Waters™

应用纪要

The Use of A Nanolockspray Electrospray Interface for Exact Mass Proteomics Studies

Alan L. Millar, James I. Langridge, Therese McKenna, Steven Pringle, Phil Young, Bateman, R., Kevin Giles

Waters Corporation

For research use only. Not for use in diagnostic procedures.

Abstract

- This application note describes the implementation of a dual sprayer NanoLockSpray source on the Waters Micromass Q-Tof micro Mass Spectrometer
- This source consists of a dual sprayer arrangement; one for the nanolitre/min flow rate from the analyte, while the second is for the reference compound
- This facilitates routine exact mass measurement for nano scale LC-MS and MS/MS studies

Benefits

- The NanoLockSpray source described here provides a routine method for obtaining enhanced mass measurements in LC-MS and MS/MS experiments

- Mass measurement using the NanoLockSpray source over a 16-hour period resulted in values ±5 ppm
- This interface provides unique functionality in a proteomics experiment, where exact mass measurement reduces false positive protein identifications and provides added confidence in those proteins identified
- Use of the NanoLockSpray source reduced the false positive identification of proteins even in the case of a BSA standard
- Identification of proteins from both standard digests and *E.coli* samples resulted in mass measurements better then 10 ppm
- *De novo* sequencing of peptides with ProteinLynx Global SERVER v2.0 is significantly enhanced through the specificity of exact mass MS/MS

Introduction

Due to the complexity encountered in the analysis of proteins obtained from mammalian systems, the primary route for the identification and characterization of the constituent proteins is electrospray (ESI) LC-MS/MS. The low endogenous levels and large dynamic range of proteins present in these samples dictate that nanoscale LC-MS/MS is often the method of choice, due to the concentration dependent nature of the ESI technique. This has led to nanoscale LC-MS/MS on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-Tof) mass spectrometer as an established technique for high sensitivity identification and characterization of proteins. Typically, these experiments employ HPLC columns that have an internal diameter of 75 µm, or less, operating at flow rates of approximately 200 nL/min. While this set-up offers the optimum sensitivity it does not allow the post-column addition of an internal reference ion, as this would detrimentally effect the resolution of the HPLC separation, resulting in peak broadening. The use of an internal reference is required to provide reliable high mass measurement accuracy. Here we report the use of a nanoflow LockSpray interface to routinely provide enhanced mass measurement in the analysis of protein digests.

Experimental

Electrospray LC-MS/MS

- Data were acquired using a Q-Tof micro, hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Figure 1).
- The analytical system used for the analysis consisted of a Waters Micromass CapLC with a ten-port valve, the stream select module, attached directly to the ZSpray source of the mass spectrometer.
- The CapLC was configured with a pre-concentration column, to allow large volume sample injection and a nanoscale analytical column. The trapping column was packed with Symmetry C_{18} stationary phase (300 µm ID x 5 mm), while the analytical column was a 150 mm x 75 µm column packed with C_{18} material.
- A splitter was employed to provide a flow through the analytical column of 200 nL/min with the pump programmed to deliver a flow of 2 µL/min. The LC gradient ran from 5% to 60% acetonitrile in 26 minutes.

Figure 1. Waters Micromass System for protein characterization.

NanoLockSpray

- The NanoLockSpray interface consists of a dual sprayer, one for the analyte and one for the reference, see Figure 2. Each spray is sampled individually by the mass spectrometer by means of an electronically controlled baffle plate.
- The software was configured such that the reference spray was sampled for one in every ten seconds. It is advantageous to separate the reference signal from the analyte signal in order to avoid ion suppression effects and possible interference to the lockmass from ions closely related in *m/z*. The reference ion used was the doubly charged ion of [Glu1]-Fibrinopeptide B at *m/z* 785.8426.
- The mass spectrometer was operated in a Data Directed Analysis (DDA) mode whereby following the interrogation of MS data, ions were selected for MS/MS analysis based on their intensity and charge state. Collision energies were chosen automatically based on the *m/z* and charge state of the selected precursor ions.

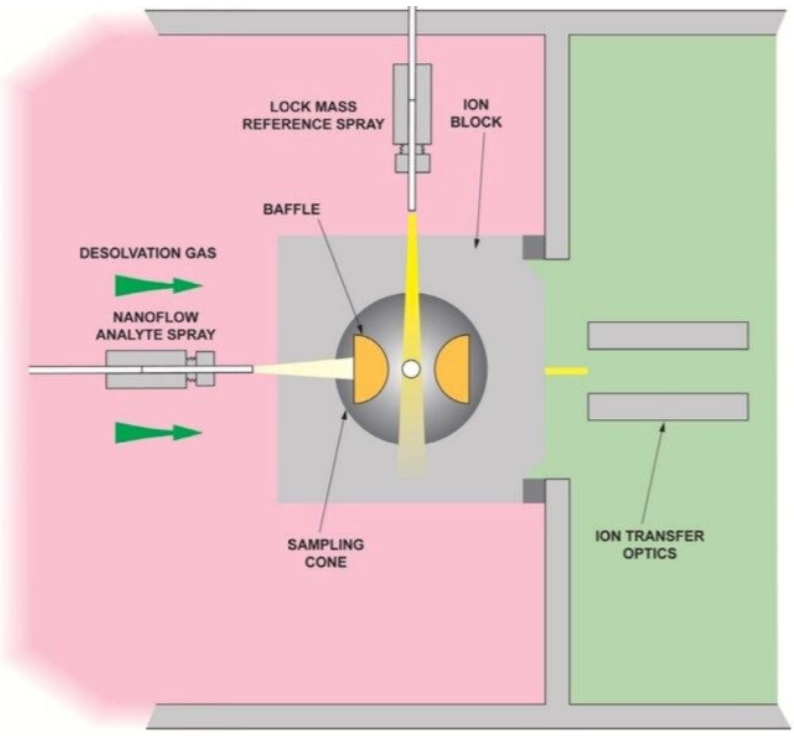


Figure 2. The NanoLockSpray source, consisting of dual sprayers for both analyte and reference spray.

Data Processing

- All data were processed using ProteinLynx Global SERVER v2.0. The processing consisted of automatically correcting the *m/z* scale of both the MS and MS/MS data utilizing the reference ion.
- The MS/MS data was also MaxEnt III processed. The purpose of the MaxEnt processing was to simplify the MS/MS spectra by deconvoluting the data to the single charge state and de-isotoping the data. After processing the data was searched against SWISS-PROT v40.

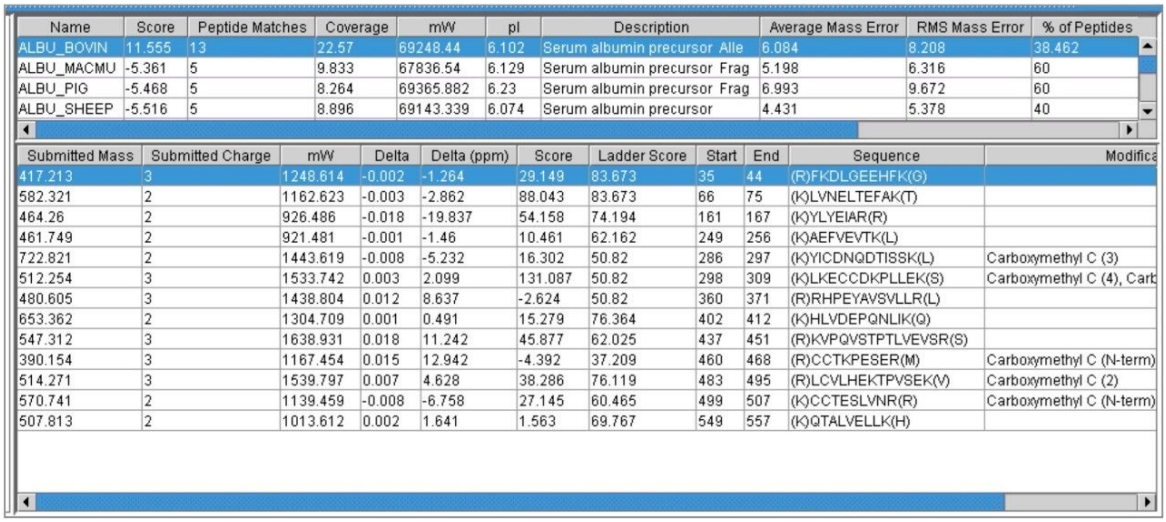
Results and Discussion

Initial results were obtained from the LC-MS/MS analysis of 500 fmoles of a tryptic digest of Bovine

Serum Albumin. Data were acquired automatically from a Q-Tof micro fitted with NanoLockSpray, using DDA.

Mass measurement accuracy and databank searching

Results from this analysis are presented in Figure 3, where the databank search was conducted with a precursor ion tolerance of 20 ppm. In this case 13 matching peptides were identified to the BSA sequence. Several other serum albumin proteins from different species were also identified. The RMS error for the peptides matching to the BSA sequence was 8 ppm.



Name	Score	Peptide Matches	Coverage	mW	pI	Description	Average Mass Error	RMS Mass Error	% of Peptides
ALBU_BOVIN	11.555	13	22.57	69248.44	6.102	Serum albumin precursor Alle	6.084	8.208	38.462
ALBU_MACMU	-5.361	5	9.833	67836.54	6.129	Serum albumin precursor Frag	5.198	6.316	60
ALBU_PIG	-5.468	5	8.264	69365.882	6.23	Serum albumin precursor Frag	6.993	9.672	60
ALBU_SHEEP	-5.516	5	8.896	69143.339	6.074	Serum albumin precursor	4.431	5.378	40

Submitted Mass	Submitted Charge	mW	Delta	Delta (ppm)	Score	Ladder Score	Start	End	Sequence	Modifica
417.213	3	1248.614	-0.002	-1.264	29.149	83.673	35	44	(R)FKDLGEEHFK(G)	
582.321	2	1162.623	-0.003	-2.862	88.043	83.673	66	75	(K)LVNELTEFAK(T)	
464.26	2	926.486	-0.018	-19.837	54.158	74.194	161	167	(K)YLYEIAR(R)	
461.749	2	921.481	-0.001	-1.46	10.461	62.162	249	256	(K)AEFVEVTK(L)	
722.821	2	1443.619	-0.008	-5.232	16.302	50.82	286	297	(K)YICDNQDTISSK(L)	Carboxymethyl C (3)
512.254	3	1533.742	0.003	2.099	131.087	50.82	298	309	(K)LKECCDKPLLEK(S)	Carboxymethyl C (4), Carb
480.605	3	1438.804	0.012	8.637	-2.624	50.82	360	371	(R)RHPEYAVSVLLR(L)	
653.362	2	1304.709	0.001	0.491	15.279	76.364	402	412	(K)HLVDEPQNLIK(Q)	
547.312	3	1638.931	0.018	11.242	45.877	62.025	437	451	(R)KVPQVSTPTLVEVSR(S)	
390.154	3	1167.454	0.015	12.942	-4.392	37.209	460	468	(R)CCTKPESER(M)	Carboxymethyl C (N-term)
514.271	3	1539.797	0.007	4.628	38.286	76.119	483	495	(R)LCVLHEKTPVSEK(V)	Carboxymethyl C (2)
570.741	2	1139.459	-0.008	-6.758	27.145	60.465	499	507	(K)CCTESLVNR(R)	Carboxymethyl C (N-term)
507.813	2	1013.612	0.002	1.641	1.563	69.767	549	557	(K)QTALVELLK(H)	

Figure 3. Databank search results for the NanoLockSpray BSA sample, with a 20 ppm precursor ion tolerance used in the search.

Analysis of the same data set using a 100 ppm window resulted in the identification of BSA as the top hit, with again 13 matching peptides, however an incorrect identification Ribonuclease R was also returned (Figure 4).

Name	Score	Peptide Matches	Coverage	mW	pI	Description	Average Mass Error	RMS Mass Error	% of Peptides Modifie
ALBU_BOVIN	11.555	13	22.57	69248.44	6.102	Serum albumin precursor Alle	6.084	8.208	38.462
RNR_BACSU	3.565	1	1.155	88694.584	6.03	Ribonuclease R EC 3 1	35.747	35.747	0

Submitted Mass	Submitted Charge	mW	Delta	Delta (ppm)	Score	Ladder Score	Start	End	Sequence	Modificatior
417.213	3	1248.614	-0.002	-1.264	29.149	83.673	35	44	(R)FKDLGEEHFK(G)	
582.321	2	1162.623	-0.003	-2.862	88.043	83.673	66	75	(K)LVNELTEFAK(T)	
464.26	2	926.486	-0.018	-19.837	54.158	74.194	161	167	(K)YLYEIAR(R)	
461.749	2	921.481	-0.001	-1.46	10.461	62.162	249	256	(K)AEFVEVTK(L)	
722.821	2	1443.619	-0.008	-5.232	16.302	50.82	286	297	(K)YICDNQDTISSK(L)	Carboxymethyl C (3)
512.254	3	1533.742	0.003	2.099	131.087	50.82	298	309	(K)LKECCDKPLLEK(S)	Carboxymethyl C (4), Carboxy
480.605	3	1438.804	0.012	8.637	-2.624	50.82	360	371	(R)RHPEYAVSVLLR(L)	
653.362	2	1304.709	0.001	0.491	15.279	76.364	402	412	(K)HLVDEPQNLIK(Q)	
547.312	3	1638.931	0.018	11.242	45.877	62.025	437	451	(R)KVPQVSTPTLVEVSR(S)	
390.154	3	1167.454	0.015	12.942	-4.392	37.209	460	468	(R)CCTKPESER(M)	Carboxymethyl C (N-term), Ca
514.271	3	1539.797	0.007	4.628	38.286	76.119	483	495	(R)LCVLHEKTPVSEK(V)	Carboxymethyl C (2)
570.741	2	1139.459	-0.008	-6.758	27.145	60.465	499	507	(K)CCTESLVNR(R)	Carboxymethyl C (N-term), Ca
507.813	2	1013.612	0.002	1.641	1.563	69.767	549	557	(K)QTALVELLK(H)	
509.296	2	1016.54	-0.036	-35.747	9.259	34.884	766	774	(R)GFQNAPKQK(R)	

Figure 4. Databank search results for the NanoLockSpray BSA sample, with a 100 ppm precursor ion tolerance used in the search..

Despite the poor quality of this identification, it would require manual verification to remove it from the protein hit list.

Figure 5 shows the results from a DDA analysis of an *Escheria coli* (*E. coli*) sample containing multiple proteins. In this case the SWISS-PROT databank was searched with the peptide precursor ion tolerance set to 20 ppm.

All of the retrieved proteins originated from *E. coli*. Identification of 17 proteins was made. The peptides matching the parent proteins had an RMS error of 10 ppm or better.

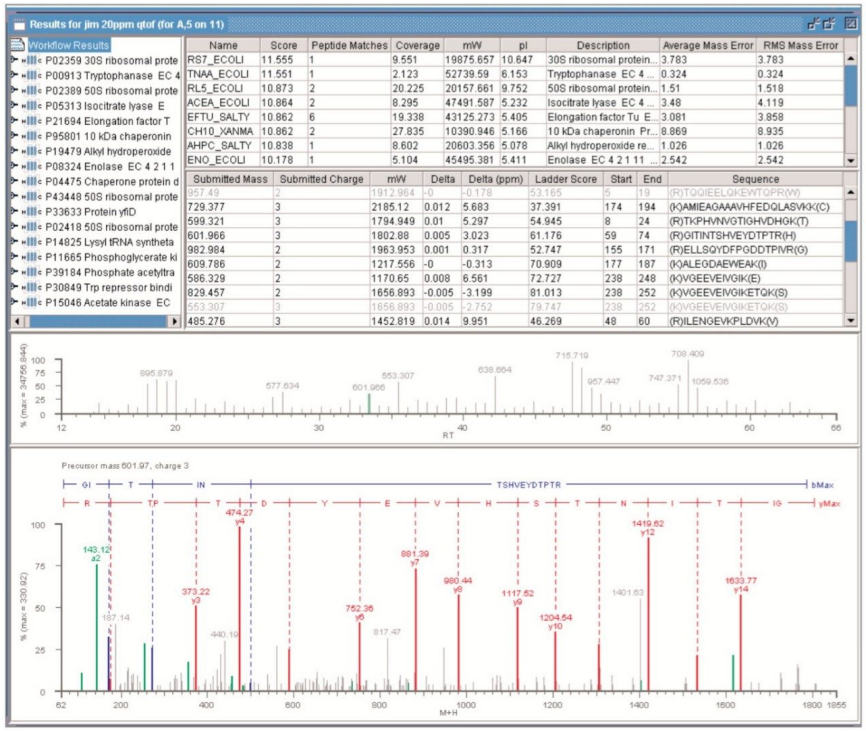


Figure 5. Databank search results obtained from the E.coli protein mixture.

Mass Measurement Stability

Stability of mass measurement obtained with the NanoLockSpray source, over an extended period of time, was investigated through the infusion of two peptides, [Glu’]-fibrinopeptide B (*m/z* 785.8426 2+) and Angiotensin II (*m/z* 523.7751, 2+) over a 16-hour period (Figure 6). Mass measurement errors obtained from the Q-Tof micro equipped with the NanoLockSpray source were determined with the external lockmass correction. It can be seen that errors, when the single point external calibration was used, were 5 ppm or better over the 16-hour period.

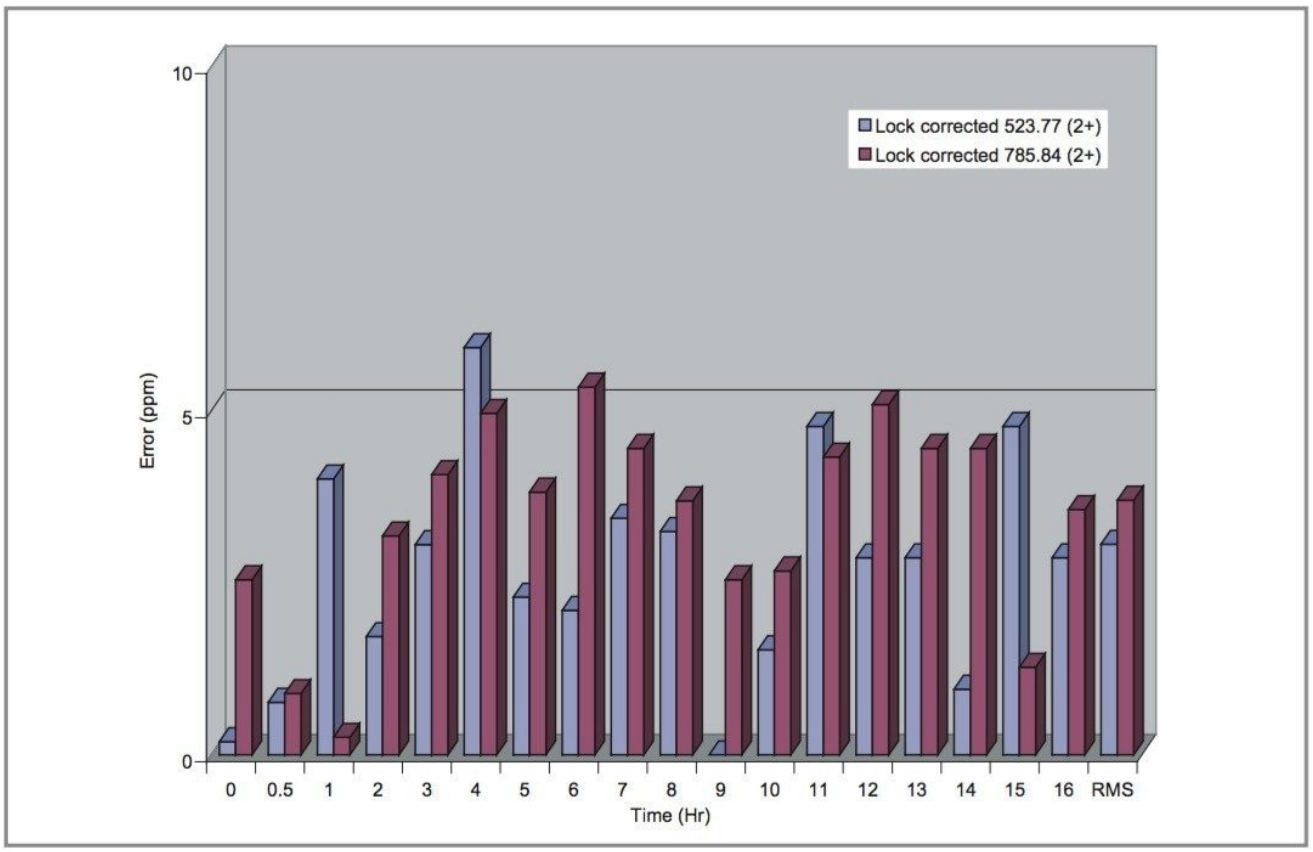


Figure 6. Mass measurement accuracy obtained from the Q-Tof micro equipped with the NanoLockSpray source. Two peptides, [Glu']-fibrinopeptide B (m/z 785.8426 2+) and Angiotensin II (m/z 523.7751 2+) were measured over a 16-hour period.

Digital Dead Time Correction (DDTC)

During a typical nanoflow LC-MS/MS experiment the tryptic peptides present exhibit a wide dynamic range and as such are detected by the mass spectrometer with varying signal intensities. This provides a challenge to achieving routine exact mass measurement as very intense peaks can cause the detector of the mass spectrometer to become saturated or move into 'dead time'.

Once a single ion has been detected there is a period during which further ion arrival events will not be detected. The result is a non-linear response between ions detected vs ions arriving at the detector. This manifests itself in a shift to a lower mass being reported for that particular ion.

An algorithm termed Digital Dead Time Correction (DDTC) can be used to correct for this effect. An

example of this effect is presented for data obtained from a Q-Tof Ultima API. This is presented in Figure 7 where a doubly charged ion at *m/z* 681.36 is shown both with and without DDTC applied during processing. The ion was assigned the sequence IQTQPGYANTLR during database searching. With DDTC applied the mass measurement for the ion was -1.9 ppm and without DDTC applied it was measured to be -19.8 ppm.

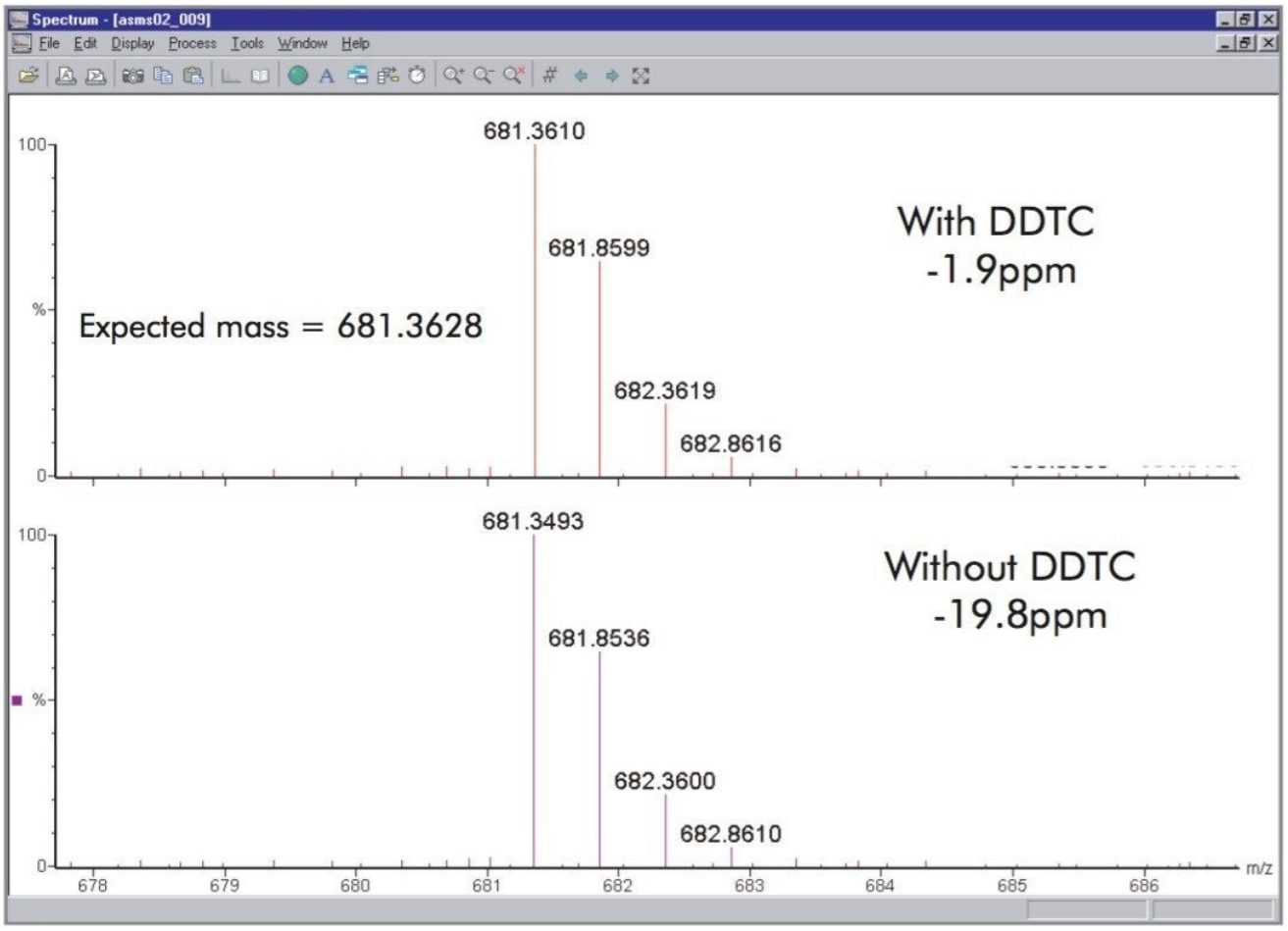


Figure 7. Doubly charged ion at m/z 681.36 processed both with and without the dead time correction algorithm applied.

Mass Measurement Accuracy and *de novo* sequencing

A challenge remains in an LC-MS/MS experiment in that often many of the MS/MS spectra acquired do not provide matches when searched against known protein sequence databanks. The nature of database

searching is such that only peptides that match exactly those within the databank will be identified. Consequently, many good quality spectra of novel peptides or of those containing a single amino acid substitution, or modification remain unmatched. Similarly, samples from organisms poorly represented in sequence databases can often produce a large number of unidentified spectra. Currently a common solution to this would be to extract these spectra manually from the data set, derive some degree of peptide sequence and perform further database searching. This, however, can be time consuming when large numbers of spectra are involved and is reliant upon the skill of the operator.

With the introduction of an automated computer-sequencing algorithm, MS/MS spectra can be identified by direct derivation of the novel peptide sequence through a double application of a Bayesian probabilistic analysis. A fragmentation model is applied to compare trial sequences against singly charged, de-isotoped MS/MS spectra. An overall confidence value for the most probable sequence, along with a confidence in the assignment of each individual residue, is calculated.

The MassSeq algorithm (ProteinLynx Global SERVER v2.0) utilizes the exact mass measurement of the MS/MS fragment ions to confidently define the most probable amino acid sequence. This is illustrated in Figure 8 where MS/MS spectra were acquired at different mass measurement accuracy using an ESI Q-Tof. It can be seen, in the resulting sequence determinations, that the data measured to less than 10 ppm provides significantly better results than that acquired at 100 ppm.

Peptide sequence from BSA	Lockspray corrected data (MassSeq 10 ppm tolerance)	Nominal mass data (MassSeq 100 ppm)
FKDLGEEHFK	FKDLGEEHFK	FKDLGEERKK
LCVLHEKTPVSEK	LCVLHEKTPVSEK	LCVLRLLLRRLD
CCTESLVNR	CCTESLVNR	CCTESLKKL
KVPQVSTPTLVEVSR	KVPQVSTPTLVEVSR	KVPQVSTPTLVEVSR
LVNELTEFAK	LVNELTEFAK	LVNELHTLPK
GFQNALLVR	(FG)QNALLVR	No result
YLYELAR	YLYELAR	YLYELAR

Green - correctly assigned residues
Red - incorrectly assigned residues
Yellow - ambiguous assignment

Figure 8. De novo sequencing of peptides is significantly enhanced through the specificity of exact mass MS/MS. The table demonstrates how greater confidence in sequence assignment is achieved with improved mass accuracy from LC-MS/MS analysis on the Q-Tof. De novo sequencing results are shown for a digest of Bovine Serum Albumin using ProteinLynx Global SERVER v2.0 Software.

Conclusion

- The NanoLockSpray source described here provides a routine method for obtaining enhanced mass measurements in LC-MS and MS/MS experiments
- Mass measurement using the NanoLockSpray source over a 16-hour period resulted in values ±5

ppm

- This interface provides unique functionality in a proteomics experiment, where exact mass measurement reduces false positive protein identifications and provides added confidence in those proteins identified
- Use of the NanoLockSpray source reduced the false positive identification of proteins even in the case of a BSA standard
- Identification of proteins from both standard digests and *E.coli* samples resulted in mass measurements better then 10 ppm
- *De novo* sequencing of peptides with ProteinLynx Global SERVER v2.0 is significantly enhanced through the specificity of exact mass MS/MS

Featured Products

ProteinLynx Global SERVER (PLGS) <https://www.waters.com/513821>

720000738, September 2003