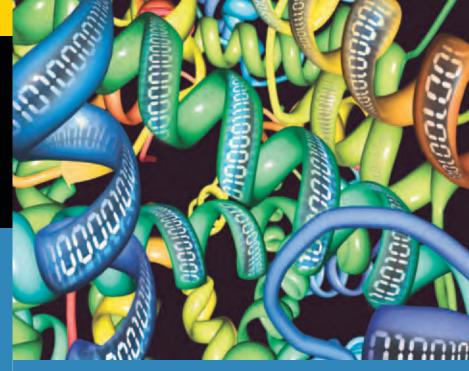
## Thermo Scientific Proteome Discoverer



Mass Informatics Platform for Protein Scientists



### A Comprehensive and Expandable Software Platform for the Analysis of Qualitative and Quantitative Proteomics Data

Outstanding sensitivity and the rapid growth of public databases have made mass spectrometry the primary method of identifying proteins in complex biological samples. Widely available database search algorithms such as SEQUEST® and Mascot<sup>™</sup> facilitate the identification of proteins from mass spectral data. Proteomics research is rapidly expanding beyond simple

Dell

identification, creating the need for more advanced and flexible tools. Thermo Scientific Proteome Discoverer software offers users an expansive range of analytical tools, the flexibility to address multiple research workflows, and an easy-to-use, wizard-driven interface.

#### With Proteome Discoverer:

- Choose the optimum database search algorithm (SEQUEST, Z-Core, Mascot, etc.,) or combine output from multiple algorithms to maximize and cross-validate results
- Merge results from multiple dissociation techniques such as ETD, HCD, and CID to identify more PTMs and increase confidence in those identifications
- Measure and report the relative expression levels of isotopically labeled peptides
- Leverage high-mass-accuracy MALDI data with a Mascot-based peptide mass fingerprinting workflow
- Validate protein IDs using False Discovery Rate (FDR) determination
- Integrate custom tools into the robust, expandable Proteome Discoverer software platform



Proteome Discover

Mass Informatics Platform for Protein Scien

- Use wizards and customizable proteomics workflows to analyze MS<sup>n</sup> data from raw spectra through protein annotation
- Apply GO annotation to illuminate the biological context of the proteins identified
- Take advantage of data standards developed by the HUPO Proteomics Standards Initiative
- Use the SRF file import wizard for a seamless transition from the Thermo Scientific BioWorks platform

# **Proteomics Data**

## Maximum Protein Sequence Coverage and Confidence – Integrating Search Engines and Dissociation Techniques

Proteome Discoverer includes multiple database search engines to complement the breadth of dissociation techniques such as CID, ETD, and HCD available with LTQ<sup>™</sup> linear ion trapbased mass spectrometers. Different dissociation techniques provide complementary results and often increase protein coverage and the number of identified proteins.

- Integrate database search results from industry-leading search engines such as SEQUEST, Z-Core, and Mascot
- Combine information from multiple dissociation techniques
- Utilize advanced tools for database search results from validation through automated false discovery rate calculations
- Exploit the benefits of high-mass-accuracy, high-resolution data

## Use the Z-Core search algorithm to exploit the power of ETD

tists

The novel Z-Core database search algorithm included in Proteome Discoverer specifically takes into account the unique characteristics of spectra generated by electron transfer dissociation (ETD). It includes a data pre-processing step that assigns a charge state to precursor ions. Using Z-Core to search ETD spectra:

- Increases throughput An ETD spectrum need only be searched once to account for all precursor charge states
- Maximizes search efficiency The number of data files searched is reduced
- Increases confidence in protein ID results The number of false positive identifications generated from multiple searches of each ETD spectrum is decreased
- Reduces time spent on data interrogation CID and ETD fragment ion information can be consolidated into one informative report

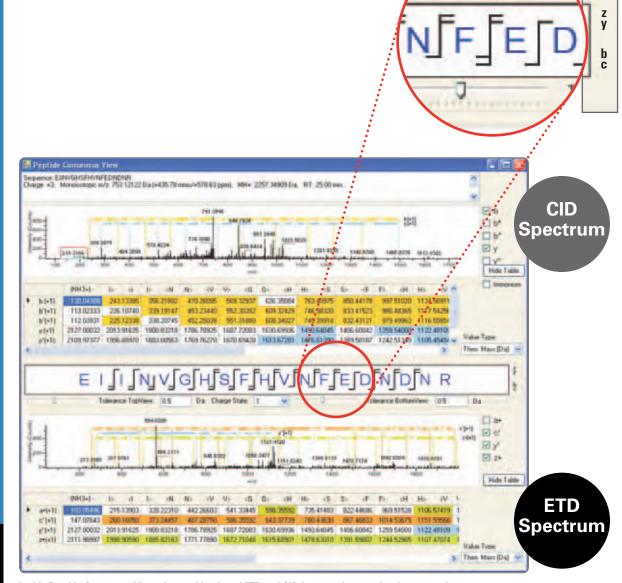
# Analysis Made Easy

3

## Combine CID and ETD data for greater sequence coverage and confidence

Proteome Discoverer includes many interactive results views that facilitate the interpretation of database search results. The Peptide Consensus View highlights the fragment ions in a peptide MS/MS spectrum that match predicted fragment masses. It color-codes the amino acid table to show the portion of the corresponding peptide sequence that was identified.

Multiple ion types (series) can be displayed, accommodating the differences in the predominant ions produced by different dissociation techniques.



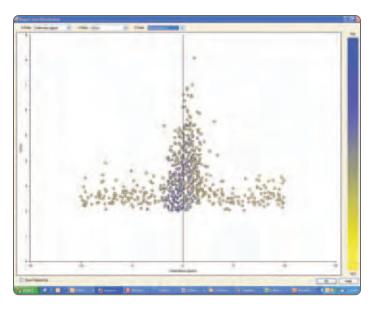
In this Peptide Consensus View, the combination of ETD and CID fragment ion matches increases the sequence coverage of a large triply charged peptide.

## Effortless Results Interpretation – Quickly Identifying Peptide Sequences and Locating Modifications

Proteome Discoverer provides a wealth of tools that make it easy to identify sequences, investigate post-translational and other modifications, and validate search results.

#### Advanced visualization tools

- Merge peptide hits into a comprehensive protein sequence coverage report using the Protein Results View
- Expand and cross-validate PTM information via direct access to the Unimod database
- Organize multiple complex data sets with an easy-to-use results table
- Export graphics for publication using a simple clipboard and common file formats



#### **New plotting capabilities**

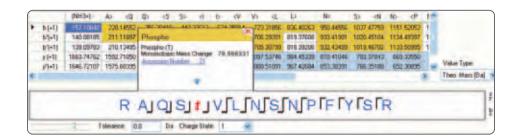
The Distribution Chart shows how closely fragment ion masses identified from a database search match theoretical expected masses, and how well the results correspond to XCorr scores and percentage of matched fragment ions.

#### Validation of MS/MS database searches

The False Discovery Rate (false positive rate) is a statistical value that estimates the number of false positive identifications amongst all identifications found in a database search, and thus is a measure of the certainty of the identification. Inclusion of False Discovery Rates (FDR) is frequently demanded by peer-reviewed publications. Proteome Discoverer includes a Decoy Database Search for determining FDR.

#### More information about PTMs

The Peptide Consensus View clearly identifies modified amino acids with lower case letters. Clicking on the letter displays a pop-up with a description of the modification. Clicking the hyperlink connects to the Unimod website for even more information.



## From Wizards to Workflows – Customizing the Process of MS and MS<sup>n</sup> Data Mining

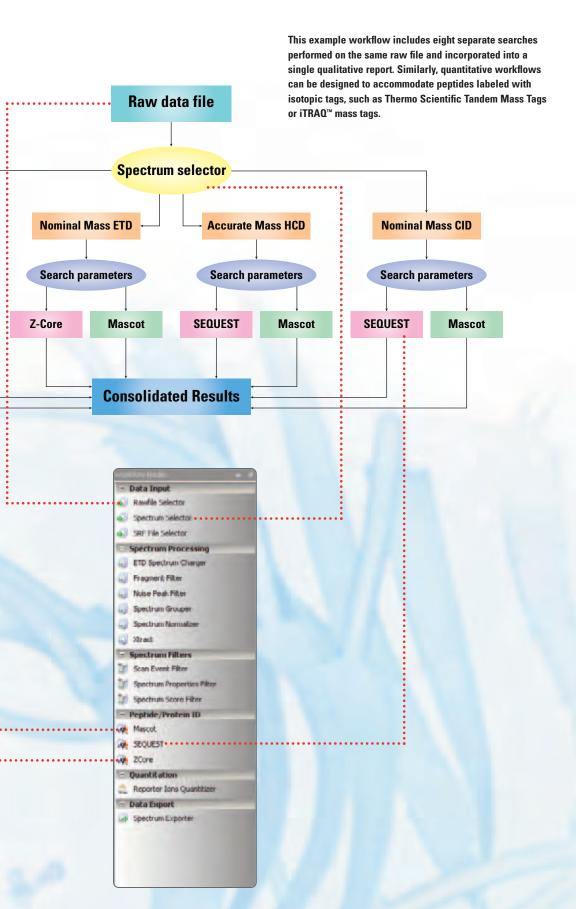
The task of analyzing proteomics data from LC/MS<sup>n</sup> experiments can be daunting, in part because of the numerous and diverse ways to acquire and subsequently interrogate the data. For example, fragment ion spectra from peptides dissociated by ETD have different characteristics than spectra from CID and must be searched differently. Accurate-mass data require different search conditions than nominal-mass data, and both benefit from being searched by complementary algorithms to increase the confidence in the proteins identified. The Proteome Discoverer workflow editor allows users to create a fully customized data analysis workflow from a comprehensive menu of search algorithms, dissociation methods, and results filters. Results can be consolidated into a single, easy-to-read report. **Search parameters** 

**Z-Core** 

Mascot

- Generate workflows for customized data analysis strategies in minutes
- Utilize comprehensive tools for sophisticated raw-data manipulation
- Import data from a variety of sources and output results in several data formats
- Integrate customized tools through the software developer's kit (SDK)



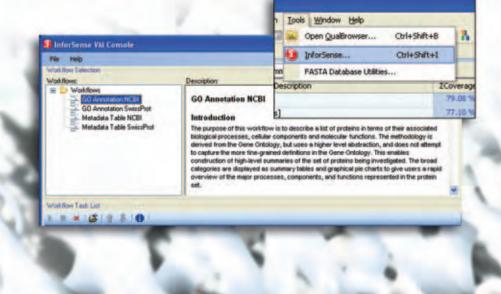


## From Protein Identification to Biological Context – Annotating Results with Information from Public Databases

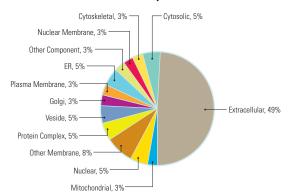
LC/MS analysis of complex protein mixtures such as whole cell digests or digests of biological fluids generally produces high numbers of protein identifications. Insight into biological meaning, however, requires extensive and specific information about each identified protein. Public databases such as NCBI Protein and UniprotKB/Swiss-Prot provide comprehensive descriptions of proteins, but without additional tools, researchers have had no alternative except querying these databases one protein at a time – a laborious and time-consuming process.

Using its integrated InforSense virtual machine (VM), Proteome Discoverer can automatically retrieve pertinent information about each identified protein – including gene ontology (GO) classifications, sites of post-translational modifications, and literature references – from the relevant public databases. Proteome Discoverer users have complete control over the choice of databases to be searched. Annotated results are easily exported to Excel®.

The InforSense virtual machine included in Proteome Discoverer can automatically extract useful metadata from public databases and integrate it with protein identity information.



**GO** Components

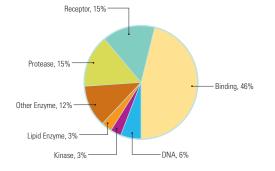


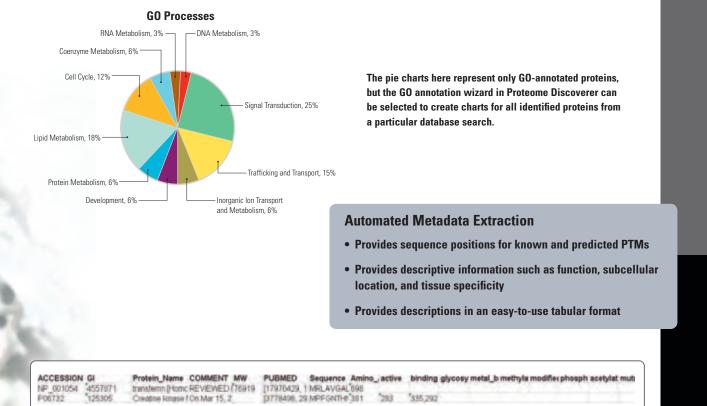
#### Automated GO annotation

- Provides specific biological context about complex protein mixtures
- Provides critical information for additional data mining steps and targeted follow-up experiments
- Produces annotations in both pie chart and tabular form

	the second second	6	0.		1	
Mail out According Number	Other According Numbers	Pratala Batte	60 Camponent	<b>Component Company</b>	60 Function	1-
inertie .	tentre .	and drive to being personners where	Contraction in a present in the state of the	Intel Management		
lan sant	Languard.	Ketterin-3	with of placeta matrix goal	Other Designation	0.000 M 000	-
Lamon P	unaut-	These in some submittee housing (2011)	statute process management	The Inches		
( and the second s		Spectra site care	spin restor & planter matter and	(But However)	and a state of the	-
COMPANY OF THE OWNER OWNE OWNER OWNE	Contract Inc.	Load east 17 being print unt 1	and to be counted and inclusive with	COM STREET, ST.		
Person .	mouth	Want over her I arough the state of	2014 design to prove a part of the local data	Television	Disk datage important right into	1200
(COMMON )		Commission Party State and South State	Tentestane - de las residents	Char Houses	Internet, Survey, print \$19.0	1.00
yora	HARD	And Provide And in concerning the local division of the local divi	I wrongstaling, into dark family and	The local and	manhed annual data indications in	1.8
Luning I		teach-	strangeness relation restorers	the Motore		
watere .	1000	CONTRACTOR OF THE OWNER WATER AND INCOME.	and placed Market Series and	the law of	where the party of the	i an
(PERMIT	10415	Income programs of present Canada present	suggestion and the lot of the lot	the last or	which parameter by printing the second	in the
(in term	(PORTS)	Country protect (Invadidated toxic receptor 1)	warment taken internet	The Name		
prober	affailer .	Company and the state of the same to go the	I wanted which the state of the local of the	Charlenge and	strends in case of the	i in
Peter	10000	Property light	Mini Dag - with the localization of the spectra	(mag)		
and the second	-182.215	Representation of the Installed Installe	activities and of phones survivored	the last as	susants arrists	-
(FRIDE)	TAUNK .	State and Canada Children to And	and the state of the second second	The lines		
and the second s		I dank in the regiment service to provide 2 with	and an other states	Diversion .	secondaries while while	18
PERATE	Helatt	Propagate of generative and \$14 \$14 \$14 \$14 \$14 \$14 \$14 \$14 \$14 \$14	Contracellate regard	Transfer .		
wept's	Autoria	there have by price are one (SP) of	and an able to get	Education		-
instant.	and the	Inclusion party of Contents (See	and a state of the state	1-transfer		
2 increase	Progetting of the local division of the loca	Address of the owner of the	and sampled at region	2.0 million	print and independence result.	
1 James a	1000	Augusturings and one Datum August	a state white regard	Emploide .	country birding	-
		Person Name print the Owner king and taxa	and the state of the state	London .		
(manufil)	- Menalet	date 1 amilutes and present of 140710	and the photoe resident	Distantion .	201000	in.
1 (month)	201708	Automa chief 118 man (M1)	eduction upon	Country .	inter links	
7 January	and and a second s	Plane or Steen 11 and its print or \$17 is	In the later spec	T-Country -	series have a strengthment where	110

**GO** Functions

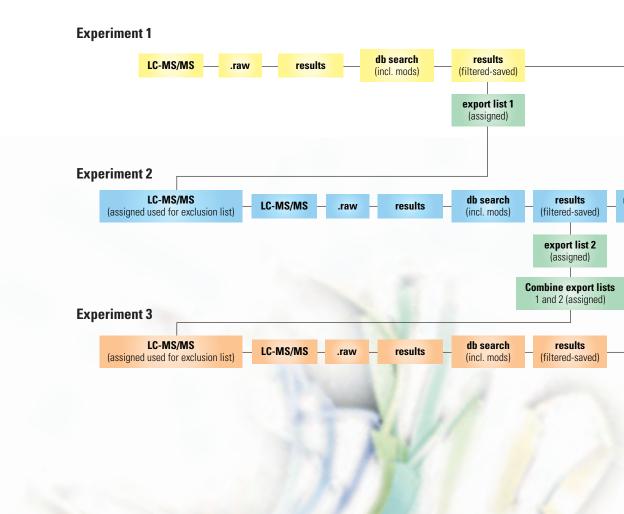




### **Digging Deeper into Each Sample – Combining Data Acquisition** with Creative Data Interrogation Strategies

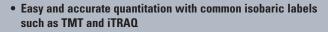
Iterative analysis strategies are a powerful tool for getting the most possible information from each sample. The combination of Proteome Discoverer and Thermo Scientific Xcalibur data acquisition software facilitates these sophisticated analyses.

- Create inclusion and exclusion lists for direct import into Xcalibur™ instrument methods
- Design iterative analysis strategies to detect low-level sample components
- Export results using standard open file formats



### Quantitation – Using Proteome Discoverer to Analyze Isotopically Labeled Peptides

Changes in protein expression levels are often indicative of biological processes such as disease. Isobaric mass tagging has become a common technique for measuring these relative changes. Proteome Discoverer can accommodate several relative quantitation methods, including Tandem Mass Tags<sup>™</sup> (TMT) which allow as many as six different samples to be multiplexed for comparison in a single experiment. The peptides identified and their relative quantitation ratios can be presented in a simple, graphical manner.



- Rigorous statistical data treatment for multiplexed assays
- Customizable filters use only unique peptides to calculate accurate relative protein quantitation
- Normalization of peptide concentration corrects for experimental errors

					126/126	127/126
						000 0.5
American	and taking processing		Concession in the local division of the loca			000 1.0
						000 1.1
And Address of the owner of	Andrew Concerning from the second		-	-	1.4	000 1.0
		Address Name and State	at in	Const Bacally Manh 19	1.0	000 1.0
1		Aller and all and an aller and all and			1.0	000 1.1
	F 13	Landa in in in			14	000 1.0
1	1255 1 400	And antipological supergrid in process Straight		and a second		000 1.1
1	100 1 10	And Spin from the spin start from the spin start of the spin start	12	100		000 1.1
		<ul> <li>1.5. State of the second distance of the last second distance of the last</li></ul>	1.000			000 1.0
· · · ·		All Children in Cases in Case II	-			***
	10.00	Auf Martin Strage game for annual Auf Martin Strage Strage Strage Strage Strage Strage The Development Strage Strage Strategy Strage St	1	111	1.0	000 1.1
1 7	100 1 10	And the second second process in the second Residence of the second seco	100	1.00		000 1.1
	1999	P.S. (2000) France Andreas - Granden Statistics and a second L.D. (2000) France Statistics - Granden Statistics (2000) France Statistics - Granden Statistics - Grande		11 1		
	iii iii	1.5 Million and the Court of American Statements     1.5 Million and the Court from the American Statements	-		1.0	000 1.0
C 100,000		<ul> <li>A. (2007) Manufacture international state of 1712 (2008).</li> <li>A.F. (2008) State of a few state on a property gate</li> <li>A.F. (2008) State of a state of a few state of a gate</li> </ul>	10	19		
C martine	100	The specific descent plants where a first state in the second state is a second state of the second state		100		
		<ul> <li>A. S. Statement of Access of Control</li> <li>A. S. Statement of Access of Control of Contro of Control</li></ul>	1.00	1.000	1.0	000 1.1
a second	100 1 10	of and and and	100		1.0	000 1.0
-	2010/ Names I	Antic Fearmine (PAD) Pall (read)		ti monitori	1.0	000 1.0
Report rati	os of relative c	hange between identified p	eptides.		5.4	000 1.0
				•••	1,0	000 1.0
					14	000 1.0
	(april)	11 11 12 12 12 19 19 19 19 19 19 19 19 19 19 19 19 19				
	1.00	ALL CONTRACTOR			1.1	000 1.1

Zoom in to view all diagnostic ions measured in a 6-plex TMT isobaric tag experiment.

Include Maned out peptide:

Include NOT tilmed out peptide

include selected search input

include of interchinged

Expert archarged mint

Retentors line window indits

retion time limit

1.00 2

0.00 2

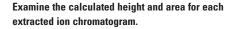
report

(filtered-saved)

C Export m/z values

merge results

(filtered-saved)



#### Laboratory Solutions Backed by Worldwide Service and Support

Tap our expertise throughout the life of your instrument. Thermo Scientific Services extends its support throughout our worldwide network of highly trained and certified engineers who are experts in laboratory technologies and applications. Put our team of experts to work for you in a range of disciplines – from system installation, training and technical support, to complete asset management and regulatory compliance consulting. Improve your productivity and lower the cost of instrument ownership through our product support services. Maximize uptime while eliminating the uncontrollable cost of unplanned maintenance and repairs. When it's time to enhance your system, we also offer certified parts and a range of accessories and consumables suited to your application.

To learn more about our products and comprehensive service offerings, visit us at www.thermo.com.

In addition to these offices, Thermo Fisher Scientific maintains a network of representative organizations throughout the world.

#### Africa

Australia +61 2 8844 9500 • analyze.au@thermo.com Austria +43 1 333 50340 • analyze.at@thermo.com Belgium +32 2 482 30 30 • analyze.be@thermo.com Canada China +86 10 8419 3588 • analyze.cn@thermo.com Denmark +45 70 23 62 60 • analyze.dk@thermo.com **Europe-Other** France +33 1 60 92 48 00 • analyze.fr@thermo.com Germany +49 6103 408 1014 • analyze.de@thermo.com India +91 22 6742 9434 • analyze.in@thermo.com Italy +39 02 950 591 • analyze.it@thermo.com Japan +81 45 453 9100 • analyze.jp@thermo.com Latin America +1 608 276 5659 • analyze.la@thermo.com Middle East +43 1 333 5034 127 • analyze.emea@thermo.com Netherlands South Africa +27 11 570 1840 • analyze.sa@thermo.com Spain +34 914 845 965 • analyze.es@thermo.com Sweden / Norway / Finland Switzerland +41 61 48784 00 • analyze.ch@thermo.com UK +44 1442 233555 • analyze.uk@thermo.com

USA +1 800 532 4752 • analyze.us@thermo.com

www.thermo.com



Thermo Fisher Scientific, San Jose, CA USA is ISO Certified.

©2008 Thermo Fisher Scientific Inc. All rights reserved. SEQUEST is a registered trademark of the University of Washington. InforSense is a trademark of InforSense Ltd. Mascot is a trademark of Matrix Science Ltd. iTRAQ is a trademark of Applera Corporation. Excel is a regis-tered trademark of Microsoft Corporation. All other trademarks are the property of Thermo Fisher Scientific the and its outwickfiering.

Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details.

