COMPLEMENTARY MASS SPECTROMETRY OR JUST QUESTIONABLE IDENTIFICATIONS?

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Scott J. Geromanos
Waters Corporation, Milford, MA 01757

OVERVIEW

- **Is it reasonable to assume that the number of identifiable peptides to a protein is directly proportional to its molecular weight and concentration?**
- **Should replicate analyzes of the same sample or similar samples result in the same proteins being identified?**
- **Should the same sample or similar sample analyzed in a different laboratory identify the same proteins?**
- **Is the concentration of certain proteins consistent across different samples?**
- **Should the fragmentation pattern of the same peptide identified from different samples be internally consistent?**
- **Or is absence of replication proof that everything that does not replicate is by definition Complementary?**

INTRODUCTION

It is generally accepted that the reason for poor reproducibility both within and between platforms on similar if not identical samples is that each analysis is complementary. Basically the belief is that data-directed/dependant analyzes are by nature serendipitous and as such it should not be considered uncommon that different proteins are identified from replicate injections on the same instrument or between instruments or across platforms. However, the protein complement, concentration and ionization efficiency of the enzymatically produced peptides of these samples are for all intents and purposes not only constant but similar in like samples. Careful inspection of the DDA switch lists reveals that ~67% of the ions switched on are the same across replicate analyzes, instruments and platforms so why the lack of reproducibility?

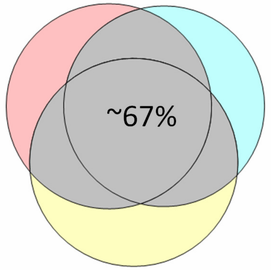
METHODS

All samples were enzymatically digested with trypsin utilizing the protocols described by Dorschel et.al ASMS 2004.

The resulting polypeptide pools were analyzed in triplicate on a QTof Premier and/or Synapt HDMS operating in a data-independent mode (MS^E) of acquisition Both mass spectrometers were interfaced with a NanoAcquity UPLC system. Column diameters, flow rates and gradient conditions varied with each sample.

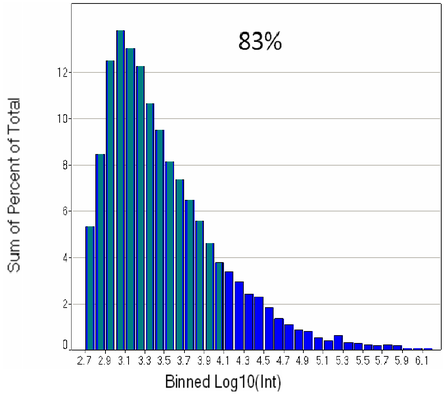
For some of the samples Data Directed/Dependent Analyzes were acquired in triplicate on both the QTof Premier and Synapt HDMS as well as an LTQ, LTQ Orbitrap, QStar Elite and Agilent 6510.

DDA

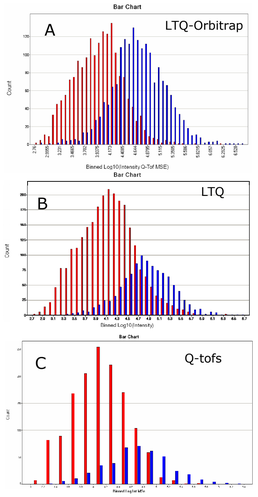


Though the duty cycle for the trapping instruments was much higher than the orthogonal-time-of-flight instruments, ~15K switches per hour of separation compared to 3—4K (instrument dependant). In all cases ~ two-thirds of the switched on ions were the same in all replicate injections.

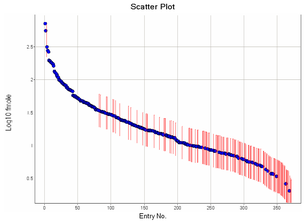
Peptide Ion Intensity Distribution



The above bar chart illustrates the percent of the total ion intensity detected in a given ion intensity bin for a properly loaded column. As is clearly illustrated ~80% of the detected ions are 2 orders of magnitude lower in intensity then the most intense ions with ~68% being 2.5 orders lower.



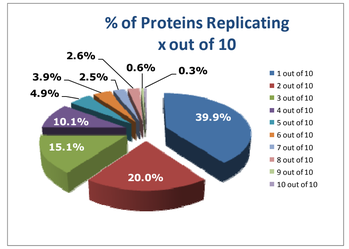
Figures A-C represent the number of matched (blue) & unmatched (red) peptide identifications per log_{10} intensity bin between the same sample analyzed by MS^E/$Identity^E$ on a Synapt MS and the DDA results from data generated on an LTQ-Orbitrap, LTQ and Q-Star searched through Mascot. The data clearly illustrates that the DDA identifications are biased to the higher intensity signal.



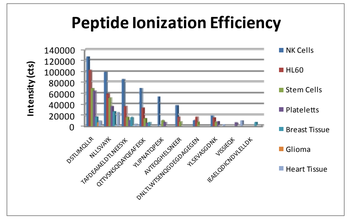
The above scatter plot illustrates the femtomolar concentration of common (blue) and uniquely identified (red) proteins between the Synapt MS data and the DDA Orbitrap data. The data is consistent with the peptide results.. DDA analyzes have problems identifying the lower intensity peptides efficiently as such the do not identify the lower abundant proteins very well either.

	femtomoles of Protein on Column - Peptides Matched												
Accession	Protein Mw	Amniotic Fluid	Breast Cancer	Depleted	Glioma	Heart Cells	HL60 Cells	NK Cells	Non Depleted	Plate-letts	Stem Cells	Urine	White BloodCells
1433B_HUMAN	47137		9-8		12-7	7-4	38-5	32-10		19-10	22-13		28-14
1433E_HUMAN	55209		12-6		21-5	9-5	72-9	152-29		25-9	31-12		25-8
1433F_HUMAN	22110		14-9		12-8	7-3	34-3	11-10		21-11	19-7		
1433G_HUMAN	28776		10-6		15-7	9-5	42-9	18-6		22-12	26-7		
1433T_HUMAN	90178		13-5		16-4	7-3	39-6	21-11		21-7	26-8		19-8
1433Z_HUMAN	41605	9-13	16-5		18-7	7-5	55-10	36-7		60-20	38-10		110-21
A26C4_HUMAN	160797	22-9	42-13		233-6	21-10	215-14	249-9		386-12	288-17		1259-23
A26C8_HUMAN	41774	21-15	40-13		223-5	20-9	206-7	239-7		370-18	276-11		1206-17
A2MG_HUMAN	53520	209-104		215-112		1-4			128-36	4-18			
ACTA_HUMAN	121370	31-9	41-11		316-6	30-17	209-11	247-5		498-11	351-15		2070-18
ACTB_HUMAN	28087	39-20	57-13		414-23	22-11	300-23	444-22		823-43	536-38	10-7	2747-48
ACTB_HUMAN	28171	19-16	35-6		147-3	16-4	141-10	318-4		298-10	269-11	7-12	1753-13
ACTC_HUMAN	27764	32-12	48-10		330-6	31-15	218-10	257-3		520-10	366-17		2160-19
ACTG_HUMAN	29173	41-20	60-13		431-18	23-13	312-21	462-25		857-36	558-34	11-7	2861-51
ACTH_HUMAN	27745	32-9	42-8		330-8	31-13	218-9	257-4		520-10	366-13		2160-16
ACTK_HUMAN	82959		45-4		189-2		158-4	330-7		314-5	228-7		980-8
ACTN1_HUMAN	36557		8-25		17-6	3-7		16-19		55-85	15-23		87-73
ACTN2_HUMAN	32818				13-4	3-19		9-16		39-21	8-22		33-24
ACTN3_HUMAN	28390				13-5	2-7		6-14		28-26	7-27		32-25
ACTN4_HUMAN	50432				16-19	2-8		19-37		48-31	11-22		69-37
ACTS_HUMAN	121363	31-11	46-7		316-7	30-13	209-8	247-4		498-11	351-14		2070-19
ALBU_HUMAN	42003	376-53	350-49			141-13		36-9	995-75	625-62		601-60	69-7
ALDOA_HUMAN	67244		16-22		17-4	32-14	131-23	123-34		17-11	18-15		109-23
ANXA2_HUMAN	28672	9-12	24-16		47-17	5-15		31-20			33-19		17-10
ANXA5_HUMAN	22540		8-7		6-4	2-4	11-8	22-20			30-15		17-9
APOA1_HUMAN	22171	203-43	26-26	264-41		11-10			491-30	19-15			
ATPA_HUMAN	14923		8-18		11-5	27-21		19-14		6-14	21-8		8-15
ATPB_HUMAN	44483		7-19		15-8	32-24		15-15		7-16	21-11		40-28
CALM_HUMAN	49907				12-3		79-8	16-9		15-6	27-7		38-9
CO3_HUMAN	50135	232-147	8-35	112-92		328-162			112-88	6-25			
COF1_HUMAN	18736		11-7		11-4	4-4	45-3	59-14		55-19	28-5		63-14
COF2_HUMAN	32850					4-2	21-5	9-5		19-6	24-5		28-9
EF1A1_HUMAN	49895		17-14		25-2	5-4	103-10	60-17			188-21		
ENO1B_HUMAN	84528		13-10		25-4		154-8	212-8		14-10	40-15		40-20
ENOA_HUMAN	103294		15-23		36-12	3-3	174-22	207-37		21-22	60-30		68-36
ENOB_HUMAN	226401		8-8			5-6	64-6	43-8		12-6	29-10		21-8
ENOG_HUMAN	36507		9-12		7-3		65-4	44-13		10-7	31-16		17-20
ENPL_HUMAN	46855		6-11		17-7	2-5	53-12	31-19		6-25	22-21		32-21
FIBA_HUMAN	49916	28-24	6-10	233-56		2-4				32-36			49-21
G3P_HUMAN	41784		19-10		53-12	31-14	203-11	463-19		51-21	80-32		198-27
GELS_HUMAN	92877	15-16	4-16	26-36					6-4	26-37	8-12	14-7	53-47
GRP78_HUMAN	18371		7-14		27-14	5-3	88-20	89-20		15-17	42-29		31-22
GSTP1_HUMAN	17881				8-6	3-4	15-7	36-13		6-11	24-9		41-12
HEMO_HUMAN	123668	193-26	13-10	149-29		5-9			78-8				
HPT_HUMAN	50151	410-41	17-18			7-12			20-7	7-10		13-13	
HS71L_HUMAN	51769		10-12		26-3	6-3	92-8	119-10		18-11	50-8		12-24
HS90A_HUMAN	38472		9-14		18-4	4-7	146-13	148-32		9-18	78-20		27-18
HS90B_HUMAN	41816		12-15		22-9	3-8	158-18	161-43		11-14	92-28		27-17
HSP71_HUMAN	50140		8-19		17-2	6-8	85-9	74-6		13-18	36-22		13-19
HSP72_HUMAN	28961		9-7		24-4	6-6	111-15	110-15		15-9	53-18		26-23
HSP76_HUMAN	47037		6-7		9-4	5-3	49-8	65-13		11-5	27-14		16-22
HSP77_HUMAN	89190		6-8		9-2	5-2	37-4	44-5			20-11		
HSP7C_HUMAN	39288		12-20		27-13	7-6	134-31	130-51		17-21	86-25		42-29
KAC_HUMAN	21541	176-11	65-16			34-7				20-5		84-8	
KPYM_HUMAN	104854		9-15		68-28	16-20	56-16	163-36		23-30	75-24		89-45
LDHA_HUMAN	184951		5-4		29-10	17-8	73-9	128-22		14-13	22-9		20-17
LDHB_HUMAN	16706		5-6		23-7	42-15	87-13	109-15		18-13	24-13		15-18
MYH9_HUMAN	49959		25-45		12-23			28-67		67-158	47-59		172-108
PGAS_HUMAN	49857				13-6		34-10	21-17		10-22	21-18		31-19
PGAM1_HUMAN	48434				11-4	9-3	38-6	50-8			28-5		27-10
PGAM4_HUMAN	26538				6-2	8-2	33-3	21-3			22-5		20-2
PGK1_HUMAN	49477		6-11		13-8	9-6	69-16	81-42		8-18	14-16		49-26
PPIA_HUMAN	103057		9-9		25-6	7-6	324-16	67-15		27-15	49-14		140-9
PRDX1_HUMAN	49776		18-4		12-6	8-6	43-4	25-11			22-18		
PRDX5_HUMAN	8564				3-2	7-6	15-5	9-4		3-6	8-9		
PRDX6_HUMAN	24903				5-3	6-7	37-4	7-6		14-16	11-6		22-17
PROF1_HUMAN	42016		9-7		16-6	4-3	124-10	72-8		115-19	64-11		279-23
RAB10_HUMAN	75181					4-5	16-4	9-7		12-7	8-6		10-7
RAB1A_HUMAN	22546		3-5		10-3	4-3	24-5	11-3		12-10	11-6		
RAB1B_HUMAN	35922		4-7		12-3	5-3	28-4	13-3		14-8	12-3		9-6
TAGL2_HUMAN	54265		3-9		7-4	2-6		25-11		49-23	18-10		75-26
TBA1A_HUMAN	41661		27-12		56-9	6-7	25-5	213-13		56-15	153-15		180-16
TBA1B_HUMAN	103854		23-12		47-10	5-7	21-5	187-27		49-15	130-14		153-13
TBA1C_HUMAN	41642		26-14		56-11	6-6	25-6	213-17		56-21	153-30		180-15
TBA3C_HUMAN	23224		20-9		48-7	5-7		203-12		55-13	116-12		163-12
TBA3E_HUMAN	83133		17-7		41-7	4-7		173-12		47-10	99-10		87-12
TBA4A_HUMAN	49295		18-9		39-7	4-5		187-14		57-19	83-10		163-27

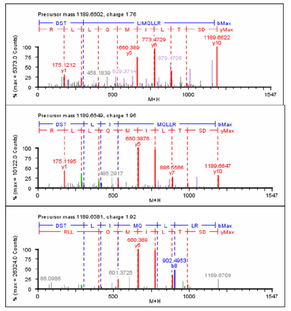
The data presented in the Table is from the combining the protein and peptide identifications from replicate analyzes of 10 different Human samples. For a protein and peptide identification to be valid per sample it must have replicated in at least two out of the three replicate injections. In total 1139 unique proteins were identified to 15,682 unique tryptic peptides comprised of ~162,124 unique product ions.



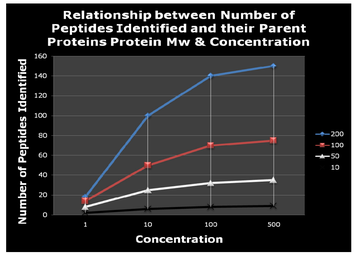
The above pie chart illustrates that ~40% of the protein identifications can be found in ~ one third of all of the analyzed samples with over 60% being identified in any two.



The peptide intensity distribution plot is generated on peptides identified to the protein 1433B_HUMAN 14-3-3 protein beta/alpha Protein kinase C inhibitor. This protein was identified in 7 of the 10 samples analyzed. A brief perusal of the data clearly illustrates a high degree ionization reproducibility of the identified peptides.



Panel A—C depict MSE time-aligned spectra of the peptide sequence DSTLIMQLLR from 1433B_HUMAN. A brief perusal clearly illustrates an exceeding similar fragmentation pattern. As an aside notice how that matched ions are consistent whereby the background ions vary.



The above plot illustrates on average the number of identifiable peptides to a protein of a given molecular weight and concentration based on the data presented in this poster.

	femtomoles on Column (norm. to C3)			
	Amniotic Fluid	Heart Tissue	Non-Depleted	Depleted
CO3_HUMAN	232 - (1)	328 - (1)	112 - (1)	112 - (1)
CO4B_HUMAN	46 - (0.2)	77 - (0.23)	22 - (0.2)	48 - (0.43)
CO4A_HUMAN	50 - (0.21)	83 - (0.25)	24 - (0.22)	52 - (0.47)
CFAB_HUMAN	49 - (0.21)	54 - (0.16)	19 - (0.17)	41 - (0.36)
CFAH_HUMAN	26 - (0.11)	55 - (0.17)	15 - (0.13)	32 - (0.28)
CO5_HUMAN	5 - (0.02)	13 - (0.04)	3 - (0.03)	7 - (0.06)
C4BPA_HUMAN	10 - (0.04)	33 - (0.1)	8 - (0.07)	17 - (0.15)
CO9_HUMAN	8 - (0.04)	13 - (0.04)	7 - (0.06)	9 - (0.08)

The above Table depicts the calculated molar amount of 8 proteins involved in the complement cascade. Interestingly normalizing the femtomolar amounts to C3 in each sample set illustrates a very high degree of similarity. This suggests the ability to determine stoichiometry of protein complexes and pathways through replicate analyzes of like samples. Of course reproducible results are paramount for success.

CONCLUSION

- **The data presented clearly illustrates a reasonable degree of similarity at both the peptide & protein levels between like samples.**
- **The data also illustrates that the best ionizing peptides are consistent to their parent protein regardless of sample.**
- **In addition the results affirm the hypothesis that the number of identifiable peptides to a protein is proportional to its concentration and molecular weight.**
- **Higher concentration protein have more peptides to match.**
- **The ability to reproducibly identify proteins from the same peptides illustrating similar ionization patterns provide the means for determining the absolute quantitation of a protein in a complex mixture.**
- **Absolute quantitation allows for normalization of datasets for comparing results across different laboratories as well as determining stoichiometry of proteins involved in different pathways and complexes.**

50 YEAR LEGACY
VISION ■ CONTRIBUTION ■ ADVANCEMENT

TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS

720002691EN