Oligosaccharides, DNA, Metabolomics, ETD

Arpad Somogyi
CCIC MSP
OSU Summer Workshop

Automated interpretation of MS/MS spectra of oligosaccharides
Haixu Tang¹,²,*, Yehia Mechref³, and Milos V. Novotny³

Fig. 4.
The structures of the oligosaccharides used for testing in this paper.
Fig. 1. The structure of oligo- and polysaccharides. (a) The open (acyclic) form structure of glucose, as an epimer of arabinose, one of the mannose derivatives, i.e. the basic unit of an oligosaccharide. (b) The cyclic form structure of glucose. The carbon atoms in the ring are not shown. (c) A disaccharide, consisting of two glucose forms forming a 1-4 glycoside bond. (d) A tetrasaccharide, consisting of four glucose forms with a branching. (e) The symbolic representation of the tetrasaccharide shown in (d). The numbers show the linkage types.

Fig. 2. The fragmentation pattern of oligo- and polysaccharides. (a) The Y-0 and Z-X ions are from the glycosidic branchings. The cross-ring fragment ions are labeled. (b) The TBDMS-protected hexasaccharide. The bonds between small masses that are not close to the corresponding Y/X ions, e.g. 577, 720, 735, 551, 562 ions will be ignored in our program.

Fig. 3. The HDM-O of oligo- and polysaccharides. (a) An oligosaccharide consisting of U-1 glucose. The glucose residues are ordered according to the partial order of nodes of the tree. For any two residues $i$ and $j$, if in the subtree rooted by $i$, nodes cannot be identified smaller than $j$. (b) The subtrates rooted by residues $\gamma_2, \gamma_3, \gamma_4, \gamma_5$. 

2
Infrared Multiphoton Dissociation of O-Linked Mucin-Type Oligosaccharides


Jinhua Zhang, Katherine Schubothe, Bensheng Li, Scott Russell, and Carlito B. Lebrilla

Department of Chemistry and School of Medicine, Biochemistry and Molecular Medicine

University of California at Davis
Infrared Multiphoton Dissociation (IRMPD) in FT-ICR

IonSpec 4.7T FT-ICR, Electrospray Ionization (ESI) Source and Ion Guide

Infrared Multiphoton Dissociation (IRMPD)

J. Beauchamp et al., J. Am. Chem. Soc., 1978, 100, 3248-3250
Comparisons between CID and IRMPD of mucin-type oligosaccharides

CID of neutral oligosaccharide XT-1415 (positive mode)

MS/MS only produced high abundance for high mass fragments

Zhang et al.,
IRMPD of neutral oligosaccharide XT-1415

Why does IRMPD yield more extensive fragments?

CID deposits energy only into the precursor ions
IRMPD excites both precursor and fragment ions

Conclusion

Compared to CID, IRMPD has advantages:

- No collision gas needed, faster
- No multistage (MS^n) needed
- Easy and good control of energy deposited
- The fragmentation efficiency of IRMPD increases with the increasing size of oligosaccharides

IRMPD provides an attractive alternative to CID in the structural elucidation of oligosaccharides

SORI-RE of oligosaccharide Alditol

ESI negative mass spectrum of an oligonucleotide

FT-ICR Instrument

7- charge state

[M-7H]^-7-
[M-8H+Na]^-7-
[M-8H+K]^-7-
[M-9H+Na+K]^-7-

974.4564 977.5947

[M-8H+K]^-7-
[M-9H+Na+K]^-7-

974.6014

[M-9H+2K]^-7-

979.8747
Table 1
Experimental mass accuracy for standard oligonucleotides analyzed by MALDI-TOF and ESI

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>MALDI-TOF Calculated mass (D)</th>
<th>MALDI-TOF Experimental mass (D)</th>
<th>MALDI-TOF Error (ppm)</th>
<th>ESI-LC-MS Experimental mass (D)</th>
<th>ESI-LC-MS Error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mer</td>
<td>2993.0</td>
<td>2992.1</td>
<td>-300.7</td>
<td>2993.1</td>
<td>33.4</td>
</tr>
<tr>
<td>20 mer</td>
<td>6168.1</td>
<td>6167.4</td>
<td>-113.5</td>
<td>6167.6</td>
<td>-81.1</td>
</tr>
<tr>
<td>40 mer</td>
<td>12300.0</td>
<td>12298.6</td>
<td>-113.8</td>
<td>12299.7</td>
<td>-24.4</td>
</tr>
<tr>
<td>60 mer</td>
<td>18410.0</td>
<td>18414.6</td>
<td>249.9</td>
<td>18410.1</td>
<td>5.3</td>
</tr>
<tr>
<td>80 mer</td>
<td>24661.1</td>
<td>24664.7</td>
<td>146.0</td>
<td>24661.0</td>
<td>-4.1</td>
</tr>
<tr>
<td>100 mer</td>
<td>30951.1</td>
<td>30978.3</td>
<td>878.8</td>
<td>30950.6</td>
<td>-16.2</td>
</tr>
<tr>
<td>120 mer</td>
<td>37031.0</td>
<td>37121.6</td>
<td>2446.6</td>
<td>37030.0</td>
<td>-27.0</td>
</tr>
</tbody>
</table>

Source: Hail et al. (2004)

DNA mass calculator program:
http://medlib.med.utah.edu/masspec/mongo.htm

Partial Sequencing of Different Products of the Human Cytochrome P450 cyp2d6 Gene of Different Individuals

Enzymes
bovine spleen phosphodiesterase and snake venom phosphodiesterase
Review

Data processing for mass spectrometry-based metabolomics

Mikko Katajamaa,1,*, Matej Orešič2,***

1 Turku Centre for Biotechnology, P.O. Box 7, FIN-20521 Turku, Finland
2 VTT Technical Research Centre of Finland, Turku 2, P.O. Box 1350, FIN-02044 VTT, Espoo, Finland

Available online 19 April 2007

Abstract

Modern analytical technologies afford comprehensive and quantitative investigation of a multitude of different metabolites. Typical metabolomic experiments can therefore produce large amounts of data. Handling such complex datasets is an important step that has big impact on extent and quality at which the metabolic identification and quantification can be made, and thus on the ultimate biological interpretation of results. Increasing interest in metabolomics thus led to resurgence of interest in related data processing. A wide variety of methods and software tools have been developed for metabolomics during recent years, and this trend is likely to continue. In this paper we overview the key steps of metabolomic data processing and focus on reviewing recent literature related to this topic, particularly on methods for handling data from liquid chromatography mass spectrometry (LC-MS) experiments.

Keywords: Metabolomics; Lipidomics; Proteomics; Normalization; Alignment; Liquid chromatography; Mass spectrometry; Feature extraction; Peak detection; Deconvolution
Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation

B. Álvarez-Sánchez, F. Prigo-Capoño, M.D. Luque de Castro

![Diagram showing the workflow of metabolomics analysis](image-url)

Figure 1. General workflow of the main steps involved in conventional metabolomics analysis.
Figure 2. Venn diagram illustrating the number of metabolites detected in the metabolome analysis of *Escherichia coli* using four different extraction procedures: methanol, ethanol, chloroform-methanol and potassium hydroxide. The diagram includes information about the number of metabolites detected with each protocol independently and with different protocols simultaneously (Reproduced with permission of the American Chemical Society, [5]).

Figure 3. Global metabolomics profiling of serum and urine samples with GC-TOF-MS using PCA and PLS-DA models. The two sets of samples were analyzed at 5% and 0.1% for serum samples at 3P, 3.1% and 2.1% for the serum, respectively. PCA loadings plot for the two sets show strong signals with PCA and PLS-DA for serum samples. The diagrams show metabolites responsible for the variance contained in the PLS-DA loadings plot, represented with percentage or absolute intensity (new view).
Metabolomics analysis II. Preparation of biological samples prior to detection

B. Álvarez-Sánchez, F. Priego-Capote, M.D. Luque de Castro

Figure 1. Automated approach for quenching mammalian cells by on-line coupling of a heat exchanger used as a fast-sampling device to a bioreactor where the metabolism is stopped (Reproduced with permission of Springer-Verlag, [16]).

Figure 2. Steps followed in simultaneous or sequential approaches used for extraction of metabolites from cells (Reproduced with permission of the American Chemical Society, [50]).
Update in Bioinformatics

Web-based resources for mass-spectrometry-based metabolomics: A user's guide

Takayuki Tobe, Alisdair R. Fernie *

Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Potsdam-Cohn, Germany

Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Fryday available&quot; software</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enzy</td>
<td><a href="http://www.enzyme.org">http://www.enzyme.org</a></td>
<td>Enzyme</td>
</tr>
<tr>
<td>KPPA View</td>
<td><a href="http://www.kppaview.com">http://www.kppaview.com</a></td>
<td>KPPA View</td>
</tr>
<tr>
<td>XEGG</td>
<td><a href="http://www.genome.g">http://www.genome.g</a> phases.gpc</td>
<td>XEGG</td>
</tr>
<tr>
<td>Metacyc</td>
<td><a href="http://www.metacyc.org">http://www.metacyc.org</a></td>
<td>Metacyc</td>
</tr>
<tr>
<td>PlantCyc</td>
<td><a href="http://www.plantcyc.org">http://www.plantcyc.org</a></td>
<td>PlantCyc</td>
</tr>
<tr>
<td>Metabolic data (samples)</td>
<td>Analogs, potato, plant</td>
<td></td>
</tr>
<tr>
<td>Arabidopsis</td>
<td><a href="http://www.arabidopsis.org">http://www.arabidopsis.org</a></td>
<td>Arabidopsis</td>
</tr>
<tr>
<td>MGCAMDB</td>
<td><a href="http://www.mgcambdb.org">http://www.mgcambdb.org</a></td>
<td>MGCAMDB</td>
</tr>
<tr>
<td>MassDB</td>
<td><a href="http://www.massdb.org">http://www.massdb.org</a></td>
<td>MassDB</td>
</tr>
<tr>
<td>Analytical data (standard compounds)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOWSEK</td>
<td><a href="http://www.embnet.net">http://www.embnet.net</a></td>
<td>MOWSEK</td>
</tr>
<tr>
<td>MBAT</td>
<td><a href="http://www.mbat.molgen.mpg.de">http://www.mbat.molgen.mpg.de</a></td>
<td>MBAT</td>
</tr>
<tr>
<td>ORBM</td>
<td><a href="http://orbmui.com">http://orbmui.com</a></td>
<td>ORBM</td>
</tr>
<tr>
<td>Publications and Literature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MolNet</td>
<td><a href="http://molnet.com">http://molnet.com</a></td>
<td>MolNet</td>
</tr>
<tr>
<td>PDB</td>
<td><a href="http://www.pdb.org">http://www.pdb.org</a></td>
<td>PDB</td>
</tr>
<tr>
<td>WebMDE</td>
<td><a href="http://www.cheminformatics.com/webmde.html">http://www.cheminformatics.com/webmde.html</a></td>
<td>WebMDE</td>
</tr>
<tr>
<td>CCSS</td>
<td><a href="http://cds.cas.ens.fr">http://cds.cas.ens.fr</a></td>
<td>CCSS</td>
</tr>
<tr>
<td>Compendia Library</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino Acid Library</td>
<td><a href="http://www.compendia.com/Amino">http://www.compendia.com/Amino</a> Acid Library</td>
<td>Amino Acid Library</td>
</tr>
<tr>
<td>Lipid Library</td>
<td><a href="http://www.compendia.com/Lipid">http://www.compendia.com/Lipid</a> Library</td>
<td>Lipid Library</td>
</tr>
</tbody>
</table>

"For a commercial 'jar' database" | | |
| Dictionary of Natural Medicines | Dictionary of natural medicines | Dictionary of Natural Medicines | Medicines |
| Microarray | http://www.microarray.org | Microarray | Genomes |
| MS | http://www.msd.com | MS | Masses |
| SED | http://www.cas-ue/csb/ED/SED/SED.html | SED | Spectroscopy |
| Web MS Library | http://www.compendia.com/WebMSlibrary.com | Web MS Library | Masses |
• b/y ion series commonly used for sequencing common with CID

• **Alternative** activation methods (ETD, ECD) generate c/z ion series
  • Can also be used to sequence peptides
1. c/z Ion formation mechanism


2. c/z Ion formation mechanism

Δ between ion series = Residue mass

Peptide bond fragment ions

Residue Mass
specific to amino acid present in sequence

Finnigan LTQ FT

Linear Ion Trap MS
- MS, MS/MS and MS^n Analysis
- AGC Control
- Secondary Electron Multiplier Detector

FTICR MS
- Ion Image Current Detector
- Accurate Mass, High Resolution
- ECD, IRMPD

FTMS Data
7 T Actively Shielded Superconducting Magnet
ECD Assembly
IRMPD Laser Assembly

Triple Ported Turbo Pump
Neutral Loss triggered ECD

- Precursor m/z added to exclusion list
- Full scan FT-ICR mass spectrum
  - CID MS/MS (IT) of most abundant ion
    - User defined neutral loss observed?
      - No
      - Yes
        - ECD MS/MS (FT) of most abundant ion

Scan event 1
Scan event 2
Scan event 3

Sweet, Anal. Chem. 2006, 78, 7563-7569

1.3 ppm mass accuracy

http://www.iscpubs.com/articles/al/a0401tay.pdf
Neutral Loss triggered ECD

a) Full FT-MS

b) CID (ion trap)

c) ECD FT-MS

Sweet, Anal. Chem. 2006, 78, 7563-7569

Neutral Loss triggered ECD

Sweet, Anal. Chem. 2006, 78, 7563-7569
Neutral Loss triggered ECD

Sweet, Anal. Chem. 2006, 78, 7563-7569