The Proteome of *Arabidopsis thaliana* Analyzed by Electron Transfer Dissociation Utilizing High Mass Accuracy Measurements

Kai Scheffler[1], Martin Hornshaw[2], Martin Zeller[3] and Bernd Müller[4]

[1] Thermo Fisher Scientific, Im Steingrund 4-6, 63303 Dreieich, Germany; [2] Thermo Fisher Scientific, Boundary Way, Hemel Hempstead, HP2 7GE, United Kingdom´; [3] Thermo Fisher Scientific, Hanna-Kunath-Str.11, 28199 Bremen, Germany; [4] Department of Biology, Ludwig- Maximilians-University, Menzingertr. 67, 80638 Munich, Germany

Overview

Purpose: Maximize sequence coverage and protein identification for the analysis of the water-soluble proteome of *Arabidopsis thaliana* using a combination of collisionally induced dissociation (CID) and electron transfer dissociation (ETD) selected via a data-dependent decision tree.

Introduction

While CID MS/MS is a very powerful peptide fragmentation technique, it imposes some limitations on investigators in their efforts to identify proteins and understand their biological function. ETD is a new method of fragmenting peptides that utilizes ion/ion chemistry [1,2]. ETD fragments peptides by transferring an electron from a radical anion to a protonated peptide. This induces fragmentation of the peptide backbone, causing cleavage of the N-Cα bond. Data sets acquired with optimized data-dependent parameter settings for CID and ETD provide complementary information. For the analysis of highly complex samples, the combination of CID and ETD experiments can significantly improve the sequence coverage and the number of identified proteins.

Methods

Leaves from 5-week-old *Arabidopsis thaliana* plants were harvested, frozen in liquid nitrogen, and ground until homogenized. The material was centrifuged and the soluble proteins from the supernatant were concentrated. The soluble fraction was digested with different proteolytic enzymes. Intact chloroplasts from *Arabidopsis thaliana* were isolated according to Aronsson and Jarvis [3]. Intact chloroplasts were lysed and stromal proteins were recovered by ultracentrifugation. Proteins were separated by SDS-PAGE followed by in-gel digestion using different proteolytic enzymes. Peptides were separated by on-line chromatography on a reversed-phase column, followed by mass spectrometric analysis performed on a hybrid linear ion trap / Orbitrap™ mass spectrometer equipped with ETD capability.

LC/MS:
HPLC System: Thermo Scientific Surveyor™ MS pump with a flow splitter
Autosampler: MicroAS
Columns: C18 analytical column (0.1 mm x 10 cm) integrafrit with C18 trapping column 20 mm x 0.1mm
Mobile Phase: A: Water/ 0.1% formic acid;
B: Acetonitrile/ 0.1% formic acid
Flow Rate: 300 nL/min, post split
Mass Spectrometer: Thermo Scientific LTQ Orbitrap XL ETD™

Spray Voltage: 1.6 kV
Capillary Temp: 200 °C
Dynamic Exclusion™: Repeat count 1,
exclusion list size 500
exclusion duration 60 s

Parameters for data dependent decision tree:
z=3 >600 CID, <600 ETD
z=4 >900 CID, <900 ETD
z=5 >950 CID, <950 ETD

Reagent AGC Target: 1E6
Reagent Injection Time: 50 ms
Reaction Time: 100 ms

Data Processing: **Thermo Scientific** Proteome Discoverer software using SEQUEST®.

Results

Due to the complexity of a dynamic proteome, different approaches have to be combined to most effectively measure protein expression and dynamics, stress and developmental responses, post-translational protein modifications, and protein interaction. So far, about 700 proteins from isolated *Arabidopsis thaliana* chloroplasts have been identified by mass spectrometry whereas the predicted number of proteins targeted to the plastid is approximately 3000. Here we present an approach for the data-dependent decision tree which allows for scan-to-scan selection of fragmentation technique based on charge and precursor mass.

FIGURE 1. ETD concept – ion / ion chemistry

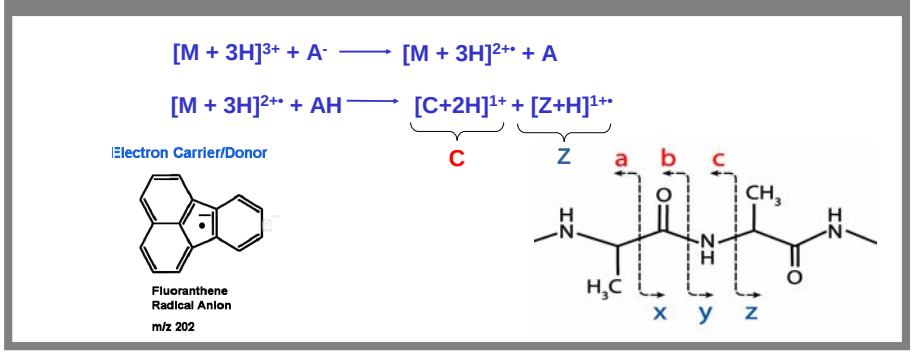


FIGURE 2. Schematic of the ETD (reagent) ion source attached to the LTQ Orbitrap XL ETD

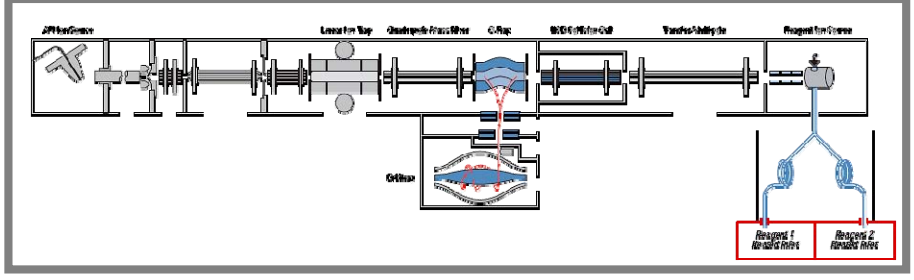
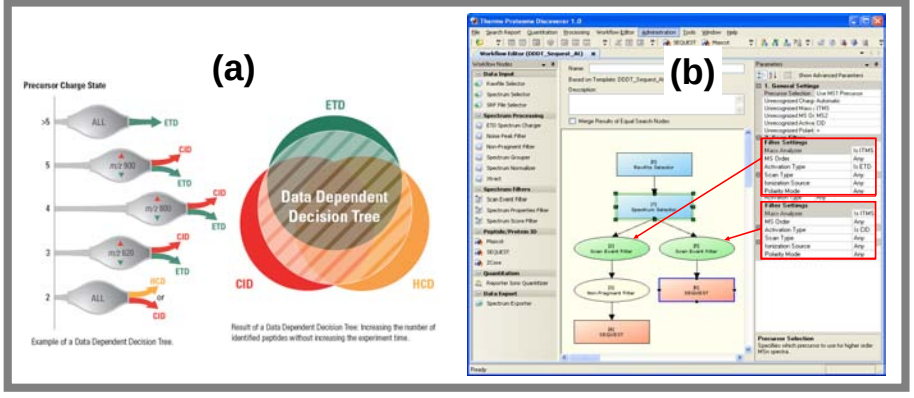


FIGURE 3. (a) Intelligent peptide sequencing using a data-dependent decision tree

(b) Proteome Discoverer data analysis workflow for parallel processing of CID and ETD spectra



The analysis of the whole soluble fraction without separation showed that the total number of identified proteins was not significantly increased using ETD over CID. This was not unexpected since the extreme complexity of the sample analyzed resulted in under sampling during the LC-MS/MS experiment. However, the coverage of most proteins was significantly increased, especially when combining CID and ETD data from different sample digests. Further benefits were achieved by using the multiple dissociation techniques and mass analyzers more efficiently with a data-dependent decision tree that let the instrument decide on the fly which fragmentation technique to use based on the *m/z* of the peptide.

FIGURE 4. (a) Overlay of base peak chromatograms from the CID-only run (I), ETD-only run (II) and mixed run with the data dependent decision tree (III)

(b) Protein ID result of database searches of the three runs and their overlap

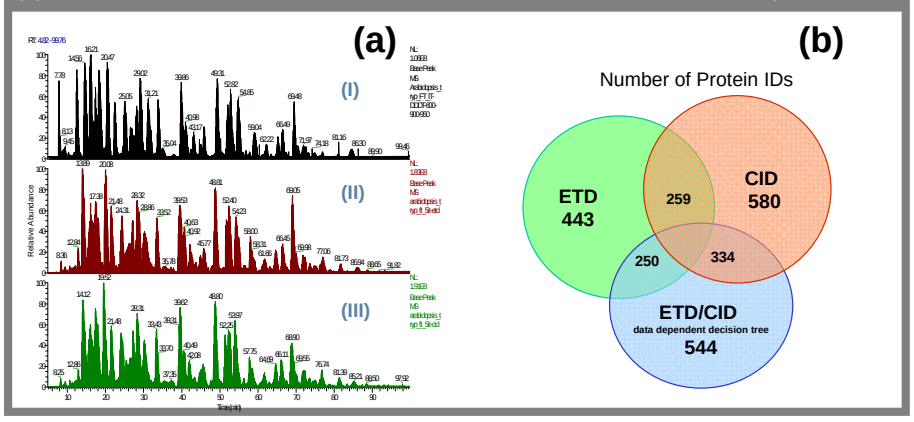
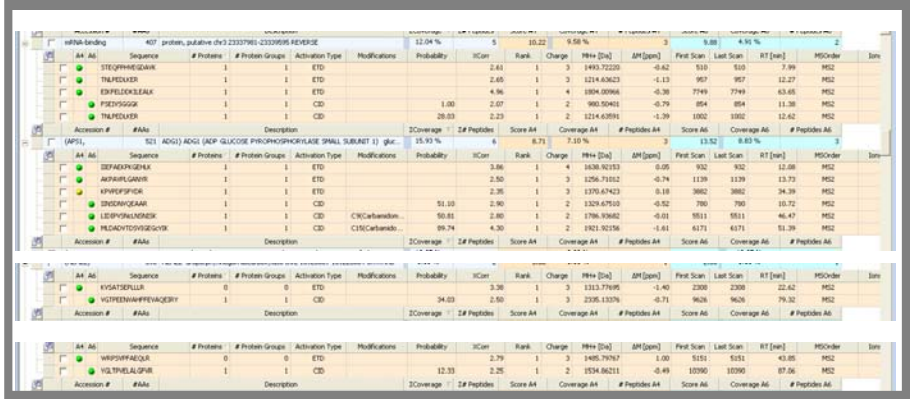


FIGURE 5. Database search result highlighting the complementary protein identification based on CID and ETD data obtained from the experiment using the data-dependent decision tree



Conclusions

- ETD on the LTQ Orbitrap XL ETD allows high-confidence protein identification based on accurate precursor mass.
- ETD has no low-mass cut off and provides for a complete interpretation of the spectrum.
- ETD enables sequencing of larger and highly charged (>$(M+3H)^{3+}$)peptides.
- Peptide sequence information from CID and ETD spectra are complementary.
- Automatic selection of fragmentation type based on charge and precursor mass improves sequence coverage and increases protein ID confidence.
- With ETD, high-quality spectra are obtained for >2+ precursor ions.
- Supplemental activation improves ETD spectral information.
- Digestion with LysC and K/R specific endoproteases yield very similar protein identifications. Other proteases such as GluC or AspN may provide additional benefits for proteome analysis with ETD.

References

1. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Proc Natl Acad Sci U S A. 2004 Jun 29;101(26):9528-33.
2. Coon JJ, Syka JE, Shabanowitz J, Hunt DF. Biotechniques 2005;38:519–521.
3. Aronsson H, Jarvis P. FEBS Letters 2002; 529: 215-220.

SEQUEST is a registered trademark of the University of Washington.
All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.