5	59683.1	1.8	1832
6	59789.4	3.7	71806
7	59817.3	2.5	2193
8	59844.6	3.6	26185
9	59904.9	2.1	2199
10	59951.8	3.3	24156
11	59993.3	5.5	48594
12	60006.6	3.9	23178
13	60048.7	3.0	17409
14	60110.7	2.3	6158
15	60153.6	3.2	32470
16	60194.3	2.3	18876
17	60197.9	4.8	2193
18	60204.4	5.6	776
19	60209.5	3.4	23328
20	60273.5	2.0	1607
21	60315.6	3.1	5816
22	60358.1	4.2	26613
23	60372.5	4.0	1349
24	60396.2	2.5	4629
25	60413.1	2.6	7801
26	60486.8	2.8	4873
27	60522.0	2.7	3375
28	60562.0	3.1	16559
29	60611.1	2.6	2599
30	60691.8	3.3	6589
31	60699.0	5.6	571
32	60721.0	3.5	2806
33	60766.6	3.6	861
34	60855.2	1.9	1041
35	60893.9	2.3	4465
36	61269.3	4.2	506
37	61480.6	7.2	340

Series Found

Series 1

Masses Calibrated Masses & Differences

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	3	0	0	2	59584.3	59585.9	2.2	82252		1.6
1	3	0	1	6	59787.5	59789.4	3.7	71806	203.6	2.0
1	3	0	2	11	59990.7	59993.3	5.5	48594	203.8	2.6
1	3	0	3	16	60193.9	60194.3	2.3	18876	201.0	0.4
1	3	0	4	24	60397.1	60396.2	2.5	4629	201.9	-0.9
1	3	0	5	29	60600.3	60611.1	2.6	2599	214.9	10.8
1	3	1	1	10	59949.6	59951.8	3.3	24156		2.1
1	3	1	2	15	60152.8	60153.6	3.2	32470	201.8	0.8
1	3	1	3	22	60356.0	60358.1	4.2	26613	204.5	2.0
1	3	1	4	28	60559.2	60562.0	3.1	16559	203.9	2.8
1	3	1	5	33	60762.4	60766.6	3.6	861	204.6	4.2
1	3	2	0	9	59908.6	59904.9	2.1	2199		-3.7
1	3	2	1	14	60111.8	60110.7	2.3	6158	205.8	-1.1
1	3	2	2	21	60315.0	60315.6	3.1	5816	204.9	0.6
1	3	2	3	27	60518.2	60522.0	2.7	3375	206.4	3.8
1	3	2	4	32	60721.4	60721.0	3.5	2806	199.0	-0.4

Series 2

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	2	0	2	8	59844.5	59844.6	3.6	26185		0.1
1	2	0	3	13	60047.7	60048.7	3.0	17409	204.1	1.0
1	2	1	2	12	60006.7	60006.6	3.9	23178		-0.1
1	2	1	3	19	60209.9	60209.5	3.4	23328	202.9	-0.4
1	2	1	4	25	60413.1	60413.1	2.6	7801	203.6	0.0
1	2	4	2	26	60493.1	60486.8	2.8	4873		-6.3
1	2	4	4	35	60899.5	60893.9	2.3	4465	407.1	-5.6
1	2	5	3	34	60858.4	60855.2	1.9	1041		-3.2
1	2	5	5	36	61264.8	61269.3	4.2	506	414.0	4.5

Series 3

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	1	2	1	7	59819.5	59817.3	2.5	2193		-2.1

Series 4

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	0	2	6	30	60689.3	60691.8	3.3	6589		2.5

Series with less than 4 components to core

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	3	-1	1	3	59625.3	59618.1	3.8	9876		-7.3
1	2	-1	2	5	59682.4	59683.1	1.8	1832		0.7
1	2	2	-1	1	59559.2	59567.0	8.6	175		7.8

Glycoprotein B Constant Peak Model

Colour Key

Identified peaks

Mass Constants

55574.60	Protein
3571.26	4xCore
146.14	Fuc
162.14	Hex
203.20	HexNAc

Analysis Summary

Error SD for Identified Peaks	3.8
Intensity of Interpreted Peaks	479214
Intensity of all Peaks	495420
% Interpreted	97
Intensity of Interpreted Peaks for a Variable Model	1028444
% Intensity Recovered: Constant/Variable Model	47

Table 4: Variable model peak identification & assignment from the charge deconvolved spectrum

Mass Peak Table

Peak #	Found M	M Err	Intensity
1	59589.1	4.3	372
2	59586.2	1.9	137621
3	59612.5	3.1	11136
4	59620.6	2.6	8047
5	59626.8	2.8	1461
6	59681.6	4.5	218
7	59687.9	2.6	3553
8	59697.9	2.0	31172
9	59727.6	7.0	431
10	59747.8	4.2	2646
11	59756.9	4.8	588
12	59790.3	1.7	121767
13	59820.8	3.8	11095
14	59844.2	2.1	44589
15	59882.2	7.0	988
16	59890.5	3.7	3163
17	59913.2	2.1	1942
18	59951.7	2.6	42359
19	59980.4	11.6	807
20	59993.1	3.8	76471
21	60006.1	2.5	70560
22	60026.9	4.0	903
23	60042.6	1.9	3394
24	60047.9	2.0	26028
25	60111.2	1.7	7870
26	60120.6	3.8	1290
27	60154.4	2.7	58166
28	60194.0	4.3	32272
29	60208.4	4.2	59476
30	60234.2	4.1	2501
31	60253.3	4.8	11531
32	60255.4	4.3	750
33	60288.6	2.4	11782
34	60313.0	5.0	9885
35	60324.3	2.9	1781
36	60389.2	2.2	62373
37	60397.1	2.0	6740
38	60407.5	6.2	2646
39	60411.9	4.6	26027
40	60455.1	3.9	708
41	60491.1	3.3	16059
42	60522.3	2.5	13701
43	60539.9	5.5	390
44	60560.4	3.3	21601
45	60573.7	1.6	18450
46	60606.3	1.2	10718
47	60656.7	3.1	1946
48	60691.4	2.9	19961
49	607 11.9	3.3	4258
50	607 20.9	2.0	10019
51	607 39.1	2.8	6631
52	607 53.3	5.8	223
53	607 66.7	5.8	14656
54	608 59.2	4.6	5931
55	608 96.3	2.9	17645
56	609 27.5	3.3	8354
57	609 42.3	7.3	311
58	609 71.2	2.3	3676
59	609 73.9	6.2	775
60	610 60.6	3.3	11918
61	610 99.3	4.4	8261
62	611 02.6	4.1	1515
63	611 21.3	2.2	1613
64	611 24.5	3.9	3121
65	611 39.4	4.4	344
66	611 76.7	3.9	344
67	612 24.8	2.9	429
68	612 46.3	3.7	731
69	612 62.4	3.1	6451
70	612 80.3	3.8	2061
71	612 85.7	4.0	762
72	613 05.2	3.6	2498
73	613 24.4	9.2	1095
74	613 96.4	6.3	326
75	614 31.1	2.7	2366
76	614 80.5	2.7	1841
77	614 61.6	2.6	716

Series Found

Series 1

Masses & Mass Differences

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	3	0	0	2	59584.3	59585.4	1.9	137621		1.1
1	3	0	1	12	59787.5	59789.5	1.7	121767	204.1	2.0
1	3	0	2	20	59990.7	59992.3	3.8	76474	202.8	1.6
1	3	0	3	28	60193.9	60193.2	4.3	32272	200.9	-0.7
1	3	0	4	37	60397.1	60396.4	2.0	6740	203.2	-0.7
1	3	0	5	46	60600.3	606048	1.2	10718	200.2	4.3
1	3	1	0	10	59745.4	59747.0	4.2	2645		0.6
1	3	1	1	18	59949.6	59950.9	2.5	42360	203.9	1.3
1	3	1	2	27	60152.8	601536	2.7	58165	202.7	0.8
1	3	1	3	36	60356.0	60358.5	2.2	62373	204.9	2.5
1	3	1	4	44	60559.2	60559.7	3.3	21607	201.2	0.4
1	3	1	5	53	60762.4	607655	5.8	14656	206.3	3.5
1	3	1	6	58	60965.6	60970.4	2.3	3676	204.5	4.8
1	3	2	0	17	59908.6	59912.4	2.1	1942		3.8
1	3	2	1	25	60111.8	601104	4.5	7870	198.0	-1.4
1	3	2	2	34	60315.0	603123	5.0	9885	201.9	-2.7
1	3	2	3	42	60518.2	605216	2.5	13704	209.3	3.4
1	3	2	4	50	60721.4	607202	2.9	10019	198.6	-1.2

Series 2

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	2	0	2	14	59844.5	59843.4	2.1	44589		-1.1
1	2	0	3	24	60047.7	60047.1	2.5	70560	203.7	-0.6
1	2	0	4	31	60250.9	602525	4.2	59476	205.4	1.6
1	2	0	5	40	60454.1	604544	3.9	708	201.9	0.2
1	2	1	2	21	60006.7	60005.3	2.5	70560		-1.4
1	2	1	3	29	60209.9	602078	4.2	59476	202.3	-2.3
1	2	1	4	39	60413.1	604112	4.8	26271	203.6	-1.9
1	2	2	4	45	60575.2	605730	1.8	18460	161.8	-2.3
1	2	3	4	51	60737.4	607393	2.9	5631	166.4	1.9
1	2	3	5	57	60940.6	609415	7.3	311	203.2	1.0
1	2	4	1	33	60289.9	602879	2.4	11782		-2.0
1	2	4	2	41	60493.1	60490.4	3.3	16059	202.5	-2.7
1	2	4	4	65	60899.5	608955	2.9	17645	405.2	-4.0
1	2	4	5	62	61102.7	611018	4.1	1515	206.3	-0.9
1	2	4	6	72	61305.9	613044	3.5	2498	202.6	-1.5
1	2	5	1	40	60462.1	604644	3.9	708		2.3
1	2	5	2	47	60665.3	606660	3.1	1945	201.6	0.7
1	2	5	3	54	60868.4	60868.4	4.5	5937	202.5	0.0
1	2	5	4	60	61061.6	610598	3.3	11918	201.4	-1.8
1	2	5	5	69	61264.8	612616	3.1	8454	201.8	-3.2
1	2	6	4	67	61223.9	612240	2.9	429		0.2
1	2	6	5	75	61427.0	614303	2.7	2366	206.3	3.3

Series 3

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	1	0	2	8	59598.4	59597.1	2.9	3112		-1.3
1	1	2	1	13	59819.5	59820.0	3.8	11895		0.5
1	1	3	1	19	59981.6	59979.6	11.8	507	199.6	-2.0
1	1	4	4	62	60753.4	607525	5.8	223	772.9	-0.8
1	1	5	5	63	61118.7	611205	2.2	1613	388.0	1.8
1	1	5	6	73	61321.9	613236	9.2	1095	203.1	1.7
1	1	6	5	70	61280.8	612795	3.8	2061		-1.3

Series 4

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	0	0	3	11	59755.4	59756.1	4.8	588		0.7
1	0	1	4	26	60120.8	601198	3.8	1290	363.7	-1.0
1	0	1	5	35	60324.0	603238	2.9	1781	203.8	-0.4
1	0	2	6	48	60689.3	606907	2.9	19961	367.1	1.3
1	0	2	8	61	61095.7	610995	4.4	9261	407.9	2.9
1	0	1	9	65	61136.8	611386	4.4	344		1.9

Series with less than 4 components to core

Core	Fuc	Hex	HexNAc	Peak #	Pred M	Calibrated M	Err	Intensity	M Diff	Cal-Pred
1	3	-1	1	5	59625.3	596250	2.8	1461		-0.3
1	2	-1	2	9	59682.4	596808	4.5	218		-1.6
1	2	-1	3	16	59885.6	598914	7.0	988	200.6	-4.2
1	3	-1	4	30	60234.9	602334	4.1	2504	352.0	-1.5
1	2	2	-1	1	59559.2	595683	4.3	372		-0.9
1	3	4	-1	22	60029.7	600261	4.0	903	457.8	-3.6

Glycoprotein B Variable Peak Model

Colour Key

Identified peaks

Peaks common with Constant Model

Mass Constants

55574.60	Protein
3571.26	4xCore
146.14	Fuc
162.14	Hex
203.20	HexNAc

Analysis Summary

Error SD for Identified Peaks	2.1
Error SD for Constant Model Peaks	2.2
Intensity of Interpreted Peaks	1028444
Intensity of all Peaks in Peak Table	1082624
% Interpreted Intensity	95
Intensity Interpreted Peaks using a Constant Model	456935
% Intensity Recovered: Variable/Constant Model	225

Conclusion

The use of a variable model to deconvolve the spectra result in the following advantages:

1. Improved mass assignments on found peaks
2. Improved detailed information content on the spectra, particularly on weak components - more interpretable spectra.
3. Reduction on the number of artefact peaks that occur as a result of over or under fitting data.

Future work will include automating the ability to assess the variable model.