Waters™

응용 자료

An Automated Processing of LC-MS Data for the Detection of Drug Impurities and Degradants in the Pharmaceutical Industry

Soraya Monté, Jose M. Castro-Perez, Steve Preece

SmithKline Beecham Pharmaceuticals, Waters Corporation

Abstract

This application note describes the use of MetaboLynx, an applications-manager from Micromass UK, which has been developed to observe trace impurities and degradants in drug substance and/or formulated product using data sets acquired by Liquid Chromatography-Mass Spectrometry (LC-MS). It has the potential to provide significant impact in the Drug Development process by reducing the data processing time of data sets that have been acquired by LC-MS. Set-up windows showing the selection criteria for peak detection parameters are shown along with the resulting data sets observable in a data browser. A number of examples from trace impurity profiling, key batch processing and degradation analyses are discussed.

Introduction

An increasing priority for pharmaceutical development is to increase the throughput of samples. Data mining is a key to speeding up the drug development process by handling large volumes of data, generated by modern techniques and automated technologies thereby aiding the making of well-informed decisions early on in the analytical process. MetaboLynx, the new automated peak detection system software developed by Micromass was initially designed to automate the detection and reporting of expected metabolites. This has now been extended to observe unexpected and/or unknown components and can therefore be used to find impurities and degradants in both drug substance and formulated product. It has the potential to provide significant impact in the drug development process by reducing the data processing time of data sets that have been acquired by LC-MS. In addition pharmaceutical companies put significant effort into ensuring that competitor companies are not infringing their patents. The search for route indicative impurities is extensive and often repetitive and this limits the number of potential patent infringements that can be studied at any one time. Manual examination of the data to confirm or refute the presence of components is time consuming and consequently ideal for automation.

MetaboLynx automates the process of repetitive plotting of mass chromatograms and examines them for significant peaks that exceed pre-set user-defined thresholds and filters. Following the integration of peaks above the threshold value, a background subtracted spectrum is produced and the resulting reduced data set viewed in the data browser.

MetaboLynx also has the potential for being used as a data reduction and pre-processing system for pattern

recognition techniques.

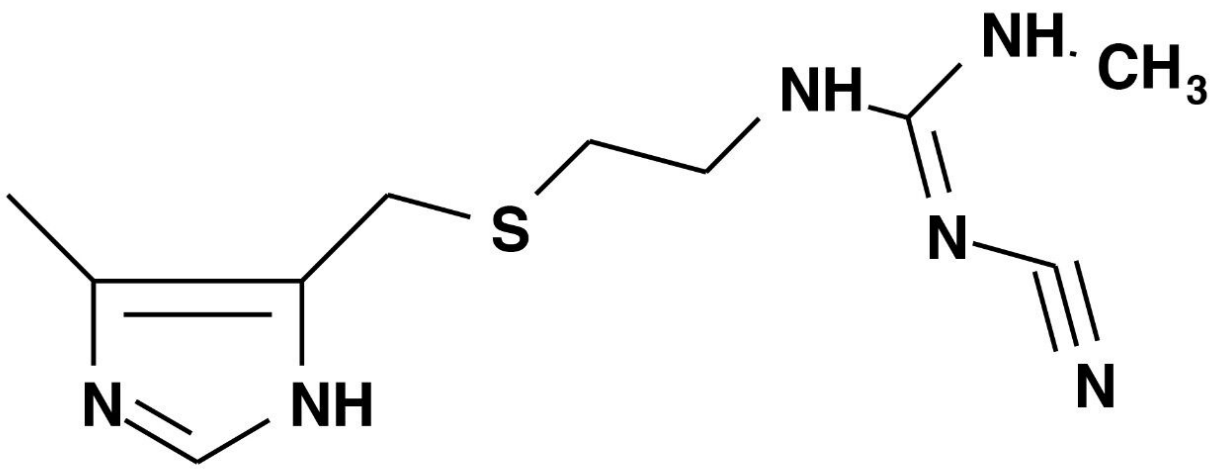
Experimental

This evaluation has been performed using MetaboLynx (Micromass UK Ltd., Altrincham, Cheshire, UK) beta-test v3.4.

All HPLC analyses were performed using Hewlett Packard HP1090/1100 (Hewlett Packard, Palo Alto, CA, USA) LC systems equipped with UV detectors. For mass spectrometric analyses, LC-MS experiments were carried out on the Quattro LC Triple Quadrupole Mass Spectrometer (Micromass UK Ltd., Altrincham, Cheshire, UK). Ionisation was carried out with a Z-SPRAY API source operating in the positive ion electrospray (ES+) mode.

All data were acquired in centroid mode.

Results and Discussion



Structure I Cimetidine (MW 252)

Figure 1 shows the raw Total Ion Current (TIC) chromatogram from the LC-MS data acquired from an extracted tablet solution of cimetidine using the Quattro LC.

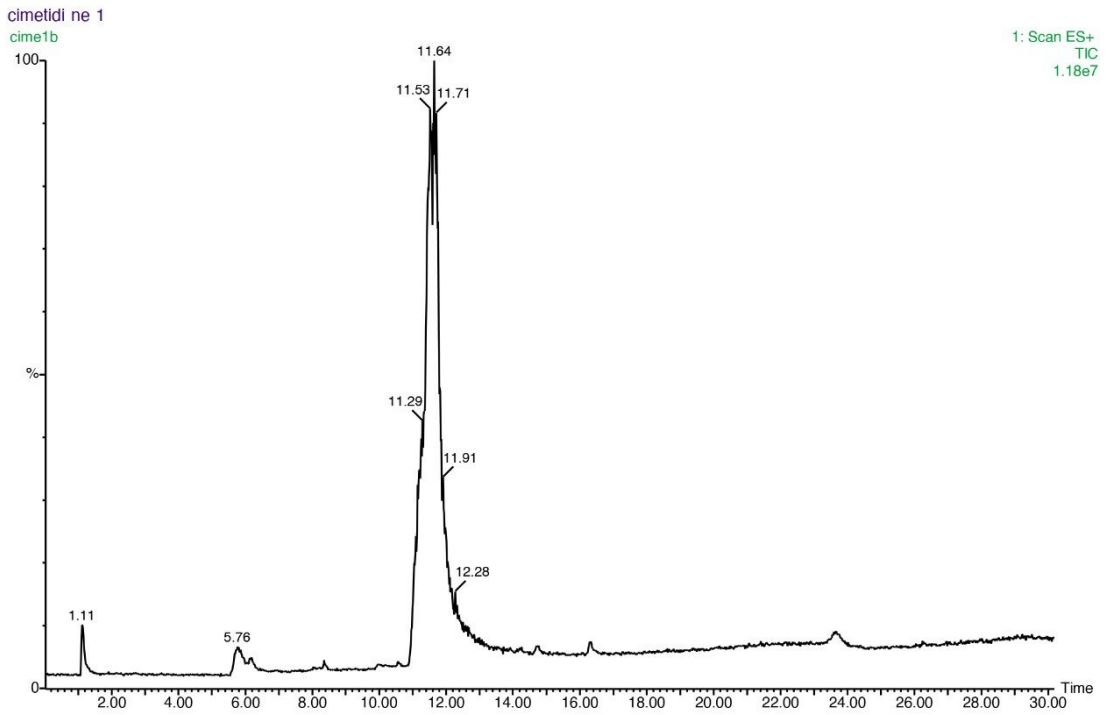


Figure 1. Total Ion Current chromatogram from the LC-MS analysis of a tablet of cimetidine.

In the analysis of trace and route indicative impurities in drug substance and/or formulated product, the TIC chromatographic traces obtained from LC-MS impurity profile analyses are often dominated by the main component and chemical noise which makes the detection of minor components difficult and time consuming. Peak overloading is often a necessity in order to detect and identify low level components in the matrix. In normal practice, often it is necessary to visually inspect hundreds of mass traces when carrying out route indicative impurity analyses on LC-MS data sets.

This same data was processed using MetaboLynx to identify the trace impurities in the tablet formulation. MetaboLynx is an application-manager designed to automate the detection and reporting of all expected and unexpected components arising in the TIC. This process involves the repetitive tasks of plotting Extracted Ion Chromatograms (XICs) for each possible component, examining them for significant peaks, producing and examining a spectrum for each peak and finally reporting the results. MetaboLynx automates these processes with the aim of removing a repetitive task from the analyst and accelerating data interpretation.

Before automated processing was initiated, a MetaboLynx method had to be set up to determine the criteria by which the data was to be judged. Peak detection parameters, which dictate how each chromatogram will be integrated, were stored in the method to produce a .mep file. Thresholds based on absolute or relative

chromatographic peak intensity can be set to filter out insignificant peaks. These criteria can be applied to MS data or separate parameter sets can be stored to allow peak detection based on photodiode array data or from the signal produced by an analog detector. In this case only mass chromatograms were used. Parameters to judge the quality of the mass spectrum produced from a detected peak were set. Figure 2 shows the set-up window where peak integration parameters based on mass chromatograms have been selected. The absence of a specified start and end time results in the entire chromatogram being processed.

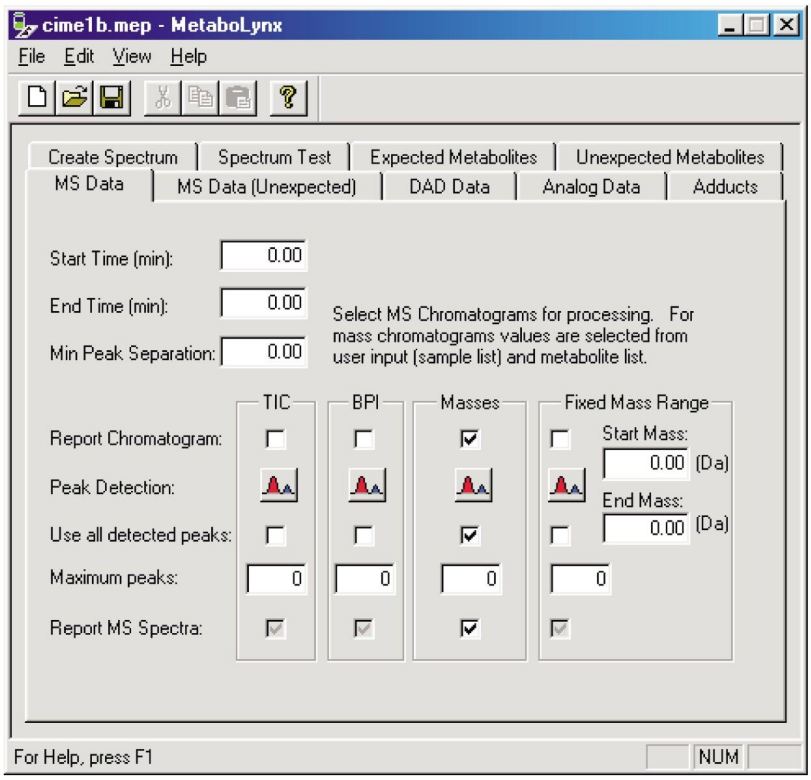


Figure 2. Set-up window showing selection criteria.

The mass chromatograms used to find the unknown components are also defined in the method. The system can be set up to integrate every mass chromatogram in the full scan LC-MS data file. Alternatively, the mass range can be divided into user-specified ranges, by using a step size of greater than one, and combined mass chromatograms can be used for peak detection.

The method also needed to include the expected component (cimetidine drug substance) in the data set. This set-up window is shown in Figure 3 and contains a description of the expected component, in this case, parent drug and the exact mass difference from the parent drug. For the analyses of forced degradation studies on drug substance or stressing studies during formulation development, expected components from a

number of common Phase I biotransformations pre-set in this menu can be selected e.g. hydroxylation (+16 amu) or demethylation (-14 amu). An example of this from a peroxide degradation study on drug substance will be shown later in this report. This is analogous to the detection of the "expected metabolites" function in MetaboLynx. For the purpose of this study, these were not selected. In addition, for peak tracking studies, new and expected entities may be created and added to the selection window.

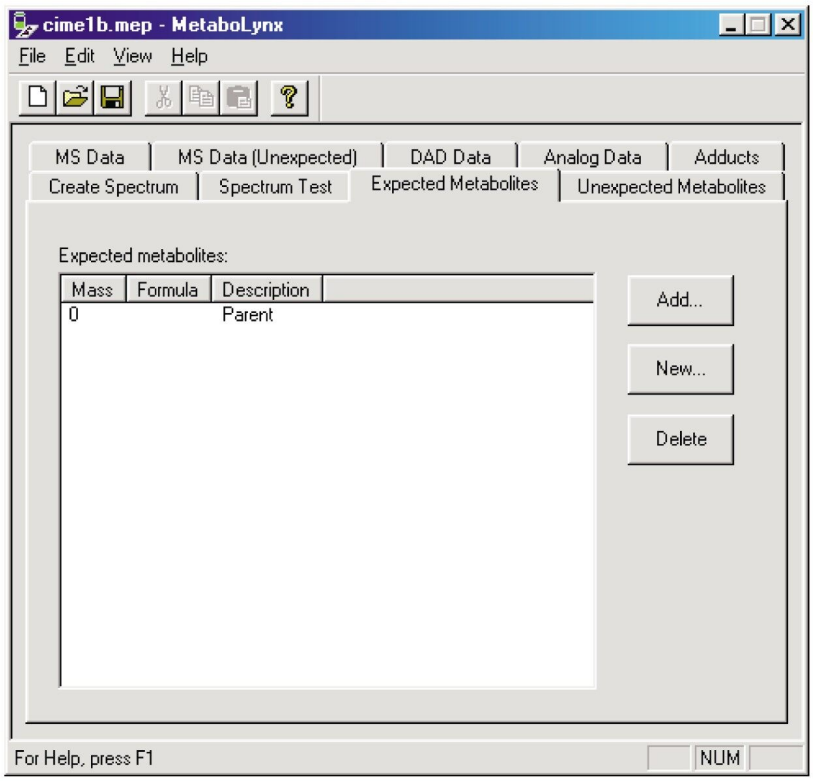


Figure 3. Set-up window showing the selection of the expected component, that of parent drug cimetidine.

Finally, the parameters required to create and report background subtracted spectra are setup. The MetaboLynx method is activated by appending it to a field in the MassLynx sample list. It is here that the molecular weight of the parent drug is also specified.

Once processing was initiated the software plotted an XIC for each mass in the acquisition scan range. Each chromatogram was integrated and, if any peaks were detected above the thresholds, a background subtracted spectrum was produced for each peak. The results of this automated processing were then written to a reduced data set. This can be viewed in a data browser and is shown in both Figures 4 and 5.

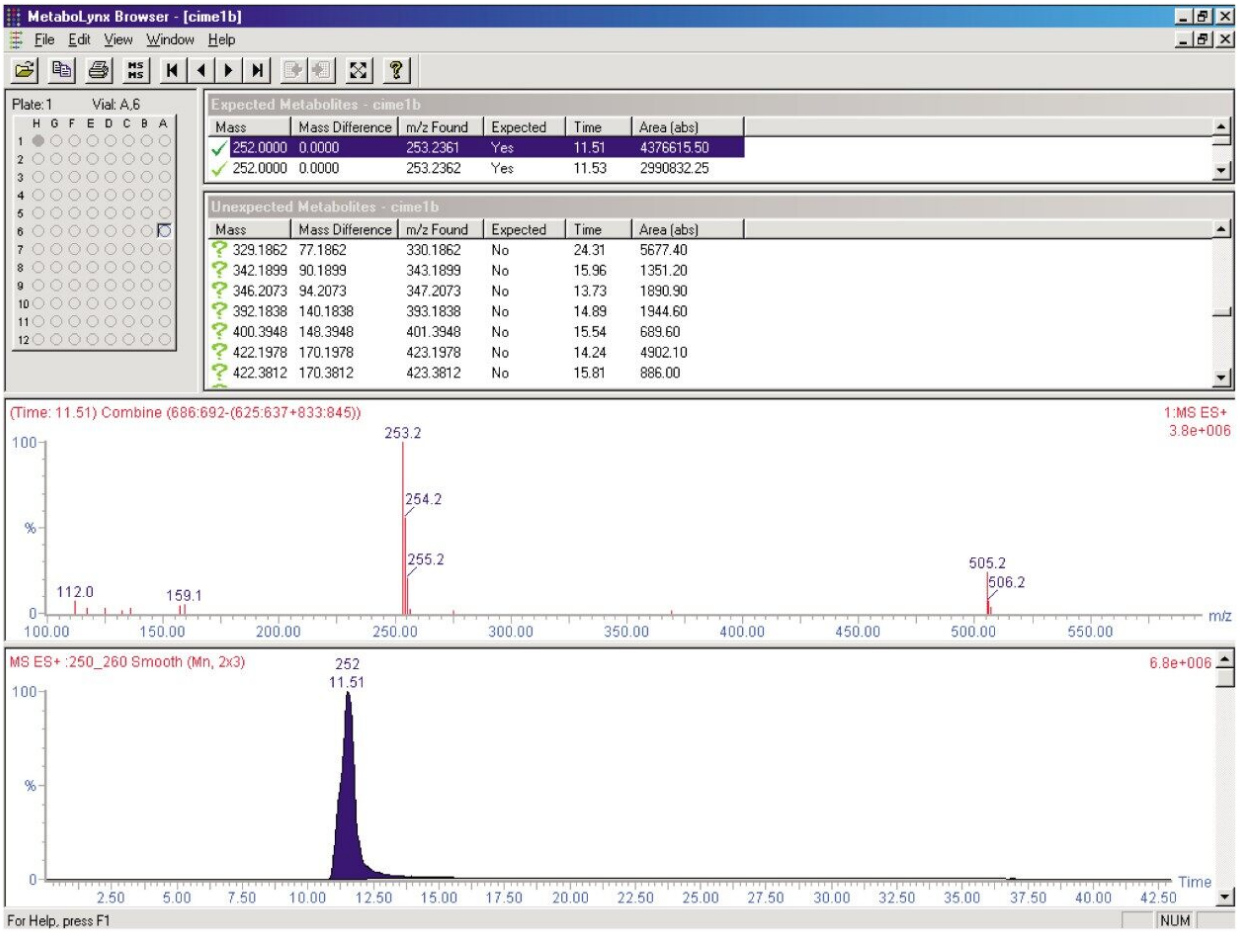


Figure 4. MetaboLynx data browser view showing results from the processing of an LC-MS data file from a tablet of cimetidine and its impurities. The above mass chromatogram and spectrum (m/z 253 $[M+H]^+$) corresponds to that of cimetidine (Structure I).

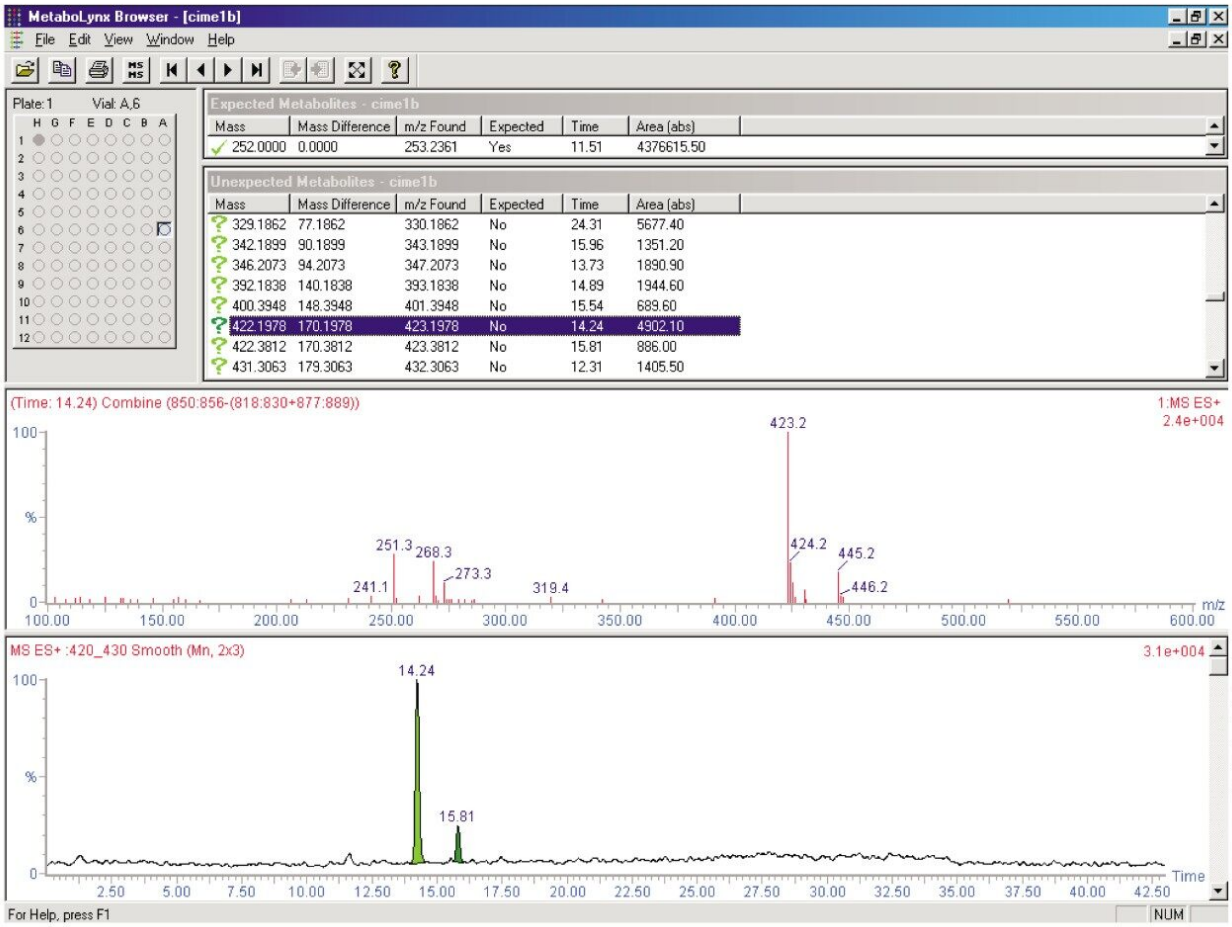


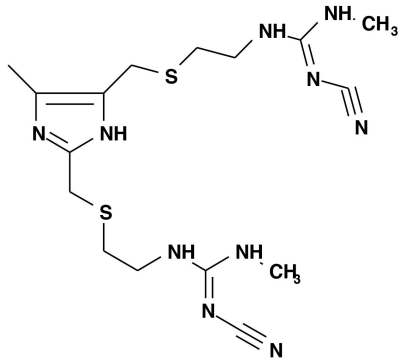
Figure 5. MetaboLynx data browser view showing results from the processing of an LC-MS data file from a tablet of cimetidine and its impurities. The above mass chromatogram and spectrum (m/z 423 $[M+H]^+$) corresponds to an impurity of cimetidine (Structure II).

The data browser contains five sections: a sample selection area which contains a representation of the autosampler bed used in the analysis, an expected component list showing parent drug, in this case cimetidine, an unexpected component list showing all detected impurities, a chromatogram view and a spectrum view. The list also includes information about the mass, retention time (RT) and peak area. If an entry in the list is highlighted, the mass chromatogram and spectrum associated with that component are displayed in the chromatogram and spectrum viewing areas.

Figure 4 shows the data browser view indicating the resulting entries, chromatogram and spectrum corresponding to the expected peak for cimetidine. The list shows that the parent drug, cimetidine has been identified with a molecular weight of 252, a mass to charge ratio of 253 at a retention time of 11:51 minutes and an associated peak area. The spectrum also highlights the $[2M+H]^+$ ion at *m/z* 505. The peak shape in the

chromatogram and the spectrum both indicate considerable overloading necessary to enable the detection of trace level impurities.

Figure 5 shows the data browser view indicating the resulting entries, chromatogram and spectrum obtained for the unknown and unexpected components. The list highlights a component with a molecular weight of 422 amu, a mass to charge ratio of 423 at a retention time of 14:24 minutes and an associated peak area. This corresponds to a known impurity of cimetidine (Structure II). The spectrum also shows the presence of the $[M+Na]^+$ adduct ion at *m/z* 445.



Structure II

The browser does not work from the raw data, but uses a reduced data set that is derived from the automated processing. This allows rapid review of the data without having to revert to reprocessing the original data. If there is any doubt about the results of automated processing the raw data is still available for examination manually, but the inclusion of a number of intelligent user selectable thresholds and filters in the method is designed to minimise this.

The software has been further developed to allow the user to rapidly browse through data which has had control sample related responses removed. The simultaneous processing of a mobile phase control sample with an analyte sample results in data only indicating those responses from the analyte sample alone. With the addition of a chromatogram and spectrum view due to the control sample, the data browser now consists of seven sections. The upper and lower chromatogram and spectrum displays, correspond to analyte and control sample results respectively. Figures 6 and 7 show the resulting displays from such an experiment using an LC-MS impurity profile data file acquired on drug substance and an associated mobile phase control. The sample display area shows blue and green circles representing the control and analyte samples in vials 4 and 5 respectively. In Figure 6, highlighting the *m/z* 224 ion in the unexpected component list shows the detection of an impurity at 12.8 minutes in the analyte sample in the analyte chromatogram

view. The absence of this peak in the control chromatogram view indicates that this component is not observed in the control sample and is consequently a sample related impurity.

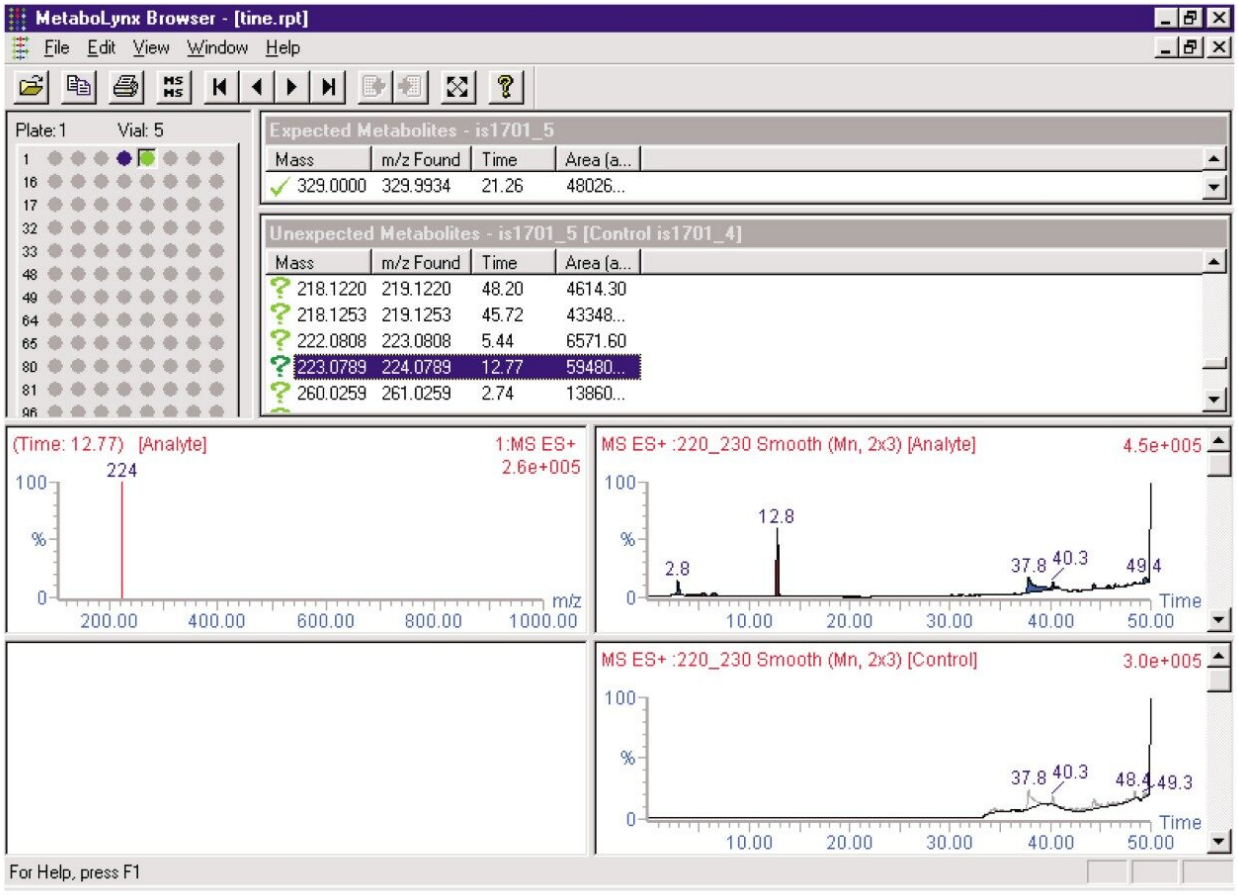


Figure 6. MetaboLynx data browser view showing presence of an analyte sample related impurity of m/z 224 with a retention time of 12.8 minutes.

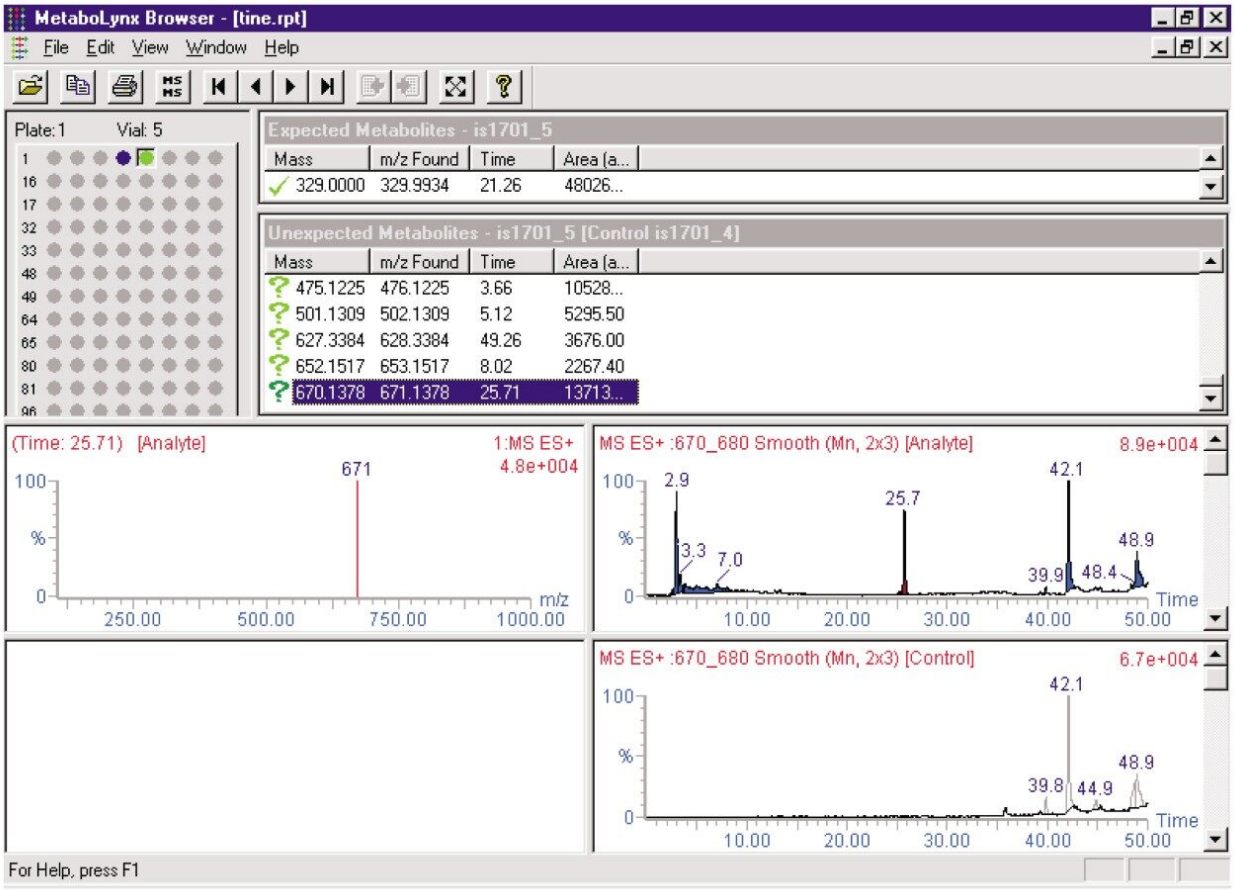


Figure 7. MetaboLynx data browser view showing analyte and control mobile phase related peaks.

In Figure 7, highlighting the *m/z* 671 ion in the unexpected component list shows the detection of four major peaks at retention times of 2.9, 25.7, 42.1 and 48.9 minutes in the analyte sample in the analyte chromatogram view. The absence of the peaks at the retention times of 2.9 and 25.7 minutes in the control chromatogram view indicate that these components are sample related whereas the peaks at 42.1 and 48.9 appear to be arising from the mobile phase control.

Identification of Impurities in Development Support

As part of our batch impurity profiling process, MetaboLynx has been used to rapidly screen thirteen LC-MS data sets from batch impurity profiles of starting material for one of our major drugs in development and is shown in Table 1.

Batch Number	unknown RRT 0.42	O-desmethyl impurity RRT 0.54 MW 268	unknown RRT 0.71 MW 267	unknown RRT 0.80	unknown RRT 0.81	Precursor RRT 0.82 MW 237	unknown RRT 0.86	unknown RRT 0.88	unknown RRT 0.96	Isomer RRT 1.02 MW 282	Dichlorinated Precursor RRT 1.09 MW 271
I	0.04	0.29		0.10		0.06					
II						0.10				0.15	0.04
III			0.08			0.20				5.44	0.05
IV			0.10			0.22				0.83	0.14
V		0.11		0.05	0.03	0.04					
VI	0.04	0.33		0.03	0.03	0.04		0.03	0.04		
VII		0.50	0.03			0.05			0.04		
VIII		0.97	0.03			0.05	0.01	0.01			
IX		0.66	0.02			0.01	0.02	0.02			
X		1.76									
XI		1.07									
XII		3.29									
XIII		0.20									

Table 1. Batch Data for a Starting Material- Impurity Profile. (Inset values are peak area ratios based on HPLC/UV profiles).

Figures 8 and 9 show the resulting data browser views indicating the detection of both an o-desmethyl impurity of the starting material and the dichlorinated impurity of its precursor, at retention times of 13.81 and 28.25 minutes respectively. The abundance of peaks in the chromatogram view in Figure 9 between retention times 15.64 to 26.37 minutes are due to increasing baseline noise. The columns in Table 1 at relative retention times of 0.42, 0.80, 0.81, 0.86, 0.88 and 0.96 indicate peak area ratios of components determined by HPLC/UV. These were not observed by mass spectrometry. Vice versa many components detected by LC-MS were not observed in the HPLC/UV profiles and remain unreported.

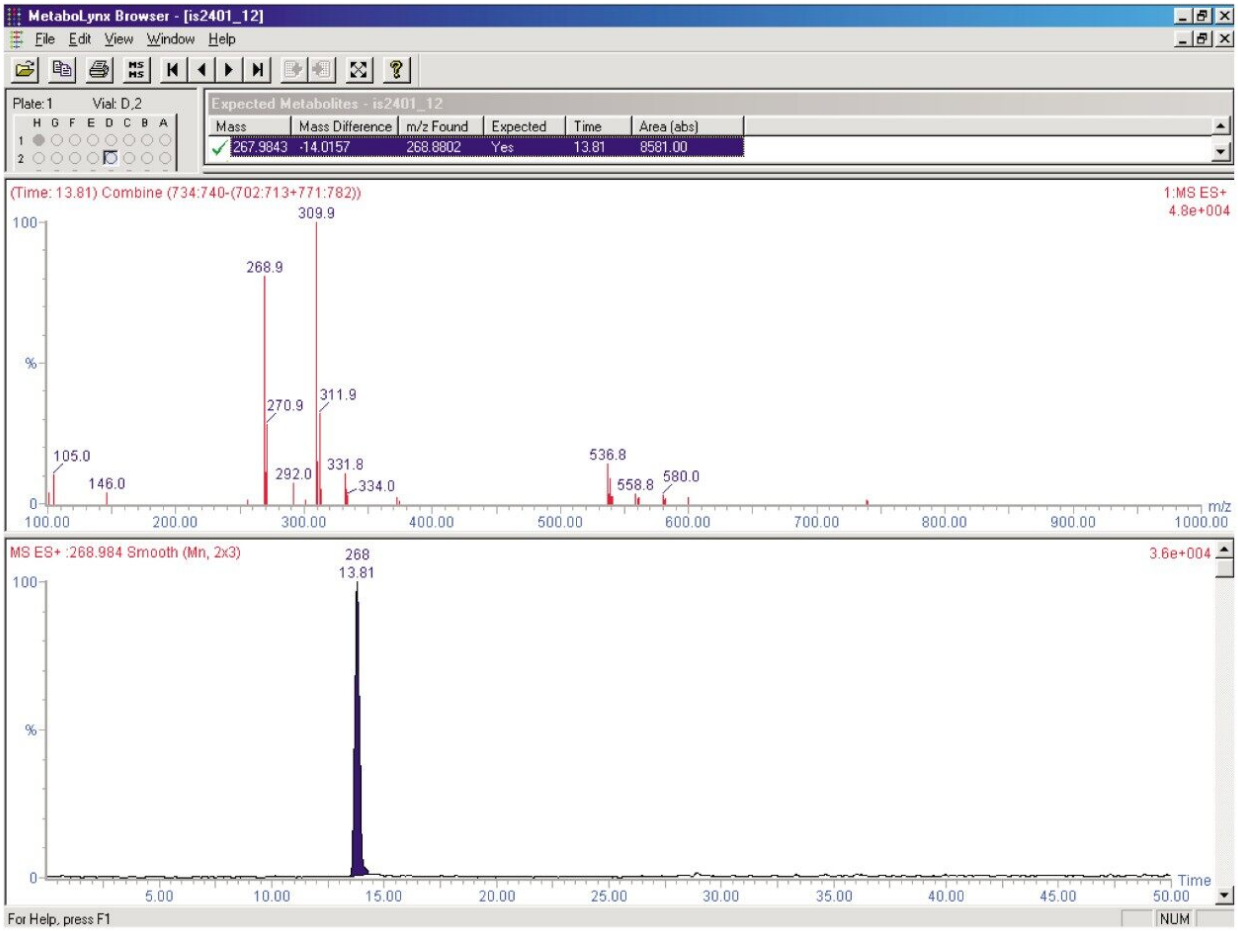


Figure 8. MetaboLynx data browser view showing the detection of an O desmethyl impurity at 13.81 minutes. Ions at m/z 269, 310 and 537 correspond to $[M+H]^+$, $[M+CH_3CN+H]^+$, and $[2M+H]^+$ respectively.

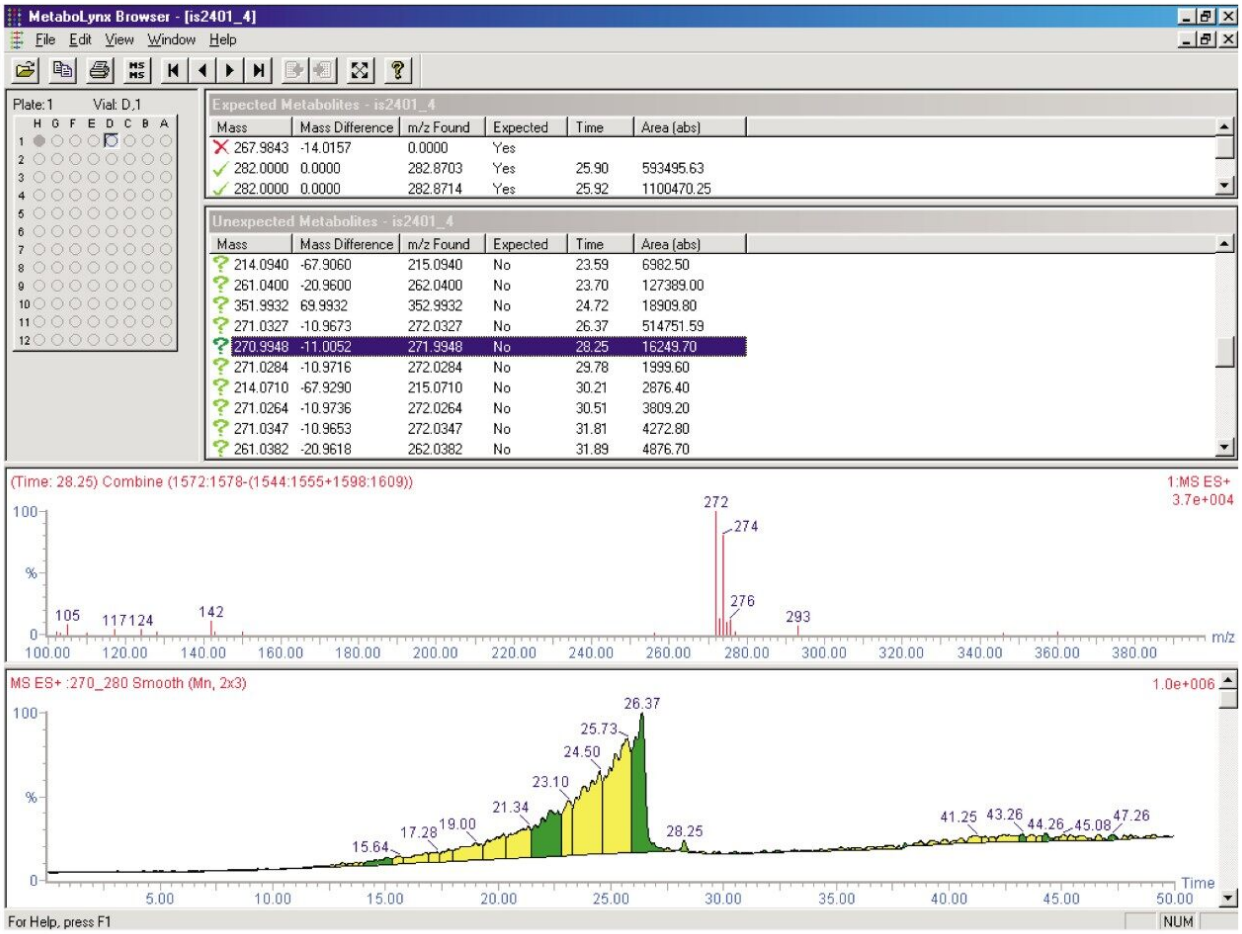


Figure 9. MetaboLynx data browser view showing the detection of a dichlorinated impurity of the precursor at 28.25 minutes (m/z 272 = $[M+H]^+$).

Identification of Degradants

Automated data processing by MetaboLynx was also carried out on a LC-MS data file acquired on a sample from a forced peroxide degradation study of drug substance. As previously described, expected components from a number of common Phase I biotransformations pre-set in the expected component set-up menu can be selected. This set-up menu is shown in Figure 10 and shows that we wish to detect expected +16 amu degradants such as n-oxides, aromatic and aliphatic hydroxylations. We also wish to determine if there is more than one oxidation reaction (+32 amu) and since this is a peroxide degradation, we also wish to determine if dehalogenation has occurred.

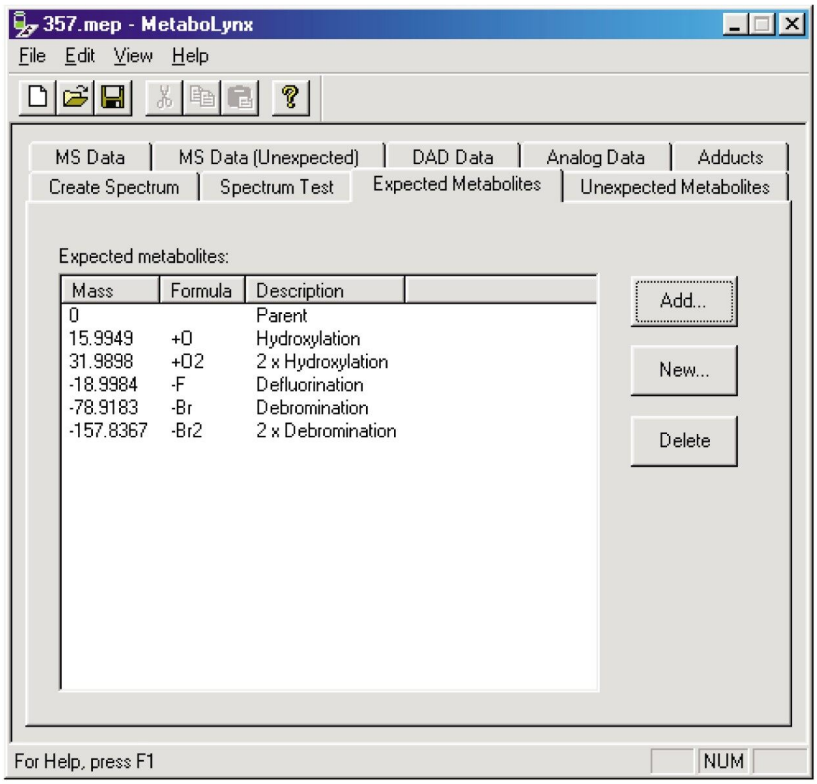


Figure 10. Set-up window showing the selection of expected components for a degradation reaction.

This LC-MS data file was processed by MetaboLynx and the resulting data browser indicating the presence of three +16 amu degradants in the chromatogram view is shown in Figure 11. The oxidative degradant at 21.60 minutes is highlighted and its associated mass spectrum with the characteristic isotope distribution pattern expected from a molecule containing two bromine atoms is observed in the spectrum view.

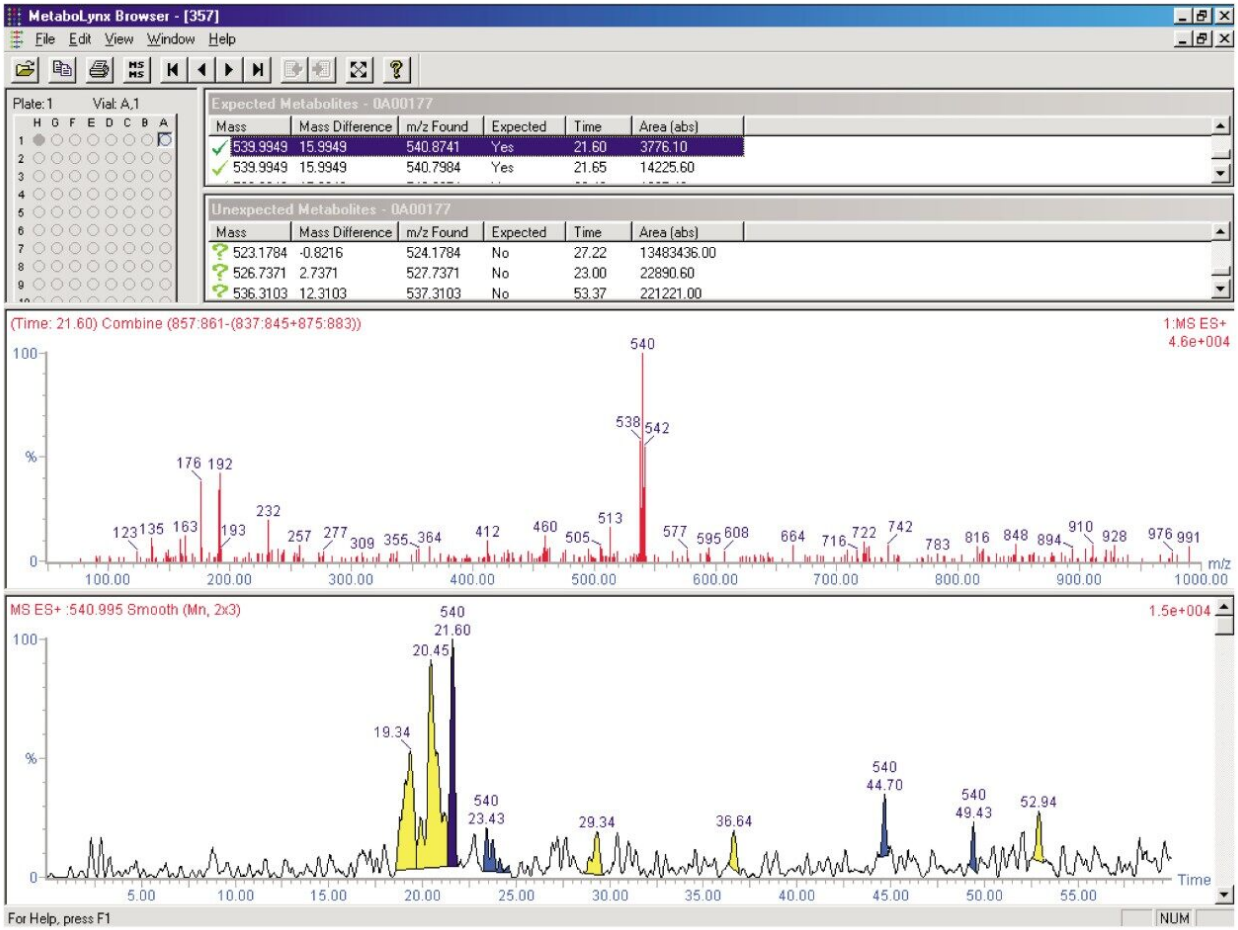


Figure 11. MetaboLynx data browser view showing three +16 amu degradants. The spectrum is that of a +16 amu degradant at a retention time of 21.60 minutes (m/z 540 = $[M+H]^+$).

This report has to date described the use of MetaboLynx as an automated LC-MS peak detection programme on low resolution centroided data sets. MetaboLynx also has the ability to process high resolution LC-MS data sets acquired on Time of Flight (Tof) analysers and will be the subject of a further communication.

In addition to the rapid reviewing of data which can be carried out by scrolling through the browser, the results of this reduced data set can be reported either in hard copy format in the form of a sample report or electronically transferred into Excel. This electronic transfer capability gives MetaboLynx the potential to be used as a preprocessing system for pattern recognition techniques using multivariate analytical methods. In addition, further evaluation of MetaboLynx to enhance its utility as a preprocessing tool prior to multivariate analysis is intended.

Conclusion

A preliminary evaluation of an automated peak detection system for processing LC-MS data obtained from trace impurity and degradation analyses has been carried out. MetaboLynx was able to detect and confirm the presence of many previously observed impurities and reported the results as a reduced data set that could be rapidly visually reviewed using a data browser. This will enable us to rapidly process large volumes of data in both the drug development and patent protection area thereby allowing us to speed up the drug development process and to increase the number of potential patent infringements that can be studied with current resources. This has the potential to provide significant impact by reducing the manual data processing time of LC-MS data sets in our mass spectrometry group.

Featured Products

MetaboLynx XS <https://www.waters.com/513803>

AN247, March 2000