Discovery of biomarkers in breath: Development and optimisation of a TD–GC×GC–TOF MS analytical platform

This study describes the development of an analytical platform based upon TD–GC×GC–TOF MS and shows its application to the collection and analysis of breath samples. The enhanced separation achieved ensures confident identification of biomarkers. Key advantages of the system described are the superior repeatability of flow modulated GC×GC and excellent linearity.

Introduction

Breath is a complex mixture of volatile and semi-volatile organic compounds, often present in trace amounts. Analysis of volatile organic compounds (VOCs) in exhaled breath offers the attractive prospect of diagnosing life-threatening diseases in a non-invasive and inexpensive fashion. In a clinical setting, breath-based tests have the potential to be used for the fast screening of large populations, improving patient survival rates and reducing costs for health services.

In this study, we develop a robust method for the collection and analysis of breath volatiles. Optimisation of the system is demonstrated – from the sampling of breath onto sorbent tubes for thermal desorption (TD), through to the data analysis workflows.

One of the key challenges in breath analysis is that hundreds of different VOCs can be present, often at trace levels – making it difficult to isolate and identify biomarkers of disease. During the biomarker discovery phase, an incorrect identification can compromise the validity of an entire trial, meaning that robust analytical techniques are required.

Comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC–TOF MS) has been proposed as a powerful technique for the analysis of such complex biological samples.
The commercialisation of consumable-free modulators for GC×GC operation has only strengthened its position for the analysis of breath volatiles – where routine analysis of large sample batches and trouble-free, long-term operation are essential.

The high sensitivity and enhanced separation of the TD–GC×GC–TOF MS system ensures that trace metabolites are not masked or overlooked and provides cleaner spectra for confident identification of potential biomarkers.

Here, we will demonstrate this enhanced performance through the analysis of sorbent tubes spiked with a suite of known biomarker standards. We will show the repeatability and linearity of the system, prior to the analysis of real breath samples. To simulate a real-world scenario, participants’ breath was measured before and after ingestion of a peppermint capsule to produce a controlled, artificial change in the breath profile and mimic a change in metabolism due to disease.

The results demonstrate a comprehensive method for the collection and analysis of VOCs in breath by TD–GC×GC–TOF MS.

**Experimental**

**Sampling:** 1 L breath samples were collected from healthy volunteers using Tedlar® breath sampling bags, prior to transfer to ‘Biomonitoring’ TD sorbent tubes (Markes International), using an ACTI-VOC™ low flow pump (Markes International).

**Standard addition:** Automated spiking of ‘Biomonitoring’ TD sorbent tubes (Markes International) with 1 μL of a standard mix of known biomarkers (as listed in Table 1) using a bespoke sample preparation robot (SepSolve Analytical). The design and workflow of the bespoke system is shown in Figure 1.

**TD:** Instrument: TD100-xr™ (Markes International) for automated desorption of up to 100 tubes. Alternatively, the Centri® platform (Markes International) can be used for automated desorption of up to 50 tubes as well as automated SPME(−trap), headspace(−trap) and/or high-capacity HiSorb™ sorptive extraction to extend analysis to sputum and other body fluids.

**GC×GC:** Modulator: INSIGHT® flow modulator (SepSolve Analytical); PM 4.0 s.

**TOF MS:** Instrument: BenchTOF-Select™; Mass range: m/z 35–600.
**Results**

1. Repeatability and separation

A major advantage of flow modulated GC×GC over thermally modulated systems (aside from the reduction in running costs) is its superior repeatability. INSIGHT’s precisely defined microfluidic design allows identical configurations to be installed across multiple instruments thus minimising uncertainty. This is in stark contrast to thermal devices, where small variations in column position can have a large impact on results.

The repeatability of the TD–GC×GC–TOF MS system was evaluated by 10 replicate analyses of sorbent tubes spiked with a standard mix of known biomarkers. Automated spiking using a bespoke sample preparation robot (SPR) speeds up workflows and ensures that human errors are minimised for increased confidence in results.

The separation of the biomarkers by GC×GC–TOF MS is demonstrated in Figure 2 – even in this simple standard, numerous co-elutions would have occurred in a conventional 1D GC system (as highlighted in the expanded region).

The results shown in Table 1 demonstrate the excellent repeatability of the entire system (from automated spiking through to GC×GC–TOF MS separation and detection) with an average relative standard deviation of <5% for the peak area. This level of repeatability is required for confident results during large-scale clinical trials, which could span many weeks, if not months.

Additional confidence in results can be obtained by secure sample re-collection. A breath sample, taken in the framework of a clinical study, captures a unique snapshot of the patient’s clinical journey. Consequently, if for any reason the sample is not correctly analysed, it will be lost. With the unique trap design of Markes’ TD systems, not only are the VOCs introduced into the GC in a narrow...
band, for optimal chromatography and high sensitivity, it is also possible to re-collect the split flow onto a clean sorbent tube for repeat analysis or archiving, meaning that TD is no longer a 'one-shot' technique.

<table>
<thead>
<tr>
<th>Peak no.</th>
<th>Compounds</th>
<th>1t&lt;sub&gt;R&lt;/sub&gt; (min)</th>
<th>2t&lt;sub&gt;R&lt;/sub&gt; (s)</th>
<th>1t&lt;sub&gt;R&lt;/sub&gt;</th>
<th>2t&lt;sub&gt;R&lt;/sub&gt;</th>
<th>Peak area</th>
<th>EIC</th>
<th>R&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Allyl methyl sulfide</td>
<td>3.4342</td>
<td>2.1984</td>
<td>0.12</td>
<td>0.71</td>
<td>4.33</td>
<td>88</td>
<td>0.9987</td>
</tr>
<tr>
<td>2</td>
<td>2-Methylbutan-2-ol</td>
<td>3.8170</td>
<td>2.1536</td>
<td>0.00</td>
<td>0.75</td>
<td>4.01</td>
<td>73</td>
<td>0.9982</td>
</tr>
<tr>
<td>3</td>
<td>Dimethyl disulfide</td>
<td>4.5906</td>
<td>2.2738</td>
<td>0.01</td>
<td>0.72</td>
<td>4.41</td>
<td>94</td>
<td>0.9979</td>
</tr>
<tr>
<td>4</td>
<td>Hexanal</td>
<td>4.7508</td>
<td>2.8655</td>
<td>0.03</td>
<td>0.56</td>
<td>5.73</td>
<td>56</td>
<td>0.9975</td>
</tr>
<tr>
<td>5</td>
<td>Butan-1-ol</td>
<td>5.5231</td>
<td>2.0955</td>
<td>0.07</td>
<td>0.76</td>
<td>4.66</td>
<td>56</td>
<td>0.9987</td>
</tr>
<tr>
<td>6</td>
<td>Cyclopentanone</td>
<td>6.3858</td>
<td>2.8784</td>
<td>0.11</td>
<td>0.58</td>
<td>3.22</td>
<td>84</td>
<td>0.9988</td>
</tr>
<tr>
<td>7</td>
<td>Heptanal</td>
<td>6.4632</td>
<td>3.3362</td>
<td>0.01</td>
<td>0.50</td>
<td>5.69</td>
<td>70</td>
<td>0.9955</td>
</tr>
<tr>
<td>8</td>
<td>Octanal</td>
<td>8.8567</td>
<td>3.7195</td>
<td>0.01</td>
<td>0.43</td>
<td>4.48</td>
<td>84</td>
<td>0.9943</td>
</tr>
<tr>
<td>9</td>
<td>Dimethyl trisulfide</td>
<td>11.0898</td>
<td>2.4811</td>
<td>0.09</td>
<td>0.60</td>
<td>5.59</td>
<td>126</td>
<td>0.9992</td>
</tr>
<tr>
<td>10</td>
<td>Nonanal</td>
<td>11.7171</td>
<td>3.9494</td>
<td>0.05</td>
<td>0.36</td>
<td>5.91</td>
<td>98</td>
<td>0.9915</td>
</tr>
<tr>
<td>11</td>
<td>Decanal</td>
<td>14.8149</td>
<td>0.0368</td>
<td>0.08</td>
<td>0.56</td>
<td>5.27</td>
<td>70</td>
<td>0.9954</td>
</tr>
<tr>
<td>12</td>
<td>Benzaldehyde</td>
<td>14.9874</td>
<td>2.5243</td>
<td>0.08</td>
<td>0.64</td>
<td>5.76</td>
<td>105</td>
<td>0.9944</td>
</tr>
<tr>
<td>13</td>
<td>Propanoic acid</td>
<td>15.8867</td>
<td>1.8832</td>
<td>0.03</td>
<td>0.88</td>
<td>5.29</td>
<td>74</td>
<td>0.9958</td>
</tr>
<tr>
<td>14</td>
<td>Butanoic acid</td>
<td>18.4190</td>
<td>1.8923</td>
<td>0.00</td>
<td>0.80</td>
<td>2.69</td>
<td>73</td>
<td>0.9949</td>
</tr>
<tr>
<td>15</td>
<td>Pentanoic acid</td>
<td>21.4002</td>
<td>1.8968</td>
<td>0.05</td>
<td>0.80</td>
<td>5.74</td>
<td>73</td>
<td>0.9979</td>
</tr>
<tr>
<td>16</td>
<td>Hexanoic acid</td>
<td>24.2774</td>
<td>1.9045</td>
<td>0.01</td>
<td>0.77</td>
<td>5.24</td>
<td>73</td>
<td>0.9918</td>
</tr>
<tr>
<td>17</td>
<td>m-Cresol</td>
<td>29.9185</td>
<td>1.8683</td>
<td>0.00</td>
<td>0.83</td>
<td>4.65</td>
<td>108</td>
<td>0.9994</td>
</tr>
<tr>
<td>18</td>
<td>4-Ethylphenol</td>
<td>31.9175</td>
<td>1.8764</td>
<td>0.04</td>
<td>0.84</td>
<td>5.38</td>
<td>107</td>
<td>0.9998</td>
</tr>
<tr>
<td>19</td>
<td>Nonanoic acid</td>
<td>32.1705</td>
<td>1.9281</td>
<td>0.01</td>
<td>0.90</td>
<td>5.20</td>
<td>60</td>
<td>0.9901</td>
</tr>
</tbody>
</table>

### 2. Linearity

Calibration curves were prepared for all compounds in the biomarker mix, with quantitation fully automated in ChromSpace software. All R<sup>2</sup> values were found to be over 0.99 (see Table 1), indicating excellent linearity even for acids, which are challenging compounds.

Figure 3 shows calibration curves for a selection of the compounds. Quantitation results are easily exported from ChromSpace, with laboratory information management system (LIMS) compatibility assisting with streamlined lab workflows.
3. Real-world application 1: ‘Peppermint oil’ experiment

Researchers working in the field of breath analysis are often faced with a major challenge: finding ‘the needle’ (i.e. biomarkers) within the ‘haystack’, represented by the wealth of features generated by the analytical platform. In fact, biomarkers need to be correctly identified and quantified, sometimes at trace levels, within a complex matrix.

Here, we demonstrate the enhanced performance of this multi-functional TD–GC×GC–TOF MS system by simulating disease biomarkers in breath by means of a controlled experiment. Two healthy volunteers (one male, one female) ingested a commercially available, food-grade peppermint oil capsule. A breath sample was taken immediately before and 30 min after capsule ingestion.

The underlying hypothesis is that VOCs contained in the capsule would enter systemic circulation and eventually reach breath after being exchanged in the lungs. Similar protocols have been adopted by breath researchers as quality control of breath analytical tool chains,[1] or as a proof of concept of substrate-based breath tests.[2] The experiment also mimics the changes in metabolism taking place, for example, as a result of disease.

Figure 4 shows the breath profiles for one of the participants before and after peppermint oil ingestion. As expected, the resulting profile is chemically diverse, including a wide variety of hydrocarbons, sulfur-containing VOCs and oxygenated compounds in a wide range of concentrations. The enhanced separation of the GC×GC system ensures that trace metabolites are not masked or overlooked and provides cleaner spectra for confident identification of the pseudo-biomarkers against existing commercial libraries, such as NIST.

A marked increase was observed for a number of terpenes (as highlighted in the expanded region of Figure 4) – which are plant metabolites commonly found in essential oils, such as peppermint oil. The peak areas for the annotated terpenes were found to increase between 2–8-fold in the breath samples taken 30 minutes after consumption of a peppermint capsule (Figure 5).
4. Real-world application 2: Natural breath spiking

In a second study, a participant’s breath was collected before and after ingestion of a cod liver oil and garlic capsule. The goal was to perform ‘natural’ breath spiking with sulfurs and aldehydes. Sulfurs are found naturally in garlic, whereas aldehydes should be obtained from the oxidative degradation of polyunsaturated fatty acids from the cod liver oil.

It is important to be able to identify these compounds confidently, as aldehydes
can be markers of oxidative stress and have been proposed as cancer biomarkers, while sulfurs have been proposed as diagnostic biomarkers of malaria or liver disease. Figure 6 shows the results of the experiment, whereby the aldehydes and sulfurs can be seen to appear in the breath profile of a participant just 10 min after ingestion of the capsule.

**Figure 6**
Breath samples from a participant before and after ingestion of a cod liver oil and garlic capsule.

## Conclusions

This study has shown that the described TD–GC×GC–TOF MS system offers numerous advantages for analysis of breath biomarkers:

- TD–GC×GC–TOF MS provides highly repeatable, linear results, for confident quantitation of biomarkers in large-scale clinical trials.
- Thermal desorption – thanks to trap focusing – ensures the introduction/injection of VOCs in a narrow band, for optimal chromatography and high sensitivity.
- Optional re-collection of a portion of the sample, either for storage or to re-analyse it to confirm the compounds identified, which is advantageous when trying to identify biometric patterns.
- INSIGHT–GC×GC provides efficient modulation of volatiles using cryogen-
free flow modulation for enhanced separation of complex breath profiles and robust biomarker discovery.

► Confident compound identification, based on high-quality spectra from BenchTOF mass spectrometers, with simple screening against existing commercial or in-house libraries.

► Automation can increase throughput and precision, with automated spiking of the sorbent tubes using gaseous or liquid standards, automated TD with unique tube caps to maintain sample integrity and automated re-collection of split flows for repeat analysis.

► Further time savings through use of ChromSpace for both instrument control and data-processing workflows.

For more information on this application, or any of the techniques or products used, please contact SepSolve.

References
