A New Universal Mass Spectrometry Data Analysis Software Suite

Graham A. McGibbon, Alexey Aminov, Vitaly Lashin, Andrey Vazhentsev

PittCon 2013
Who Needs Informatics?

88% of R&D organizations lack adequate systems to automatically collect data for reporting, analysis, and decision-making\(^1\)

\(^1\)Scientific Computing Research Study 2011
An Intelligence-from-Information ‘Live’ Cycle

ACD/Spectrus

Manage Unified Information
Generate Knowledge
Create Intelligence to Gain Insight

Unified Laboratory Intelligence

To connect chemical content with context for industries where chemistry matters
Manage Unified Information
Convert Heterogeneous Data to Homogeneous Structured Data

Data Systems and Formats
- ASCII
- ACD/Chrom
- ChemStation
- Chromeleon
- Empower
- MATLAB DSO
- Millennium
- TotalChrom
- Turbochrom
- Unidata netCDF

Chromatography
- Analyst
- ACD
- ChemStation
- Chromeleon
- Empower
- Data Explorer
- Empower LCMsSolution
- JCAMP
- GCMSolution
- MASSPEC
- MassHunter
- MassLynx
- MATLAB DSO
- Millennium
- MzML
- NIST MS & SDF
- Unidata netCDF
- Xcalibur

Mass Spectrometry
- Cary UV
- ChemStation
- Chromeleon
- Empower
- FTIR
- Galactic
- JCAMP
- LabControl
- MassLynx
- MATRIX DSO
- Massspec
- Millennium
- NIST OMNIC
- NuSpec
- OPUS
- Spectra Manager
- SpectraSuite
- Spectrum
- Vision

Optical Spectroscopy
NMR Spectroscopy
- Acorn MacNuts
- Acorn WinNuts
- ACD
- Delta
- DSpect
- Felix (MSI)
- GRAMS/Al
- JCAMP
- Nicolet
- TopSpin
- UX_NMR
- VNM
- WIN-NMR

Thermal Analysis
- Pyris
- Stare
- Thermal Advantage

X-Ray Powder Diffraction
- ASCII
- DIFFRACT-AT
- DIFFRACT-PLUS
- PanAnalytic
- XRDML
- Rigaku D/MAX-B

Homogeneous information management software
Structured, Homogenized “Live” Data
Single Processing Interface

- 1D or 2D NMR
- IR, XRPD
- MS
- GC/MS
- LC/UV/MS

- All data in one executable program interface
Spectrus Informatics for Mass Spectrometry

ACD/MS Workbook Suite

IXCR

ACD/MS-Intelli Workbook
(Find components)

ACD/MS-ID Workbook
(Characterize)

ACD/MS Fragmenter

In-house
Commercial Content

ACD/Spectrus Processor
Why Screening?

• Rapid identification of contamination, residues, adulteration, agrochemicals, drugs and potentially toxic compounds

• Protect, Assess, Improve:
  ◦ health
  ◦ forensic samples
  ◦ food
  ◦ water
Component Finding

• Targeted or Generic Screening?
  ◦ Range of suspects

• Confirm, Verify, Identify, Elucidate?
  ◦ Dereplicate
• Collaboration to focus on the discovery of new lead matter for antibiotic, anti-inflammatory, & CNS development, nutritional and cosmetic agents
• Project places special focus on using and developing special methodology
  ◦ Dereplication
  ◦ Structure Elucidation
  ◦ Separations
  ◦ Data sharing
University of Leuven (Belgium)  
The University of Aberdeen (UK)  
Aquapharm Biodiscovery Ltd (UK)  
University of Tromsø (Norway)  
eCoast Research Centre (Belgium)  
Biobridge Ltd (UK)  
Fundación MEDINA (Spain)  
University College Cork, National University of Ireland  
BIOCOM AG (Germany)  
Stazione Zoologica Anton Dohrn (Italy)  
Consiglio Nazionale delle Ricerche ibp-cnr (Italy)  
University of Santiago de Compostela (Spain)  
The Royal Society of Chemistry (UK)  
c-LEcta GmbH (Germany)  
Technical University of Denmark  
Deep Tek Ltd (UK)  
Advanced Chemistry Development UK Ltd  
Wuhan University (China)  
Institute of Microbiology – Chinese Academy of Sciences (China)  
University of the Western Cape (South Africa)  
Institute for Cell Dynamics and Biotechnology (Chile)  
National Biodiversity Institute of Costa Rica  
International Union for Conservation of Nature (Switzerland)  
University of Waikato (New Zealand)
Flavour of the Month

Phenylbutazone

- NSAID used for short-term treatment of pain and fever in *animals* but no longer approved for human use in US and UK
- Can cause severe adverse effects such as suppression of white blood cell production and aplastic anemia
Food Safety

• Melamine scare prompted US government to be more proactive

• Food Safety Modernization Act of 2010, (FDA and USDA) shifting food safety regulation to prevention

• Companies will be required to develop preventive food safety controls and share these plans with the FDA

• FDA labs monitoring food samples
Screening Strategy

- Set up Targets
- Process
  - Extract
  - Componentize
- Interpret

Then IntelliXtract Generic Screening

Automation Amenable
Combining targeted and nontargeted data analysis for liquid chromatography/high-resolution mass spectrometric analyses

Croley, T. R., 1,* White, K. D., 1 Wong, J., 1 Callahan, J. H., 1 Musser, S. M., 1 Antler, M., 2 Lashin, V. 2 and McGibbon, G. A. 2

1 U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD, USA
2 Advanced Chemistry Development, Inc., Toronto, Ontario, Canada

Abstract: Increasing importation of food and the diversity of potential contaminants have necessitated more analytical testing of these foods. Historically, mass spectrometric methods for testing foods were confined to monitoring selected ions (SIM or MRM), achieving sensitivity by focusing on targeted ion signals. A limiting factor in this approach is that any contaminants not included on the target list are not typically identified and retrospective data mining is limited. A potential solution is to utilize high-resolution MS to acquire accurate mass full-scan data. Based on the instrumental resolution, these data can be correlated to the actual mass of a contaminant, which would allow for identification of both target compounds and compounds that are not on a target list (nontargets). The focus of this research was to develop software algorithms to provide rapid and accurate data processing of LC/MS data to identify both targeted and nontargeted analytes. Software from a commercial vendor was developed to process LC/MS data and the results were compared to an alternate, vendor-supplied solution. The commercial software performed well and demonstrated the potential for a fully automated processing solution.

Keywords: Automatic data processing; High-resolution mass spectrometry; Nontargeted analysis
Formula Generation

Can be complemented by Spectrum Simulation
# Accurate HRMS: HT-2 toxin

<table>
<thead>
<tr>
<th>No.</th>
<th>Formula</th>
<th>Monoisotopic Mass</th>
<th>Error</th>
<th>Pattern</th>
<th>RDBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C22H32O8</td>
<td>424.209</td>
<td>0.001</td>
<td>425.493</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>C29H28O3</td>
<td>424.203</td>
<td>0.006</td>
<td>425.542</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>C33H28O3</td>
<td>424.219</td>
<td>-0.009</td>
<td>425.588</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>C26H32O5</td>
<td>424.224</td>
<td>-0.015</td>
<td>425.539</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>C18H32O11</td>
<td>424.194</td>
<td>0.016</td>
<td>425.446</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>C19H36O10</td>
<td>424.230</td>
<td>-0.021</td>
<td>425.490</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>C25H28O6</td>
<td>424.188</td>
<td>0.022</td>
<td>425.495</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>C32H24O</td>
<td>424.182</td>
<td>0.028</td>
<td>425.545</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>C30H32O2</td>
<td>424.240</td>
<td>-0.030</td>
<td>425.585</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>C23H36O7</td>
<td>424.246</td>
<td>-0.036</td>
<td>425.536</td>
<td>6</td>
</tr>
</tbody>
</table>

![Chemical structure of HT-2 toxin](image)
Recognizing Components:

**LC**

De-noising & Deconvolution?

**GC**

IXCR with Wiley/NIST database searching
Pre-process on Import
Structure: Why and How?

• Given or to determine?
• In-house and/or public Sources?
• Directly or indirectly
  ◦ Names
  ◦ Identifiers
  ◦ Properties
  ◦ from Data & ‘Chemical Context’
Fragmentation Prediction
Knowledge Management
Impurities or Degradants Study

DMPK Biotrans Metabolite Database

Materials Characterisation

Chiral/Achiral Screening

Automated NMR Structure Verification

ACD/Spectrus Processors

ACD/Spectrus WorkBooks

ELN