Waters[™]

Application Note

Visualization and Comparison of SYNAPT G2-S LC-MS Data with Scaffold Software

Lee A. Gethings

Waters Corporation



For research use only. Not for use in diagnostic procedures.

This is an Application Brief and does not contain a detailed Experimental section.

Abstract

This application brief demonstrates to attain precise large scale label-free analysis of high-density proteomic LC-MS data combined with a qualitative protein strategy that can confidently identify proteins and software to accurately quantify proteins with high reproducibility.

The depth that a complex sample can be interrogated with minimal technical variation is crucial as this characterizes the lowest abundance limit of proteins that can be both identified and quantified. This is of equal importance for the software that analyzes LC-MS data since it ultimately defines protein and peptide false discovery rates. An integrated workflow involving ProteinLynx Global SERVER (PLGS) and Scaffold Software (Proteome Software, Portland, OR, U.S.) provides protein identifications with increased confidence at low false discovery rates.

Benefits

Informatics for visualizing and validating complex peptide data was successfully utilized for the qualitative analysis of *Escherichia coli*.

Introduction

Large scale LC-MS-based discovery experiments investigating increasingly complex proteomic samples are conducted to assess variation, either experimental or biological, profile samples or quantitatively gauge differences in protein abundance. This not only places a demand on the performance of the analytical LC-MS system but also on the bioinformatics software required to analyze the data. Means of visualizing and validating data, using parameters such as probability and confidence scores, allows biological relevance to be assessed for large scale studies. Results from the SYNAPT G2-S for differentially spiked protein standards in a complex cell lysate of *Escherichia coli* are used to demonstrate the application of the computational and statistical tools of Scaffold Software for data independent HDMS^E experiments.

Results and Discussion

A 100-ng *Escherichia coli* sample was injected on-column. The peptides were separated and analyzed using a nanoACQUITY UPLC System coupled with a SYNAPT G2-S operating at a precursor and product ion mass resolution of >20,000 FWHM, and data acquired in LC-HDMS^E scanning mode. Searches were conducted with ProteinLynx Global SERVER v. 3.0 and visualized using Scaffold v. 3.6.

Figure 1 shows the plug-in export tool of ProteinLynx Global SERVER capable of automatically creating one or more Scaffold compatible data files. Following the upload of the search results, protein identifications can be reviewed within Scaffold using the Samples view. Criteria such as minimum protein probability, minimum peptide number, and minimum peptide probability can be useful criteria for filtering identifications. Figure 2 represents *Escherichia coli* identifications that were filtered accordingly, along with associated gene ontology (GO) terms to define biological function of the protein(s). Sequence information can be obtained by reviewing fragmentation data.



Figure 1. Batch-enabled Scaffold Software plug-in export/convertor functionality of ProteinLynx Global SERVER.

		80.0%	 Min # Pept 	ides: 1 Min Pepl	tide: 95%	•] @				
AuliD	Display Options: Protein Identification Probability Req Mod	ds: No F	Filter] Search:		Q				
Load Data Samples	Probability Legend: over 95% 80% to 94% 50% to 79% 20% to 49% 0% to 19%	r Weight	srouping Ambiguity Ny	I regulation mortal process ment of localization on c process c process	Janism process Mar organismal process titue process a to stimulus roduction	m teton Iular region Per organelle ner odrion	e part embrane ant activity ant activity	activity carrier activity r function r transducer activity in molecule activity is molecule activity	non regulator activity B on regulator activity ter activity Tech Brief	
The	Image: Signal Stress Image: Signal Stress Image: Signal Stress Im	151 kDa	Protein Escheri	 biologica cellular p develop establish localizati locomoti metabol 	multi-org multicelli reprodu respons viral rep	 cytoplas cytoskel extraceli intracelita membra mitochoi 	nucleus organell plasma r ribosom antioxid antioxid	 catalytic catalytic electron enzyme molecula molecula structur 	transtrig translati transpoi 5caffold	
Proteins	2 DNA-directed RNA polymerase su RPOC_ECOLI	155 kDa	* Escheri	0		• •		• •	100%	
-	3 7 chaperone protein ClpB OS=EscheCLPB_ECOLI	96 kDa	Escheri				۲	•	100%	
FED	4 🗹 🗠 Chaperone protein DnaK OS=Esch DNAK_ECOLI	69 kDa	Escheri	•	•	•	• •		100%	
	5 V Pyruvate dehydrogenase E1 com ODP1_ECOLI	100 kDa	Escheri	•		•	۲	• •	100%	
for al ar for	6 V Formate acetyltransferase 1 05= PFLB_ECOLI	85 kDa	* Escheri			•			100%	
Similarity	7 V Aldehyde-alcohol dehydrogenase ADHE_ECOLI	96 kDa	Escheri						100%	
	Alanyi-tRNA synthetase 05=Esch SYA_ECOLI	96 KDa	Escheri						100%	
	9 Phosphoridosylformylgiycinamid PUR4_ECULI	102 kDa	Escheri						100%	
	10 Protein transiocase subunit Seca SECA_ECULI	04 kDa	Escheri						100%	
	12 Acomitate invariates 2 03-EscherACON2_ECO	105 kDa	Escheri						100%	
Quantify	12 V 2-0X09Intarate deliverogenase E 0001_ECOLI	103 kDa	+ Ecchari						100%	
quartery	14 Chaparana aratain htsC 05-Esch HTSC ECOLI	71 kDa	Escheri						100%	
E	17 Chaperone protein https://www.hesis.e. 0000.ECOLT	57kDa	Escheri			•			100%	
	15 Manufactional purific biosynthesis p PORS_ECOLI	00 kDa	Escheri						1005	
(22)	10 Annihopepudase N 05-Escherichi CH60 ECOLI	57 kDa	Escheri						100%	
	17 W 00 Kba chaperonin 05-Escherichi Choo_ccoli	104100	Escheri						100%	
Publish	10 Soleucy-trivia synthetase 05-EsSH_ECOLI	80 kDa	Escheri						100%	
	20 Catalase-peroxidase US-EschericKATG_ECULI	07 kDa	Escheri						100%	
A	20 V Transiduon initiation lactor IF-2 0IF2_ECOLI	90 kDa	Escheri						100%	
	21 V Platate synthase G 05-Escherichi PIASZ_ECULI	7740-	Eschert				-		100%	
	22 V Polyribonucleotide nucleotidyitra PNP_ECULI	61 kDa	Escheri						100%	
Statistics	24 W Sush tona protein SI US-Escha EVV1 ECOLT	58 kDa	Escheri						100%	
Statistics	24 V Chysie debudeseesse [deserte CCEP ECOLI	104 10-	Escheri.			-			100%	
	25 Glycine deliverogenase [decarbo GCSP_ECOLI	77 10-	Escheri						100%	
	20 Givernal linease Of a Fasherichia a CLBK ECOLI	F6 kDa	Escheri			-			100%	
	28 Acparted to the synthesize OS-Esciencial C GLPK_ELULI	66 kDa	Escheri						1009	
	29 7 Leurod-tDNA synthetase 05=Esch SVI ECOLI	97 kDa	Escheri						100%	
	30 December of the second seco	99 kD=	+ Escheri						100%	
	31 Charles and the second seco	78 kDa	+ Escheri						100%	
	32 I on protease 05=Escherichia coli 104 ECOLI	87 kDa	Fschari						1004	
	33 V ANDP-dependent malic enzyme 0 MAO2 FCOLT	82 kDa	Escheri			0			100%	
	34 Gutaminyl-tRNA synthetase 05= SV0 FC011	63 kDa	Escheri						100%	
	35 V Tryntonhanase 05=Escherichia c TNAA ECOLT	53 kDa	Escheri						100%	
	Protein Information:	20100	Louiella		Gene Ontol	oov:			1001	
					Gene Onto	·91.				
	Lookup Accession Number In: NCBI (le:gi 1351907, ALBU_BOVIN, P02769)				cellular	process				
	CLP6_ECOLI					cellular metabolic process cellular macromolecule metabolic process cellular protein metabolic process protein modification process				
	Preferred Accession Number: CLPB_ECOLI Protein Name: Chaperone protein ClpB OS=Escherichia coli (strain K12) GN=clpB PE=1 SV=1				•					

Figure 2. Samples view providing a summary of the filtered Escherichia coli protein identifications with GO terms derived from the EBI database.

Figure 3 shows the spectrum of a peptide originating from the CLPB chaperone protein. Full sequence information is provided, thereby providing a high degree of confidence to the identification. Additional retention time information relating to the peptide of interest can be visualized from the low energy data, illustrating other species detected close to the chromatographic apex or (partially) co-eluting with the peptide of interest.



Figure 3. Example of peptide fragmentation spectrum for CLPB chaperone protein peptide GGESVINDQGAEDQR. Additional retention time summary information for other co-eluting peptides, based on the low energy MS1 data, is also provided.

Conclusion

Qualitative analysis of *Escherichia coli* has been demonstrated using a SYNAPT G2-S operating in LC-HDMS ^E mode of acquisition (LC-IM-DIA-MS). Data were processed using ProteinLynx Global SERVER with the results reviewed and visualized using Scaffold Software.

References

1. Li *et al.* Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics*. 2009 Mar;9(6):1696-719.

 Searle BC. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics*. 2010 Mar;1 (6):1265-9.

Featured Products

ACQUITY UPLC M-Class System <https://www.waters.com/134776759> ProteinLynx Global SERVER (PLGS) <https://www.waters.com/513821>

720004599, February 2013

©2019 Waters Corporation. All Rights Reserved.