Asilomar
2005

Qualitative and Quantitative Analysis of *E. coli* by Time-Resolved Mass Spectrometry

Jeffrey C. Silva, Richard Denny, Craig Dorschel, Guo-Zhong Li, Marc V. Gorenstein, Keith Richardson, Daniel B. Wall, Scott J. Geromanos

Waters Corporation, Core Technology/Proteomics, 34 Maple Street, Milford MA 01757

OVERVIEW

Purpose:

?? Use a novel, ***label-free*** LCMS methodology to monitor protein levels in complex protein mixtures.
?? Use *E. coli* as a model system to study proteins influenced by carbon source utilization.
?? Illustrate the ability to simultaneously ***identify and quantify*** proteins among complex samples.

INTRODUCTION

Profiling of expressed microbial proteins has diverse applications ranging from the response of bacteria to external stimuli to the bacterial and fungal response to pharmaceuticals and disinfectants. Separation methods have been integrated with mass spectrometry to reduce the complexity of biological, digested protein mixtures for efficient monitoring of differential protein expression, but these methods can be time consuming, laborious and not amenable to accurate quantification. A robust protein profiling method should provide accurate, rapid and reproducible detection of changes in the level of protein expression or presence of new proteins. The intent of this poster will be to demonstrate a novel, label-free methodology (**Figure 1**) for reproducibly monitoring protein expression levels and illustrate its ability to simultaneously identify proteins/peptides in a complex matrix such as *E. coli*.

METHODS

LC Conditions

?? Waters CapLC HPLC System at 5.0 microliters/min flow rate.
?? Atlantis™ Column (350 ?m X 15 cm, 5 ?m particles).
?? Gradient: 6% to 40% Acetonitrile/0.1% Formic Acid over 100 min.

MS Conditions

?? Prototype QTof Ultima mass spectrometer in V-mode (?12K FWHM)
?? A Nano-LockSpray ion source (~ 5ppm mass accuracy).
?? Alternate scanning acquisition, 1.85 seconds for the low (8 eV) and elevated (28-35 eV) collision energy channels.[1,2]

Sample Preparation

Media and Growth Conditions:

? *E. coli* (ATCC10798, K-12)
? M9 minimal media with either 0.5% glucose, lactose or acetate.

Protein Extract Preparation:

? Cells were suspended in 5mL/1gm biomass in lysis buffer (50 mM Ammonium Bicarbonate, pH 7.5, 1 mM EDTA).
? Cells were sonicated using a Microson XL Ultrasonic Cell Disrupter.
? Debris was removed by centrifugation at 15k rpm for 30 min at 4° C.

Protein Digest Preparation:

? Protein was reduced (10 mM DTT) and alkylated (30 mM iodoacetamide) in presence of 0.05% Rapigest.
? Protein was digested using modified trypsin (Promega) at a concentration of 50:1 (*E. coli* protein to trypsin).
? Tryptic peptide solution was centrifuged at 13,000 rpm for 10 min before LCMS analysis.

Data Processing

The Waters Protein Expression Informatics software is used to:

? ***Process raw data***: Generate an inventory of peptides from each LCMS^E analysis.
? ***Align peptides***: Cluster detected peptides from each sample into a single matrix by accurate mass and retention time.
? ***Perform label-free quantitation***: Normalize all data to peptides of TUFA and determine relative ratios of identical peptides from each condition.
? ***Identify peptides/proteins***: Accurate mass measurement of precursor and associated fragments and/or PMF-search using clustered peptides from the quantitative results.

Alternate Scanning LCMS: LCMS^E

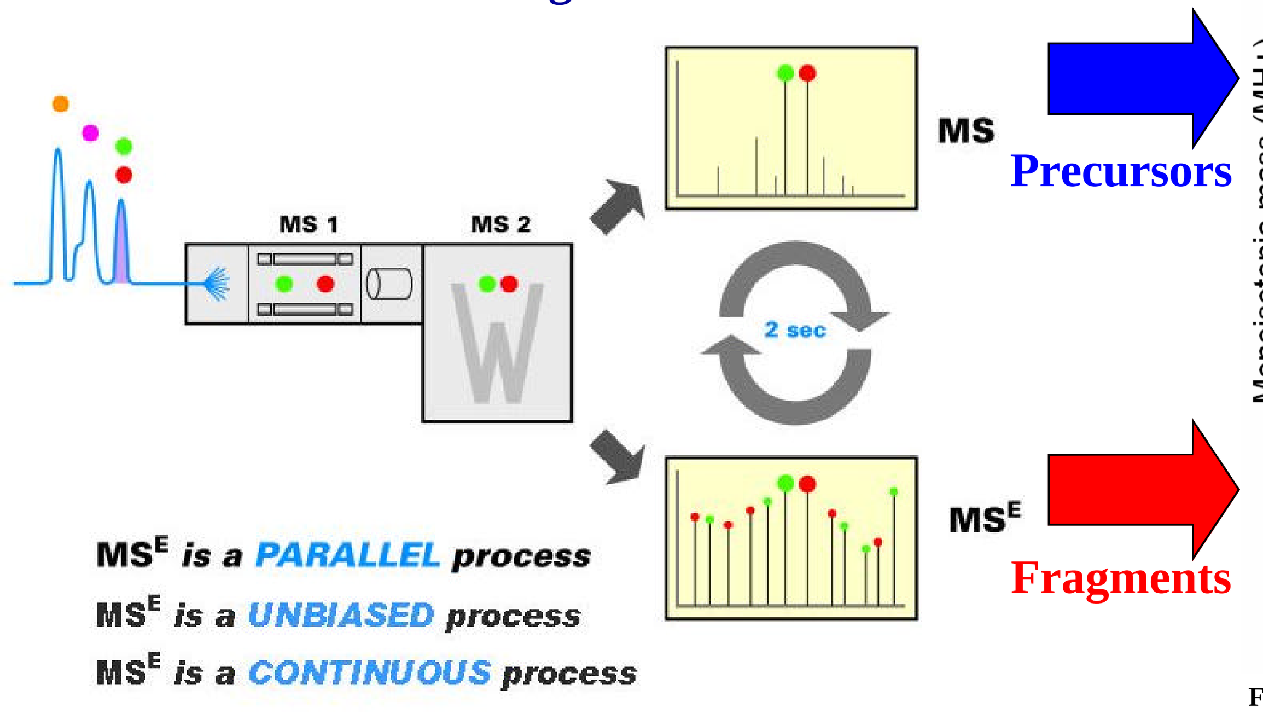
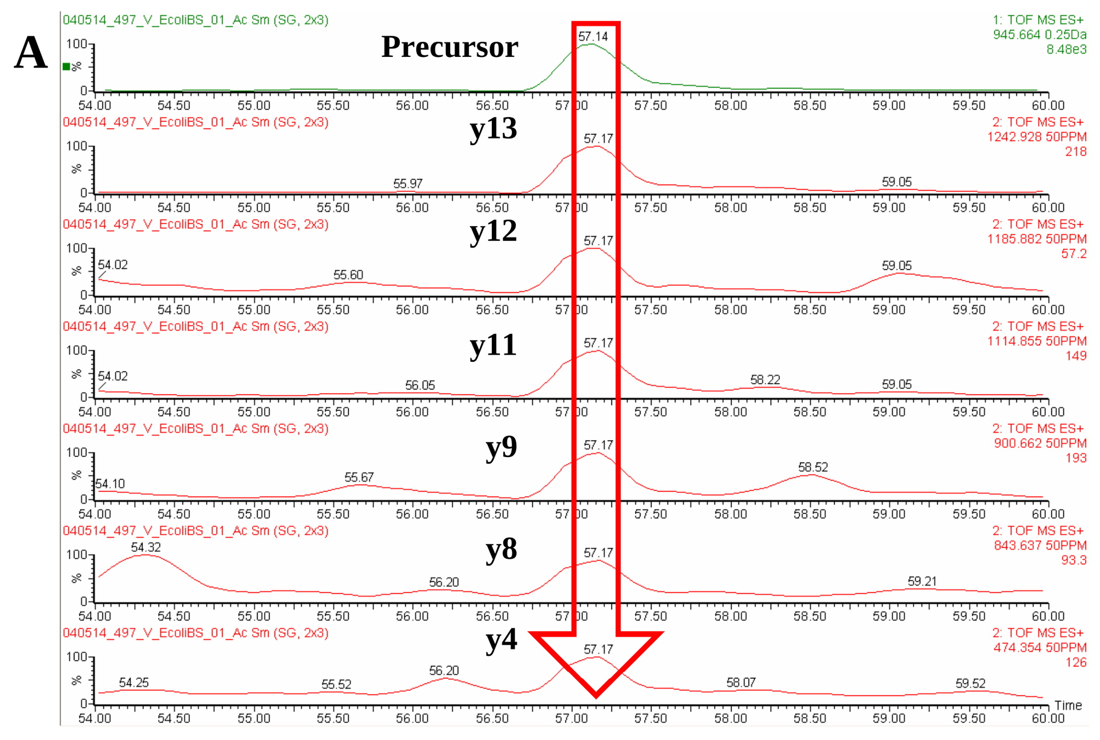


Figure 1. LCMS^E Acquisition. Simultaneous acquisition of precursor and fragment ions for label-free quantitation and identification of peptides and proteins.

RESULTS

Alignment of Precursor and Fragments

A



B

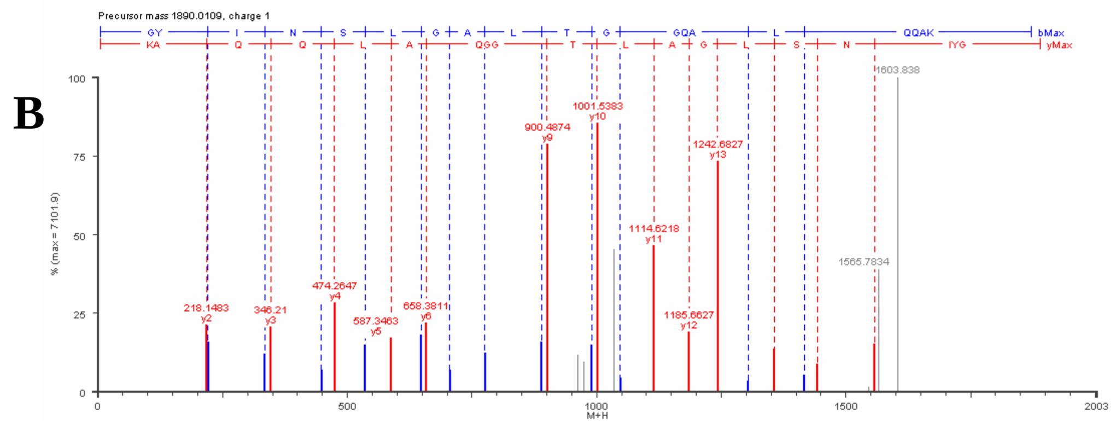


Figure 3. Alignment of Precursor and Associated Fragments. A) An SIC of a doubly-charged precursor peptide of 945.664 *m/z* from the low energy channel (Precursor, function 1) having an apex retention time of 57.14 min. Six SICs from the elevated energy channel (474.354 *m/z*, 843.637 *m/z*, 900.662 *m/z*, 1114.855 *m/z*, 1185.882 *m/z* and 1242.928 *m/z*, function 2) share the same apex retention time of 57.17 min which is within one scan of the highlighted precursor, **B)** The elevated energy spectrum (***time-resolved mass measurements***) associated with the precursor ion described above. An *E. coli* databank search by PLGS of this ion (monoisotopic/lock-mass corrected MH+ = 1890.0109) and its associated fragments identifies this peptide, GYINSLGALTGGQALQQAK, to isocitrate lyase (ACEA). The observed mass was determined to be 1.23 ppm less than the theoretical mass. The six SICs illustrated above correspond to the six corresponding fragment ions to the identified peptide: y4, y8, y9, y11, y12 and y13.

Time-Resolved Mass Measurements of Precursors and Fragments

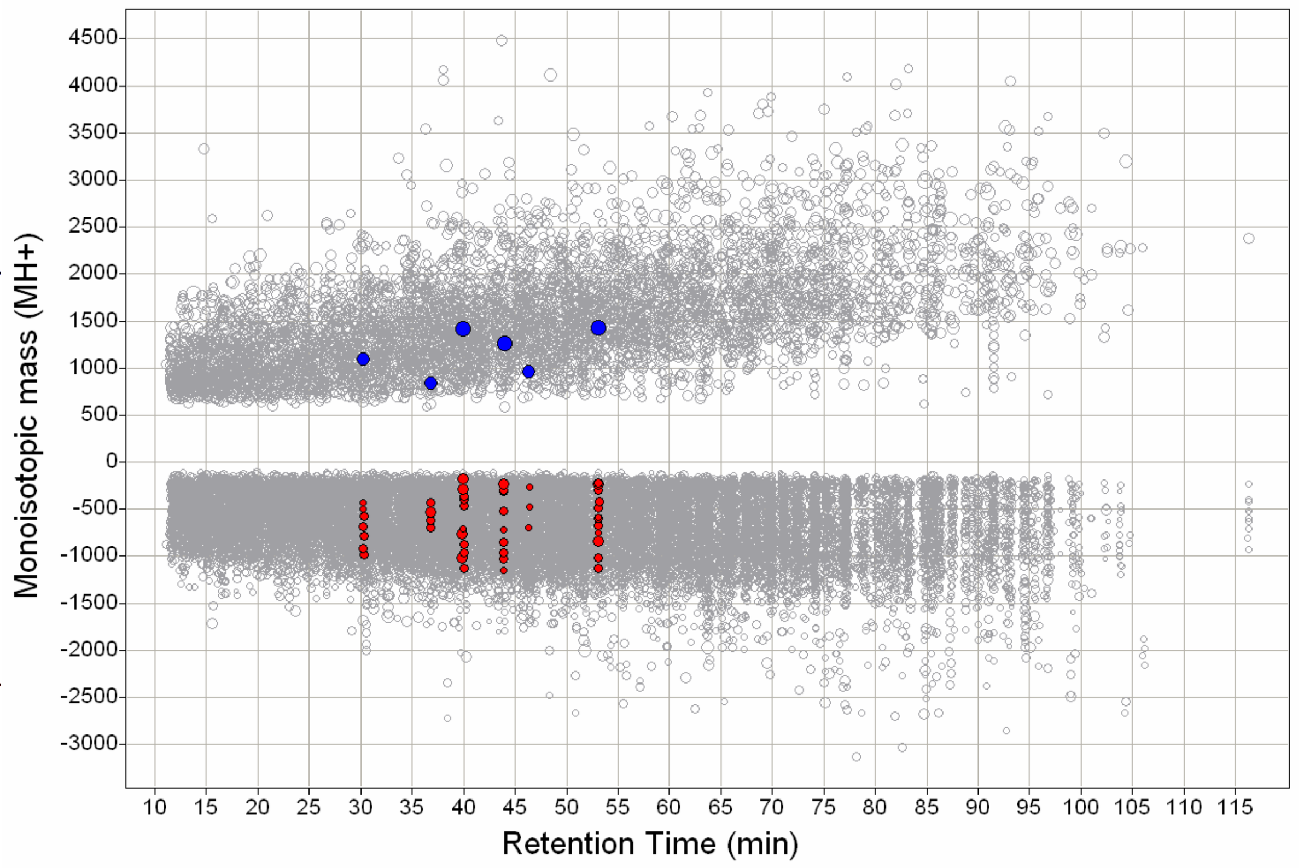


Figure 2. Ion Detections from an LCMS^E Analysis of *E. coli*. A total of 7147 and 25273 accurate mass measurements from the low and elevated energy channels, respectively. The six highlighted precursors (blue) are identified to ribosomal protein RL9 and constitute approximately 45% protein sequence coverage. The fragmentation data obtained from each precursor is highlighted in red.

RESULTS

Peptide Identification by LCMS^E

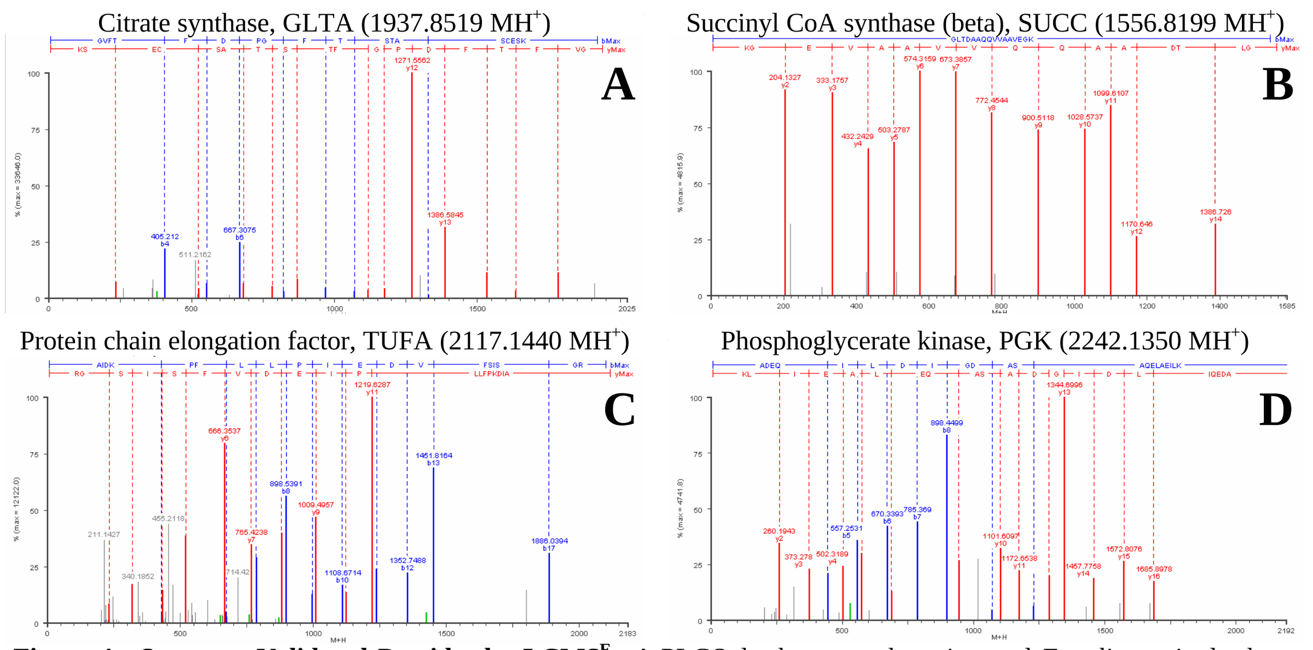


Figure 4. Sequence Validated Peptides by LCMS^E. A PLGS database search against and *E. coli* protein database identified the following peptides to: **A)** citrate synthase (GLTA, GVFTFDPGFTSTASCESK), **B)** succinyl CoA synthase (SUCC, GLTDAAQQVVAAVEGK), **C)** protein elongation factor EF-Tu (TUFA, AIDKPFLLPIEDVFSISGR), **D)** phosphoglycerate kinase (PGK, ADEQILDIGDASAQELAEILK). The mass error was 3.65, 5.49, 9.58 and 3.49 ppm from theoretical, respectively, for each of the corresponding peptides.

Quantitative Comparison of Peptides/Proteins

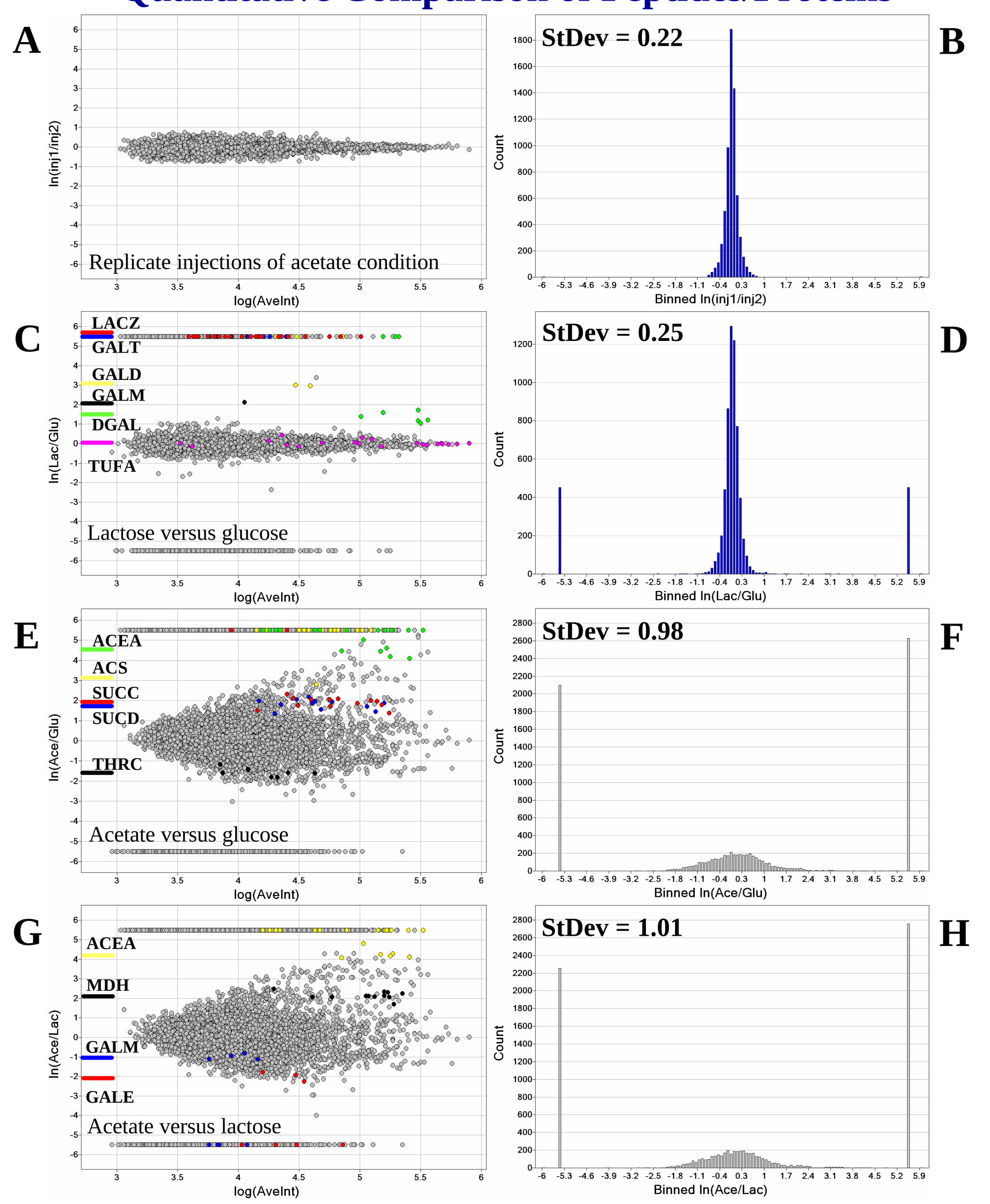

A
B
C
D
E
F
G
H
StDev = 0.22
StDev = 0.25
StDev = 0.98
StDev = 1.01
Replicate injections of acetate condition
LACZ
GALT
GALD
GALM
DGAL
TUFA
Lactose versus glucose
ACEA
ACS
SUCC
SUCD
THRC
Acetate versus glucose
ACEA
MDH
GALM
GALE
Acetate versus lactose
log(AveInt)
Binned ln(Ace/Lac)

Figure 5. Differential Protein Expression. Scatter plots (**A**, **C**, **E** and **G**) and corresponding histograms (**B**, **D**, **F** and **H**) of the natural logarithm of the relative intensity ratios of matched peptides. Multiple peptides corresponding to a subset of identified proteins are highlighted in the scatter plot for each binary comparison. The standard deviation for the ln(intensity ratio) of the matched peptides are indicated in the histogram plot for each binary comparison. **A**) The ln (intensity ratio) of matched peptides between two replicate injections of detected peptides from *E. coli* grown on acetate, **C**) Lactose versus glucose: red-LACZ, blue-GALT, yellow-GALD, black-GALM, green-DGAL and pink-TUFA), **E**) Acetate versus glucose: green-ACEA, red-SUCC, blue-SUCD, yellow-ACS, black-THRC, **G**) Acetate versus lactose: red-GALE, blue-GALM, yellow-ACEA, black-MDH. Unique peptides to each condition is indicated at the extreme values in both the scatter plots and histogram plots (top or left = numerator condition, bottom or right = denominator condition.).

RESULTS

Multiple Peptides for Relative Quantitation

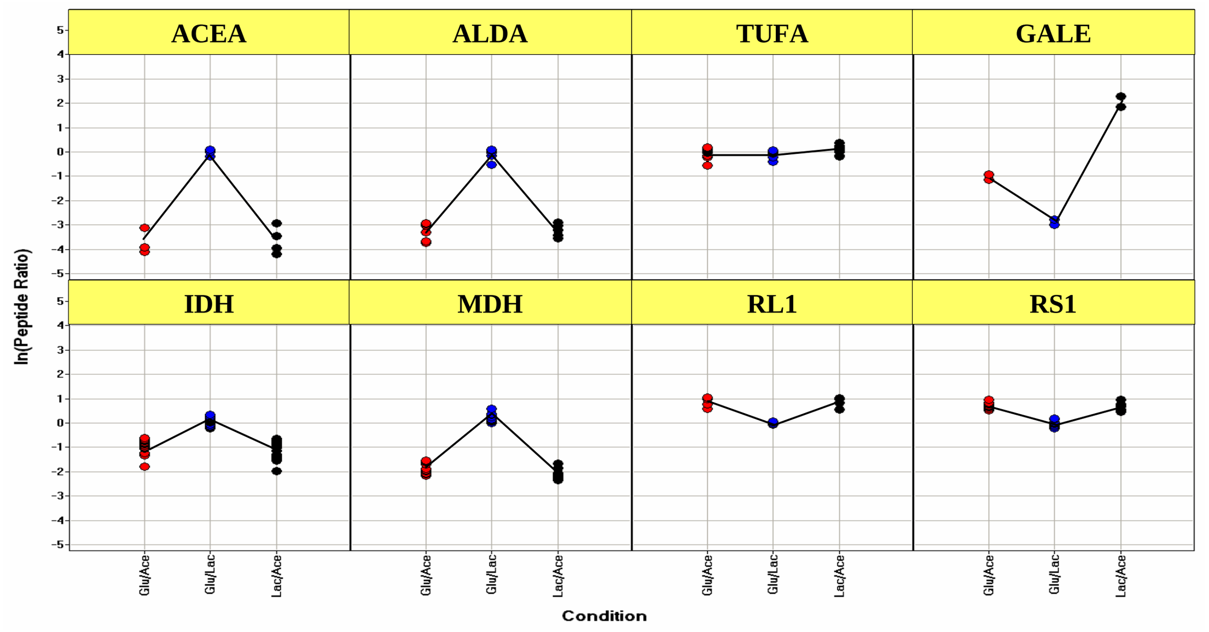


Figure 6. Differential Expression of Peptides. Clustering and quantitative analysis of the LCMS^E data shows that identified peptides in each condition corresponding to differentially expressed proteins have expression ratios within a narrow range. The relative expression of the protein is determined from these multiple peptide ratios, pro-

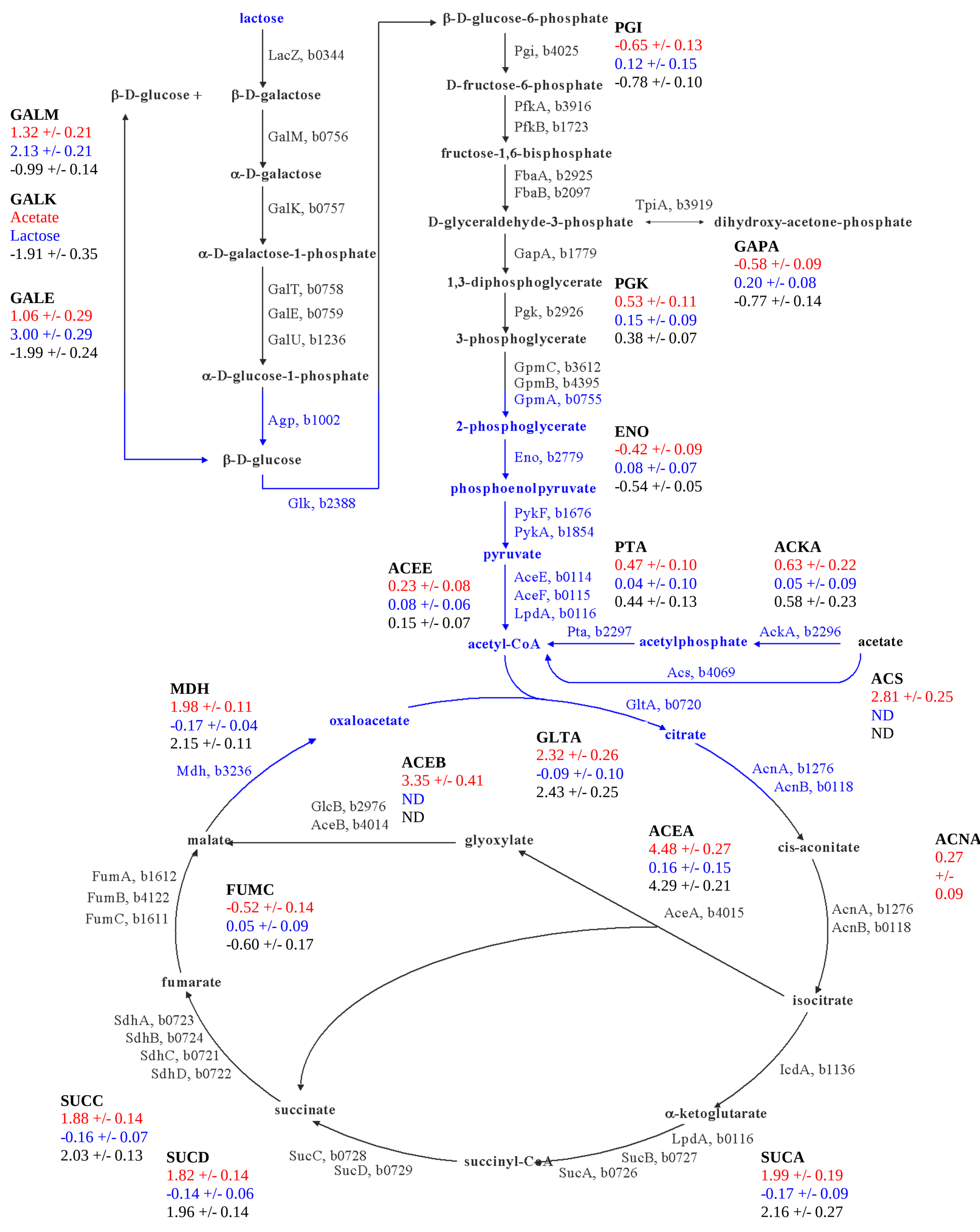

lactose
LacZ, b0344
β-D-glucose +
β-D-galactose
GalM, b0756
α-D-galactose
GalK, b0757
α-D-galactose-1-phosphate
GalT, b0758
GalE, b0759
GalU, b1236
α-D-glucose-1-phosphate
Agp, b1002
β-D-glucose
Glk, b2388
β-D-glucose-6-phosphate
Pgi, b4025
D-fructose-6-phosphate
PfkA, b3916
PfkB, b1723
fructose-1,6-bisphosphate
FbaA, b2925
FbaB, b2097
TpiA, b3919
D-glyceraldehyde-3-phosphate
dihydroxy-acetone-phosphate
GapA, b1779
1,3-diphosphoglycerate
Pgk, b2926
3-phosphoglycerate
GpmC, b3612
GpmB, b4395
GpmA, b0755
2-phosphoglycerate
Eno, b2779
phosphoenolpyruvate
PykF, b1676
PykA, b1854
pyruvate
AceE, b0114
AceF, b0115
LpdA, b0116
Pta, b2297
acetyl-CoA
acetylphosphate
AckA, b2296
acetate
Acs, b4069
GltA, b0720
oxaloacetate
citrate
Mdh, b3236
AcnA, b1276
AcnB, b0118
malate
GlcB, b2976
AceB, b4014
glyoxylate
cis-aconitate
FumA, b1612
FumB, b4122
FumC, b1611
AceA, b4015
AcnA, b1276
AcnB, b0118
fumarate
isocitrate
SdhA, b0723
SdhB, b0724
SdhC, b0721
SdhD, b0722
IcdA, b1136
succinate
α-ketoglutarate
LpdA, b0116
SucC, b0728
SucD, b0729
succinyl-CoA
SucB, b0727
SucA, b0726
PGI
-0.65 +/- 0.13
0.12 +/- 0.15
-0.78 +/- 0.10
GALM
1.32 +/- 0.21
2.13 +/- 0.21
-0.99 +/- 0.14
GALK
Acetate
Lactose
-1.91 +/- 0.35
GALE
1.06 +/- 0.29
3.00 +/- 0.29
-1.99 +/- 0.24
GAPA
-0.58 +/- 0.09
0.20 +/- 0.08
-0.77 +/- 0.14
PGK
0.53 +/- 0.11
0.15 +/- 0.09
0.38 +/- 0.07
ENO
-0.42 +/- 0.09
0.08 +/- 0.07
-0.54 +/- 0.05
ACEE
0.23 +/- 0.08
0.08 +/- 0.06
0.15 +/- 0.07
PTA
0.47 +/- 0.10
0.04 +/- 0.10
0.44 +/- 0.13
ACKA
0.63 +/- 0.22
0.05 +/- 0.09
0.58 +/- 0.23
ACS
2.81 +/- 0.25
ND
ND
MDH
1.98 +/- 0.11
-0.17 +/- 0.04
2.15 +/- 0.11
ACEB
3.35 +/- 0.41
ND
ND
GLTA
2.32 +/- 0.26
-0.09 +/- 0.10
2.43 +/- 0.25
ACEA
4.48 +/- 0.27
0.16 +/- 0.15
4.29 +/- 0.21
ACNA
0.27
+/-
0.09
FUMC
-0.52 +/- 0.14
0.05 +/- 0.09
-0.60 +/- 0.17
SUCC
1.88 +/- 0.14
-0.16 +/- 0.07
2.03 +/- 0.13
SUCD
1.82 +/- 0.14
-0.14 +/- 0.06
1.96 +/- 0.14
SUCA
1.99 +/- 0.19
-0.17 +/- 0.09
2.16 +/- 0.27

Figure 8. Glycolysis, Citric Acid Cycle and Glyoxylate Shunt. The relative quantitation of each protein is color-coded for each binary comparison. The following color-coding scheme is used: Acetate versus Glucose (red), Lactose versus Glucose (blue) and Acetate versus Lactose (black). The relative quantitation is reported as: ln(ratio) +/- 95% Confidence Interval. ND = not determined, Lactose = unique to lactose, Acetate = unique to acetate, Glucose = unique to glucose

RESULTS

Patterns in Chaos

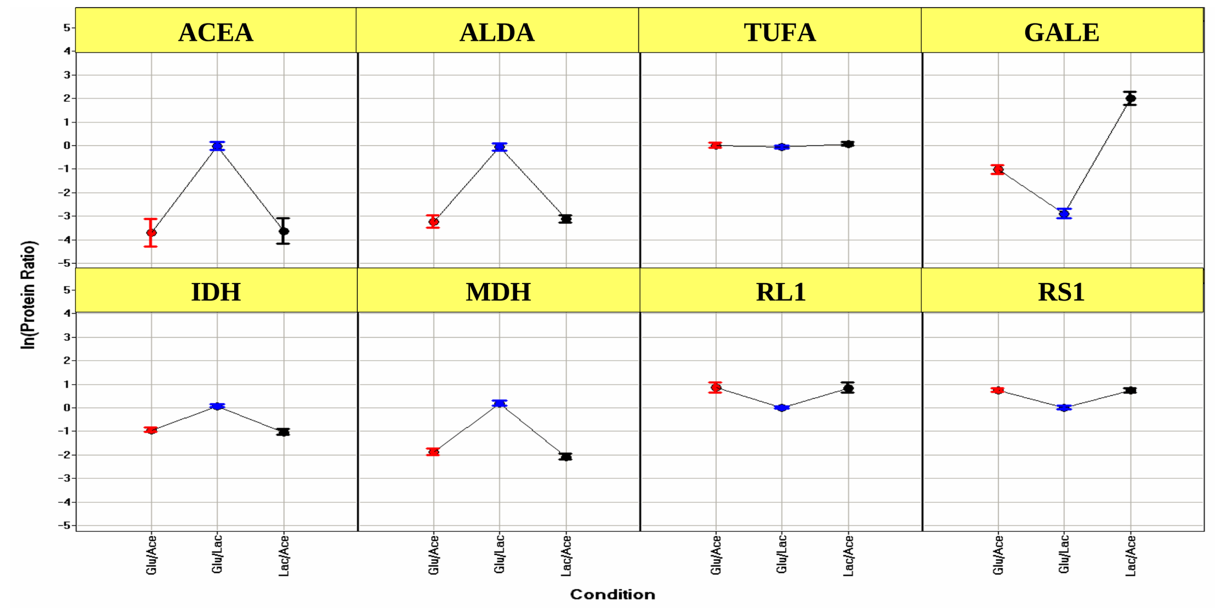


Figure 7. Differential Expression of Proteins. The multiple peptide measurements to each protein provides a means to obtain a 95% confidence interval for the relative expression level for each binary comparison. The pattern obtained from the relative abundance of each protein from each condition provides a mechanism to group related proteins according to their response to the applied perturbation. Ribosomal proteins RL1 and RS1 show a pattern across the various conditions, as do a majority of the other identified ribosomal proteins in this study. Other proteins such as ACEA, ALDA and MDH share a similar pattern which can be explained by their role in carbon utilization. GALE has a unique pattern which can be attributed to its role in lactose catabolism.

***E. coli* on Acetate versus Glucose**

Protein	Blatner	Description	Function	Peptides[A]	ln(ratio)[A]	ln(ratio)[B]	ln(ratio)[C]
ACS	b4068	Acetyl CoA synthase	Acetate Utilization	19	Acetate	2.25	ND
ACEA	b4015	Isocitrate lyase	Acetate Utilization	21	4.48	3.22	2.40
ACKA	b2296	Acetate kinase	Acetate Utilization	8	0.63	-0.67	-0.51
MASY	b4014	Malate synthase A	Acetate Utilization	14	3.35	2.83	2.33
ACO1	b1276	Aconitate hydratase 1	Glycolysis/Citric Acid Cycle	13	0.27	0.41	0.74
ACO2	b0118	Aconitate hydratase 2	Glycolysis/Citric Acid Cycle	22	1.48	1.93	1.19
ALDA	b1415	Aldehyde dehydrogenase A	Glycolysis/Citric Acid Cycle	26	3.25	ND	ND
CISY	b0720	Citrate synthase	Glycolysis/Citric Acid Cycle	15	2.32	1.59	1.50
ENO	b2779	Enolase	Glycolysis/Citric Acid Cycle	21	-0.42	-0.62	ND
F16P	b4232	Fructose-1,6-bisphosphatase	Glycolysis/Citric Acid Cycle	9	1.09	1.19	0.92
G3P1	b1779	Glyceraldehyde 3-phosphate dehydrogenase A	Glycolysis/Citric Acid Cycle	10	-0.58	-0.80	-0.69
G6PI	b4025	Glucose-6-phosphate isomerase	Glycolysis/Citric Acid Cycle	10	-0.65	-0.12	-0.51
GPMA	b0755	2,3-Bisphosphoglycerate-dependent phosphoglycerate mutase	Glycolysis/Citric Acid Cycle	13	0.56	0.53	-0.51
IDH	b1136	Isocitrate dehydrogenase [NADP]	Glycolysis/Citric Acid Cycle	17	1.08	0.59	0.41
KPY1	b1617	Pyruvate kinase I	Glycolysis/Citric Acid Cycle	8	-1.02	-1.51	-1.20
MDH	b3236	Malate dehydrogenase	Glycolysis/Citric Acid Cycle	12	1.98	1.36	1.22
ODO1	b0726	2-oxoglutarate dehydrogenase E1 component, Alpha-ketoglutarate dehydrogenase	Glycolysis/Citric Acid Cycle	26	1.99	0.79	1.48
ODP1	b0114	Pyruvate dehydrogenase E1 component	Glycolysis/Citric Acid Cycle	30	0.23	-1.24	-0.51
PPSA	b1702	Phosphoenolpyruvate synthase	Glycolysis/Citric Acid Cycle	9	1.45	2.48	2.12
SUCC	b0728	Succinyl-CoA synthetase beta chain	Glycolysis/Citric Acid Cycle	17	1.88	1.13	1.48
SUCD	b0729	Succinyl-CoA synthetase alpha chain	Glycolysis/Citric Acid Cycle	12	1.82	1.13	1.34
TPIS	b3919	Triosephosphate isomerase	Glycolysis/Citric Acid Cycle	11	0.42	-0.05	-0.11
DNAK	b0014	Chaperone protein dnaK	Translation Machinery	25	0.34	ND	ND
EFG	b3340	Elongation factor G	Translation Machinery	34	-0.03	ND	ND
EFTS	b0170	Elongation factor Ts	Translation Machinery	13	-0.50	ND	ND
EFTU	b3980	Elongation factor Tu	Translation Machinery	21	0.00	ND	ND
IF3	b1718	Translation initiation factor IF-3	Translation Machinery	8	-0.59	ND	ND
RS1	b0911	30S ribosomal protein S1	Translation Machinery	18	-0.73	ND	ND
RL2	b3317	50S ribosomal protein L2	Translation Machinery	17	-1.12	ND	ND
RL10	b3985	50S ribosomal protein L10 (L8)	Translation Machinery	12	-0.93	ND	ND
RL11	b3983	50S ribosomal protein L11	Translation Machinery	11	-0.88	ND	ND
RL3	b3320	50S ribosomal protein L3	Translation Machinery	8	-0.86	ND	ND
RL1	b3984	50S ribosomal protein L1	Translation Machinery	8	-0.60	ND	ND
RL15	b3301	50S ribosomal protein L15	Translation Machinery	7	-0.89	ND	ND
RS4	b3296	30S ribosomal protein S4	Translation Machinery	7	-0.73	ND	ND
RL13	b3231	50S ribosomal protein L13	Translation Machinery	6	-0.90	ND	ND
RL9	b4203	50S ribosomal protein L9	Translation Machinery	5	-0.81	ND	ND

[A] Protein levels as determined by label-free quantitation from this study.[2]
[B] Transcript levels as determined by M. K. Oh *etal*.[3]
[C] Protein levels as determined by 2-D gel electrophoresis by L. Peng etal.[4]
ND = not determined, Acetate = Unique to acetate condition

Table 1. Comparison of Label-Free Relative Quantitation to 2DGE and Transcriptional Profiling.

CONCLUSIONS

? A parallel LCMS method (LCMS^E) was used to **simultaneously quantify and identify** proteins involved in carbon source utilization.
? The parallel LCMS method provides a means to obtain **time-resolved accurate mass measurements** for precursors and fragments.
? **Protein identifications** from several metabolic pathways of *E. coli* were made using the accurate mass measurements of the detected precursors and their associated fragments.
? A **label-free quantitation method** was used to determine the relative levels of these proteins among the different growth conditions.
? Results are consistent with those obtained by other methods.[3,4]

REFERENCES

[1] R. H. Bateman *etal.* (2002) U. K. Patent 2,385,918A.
[2] J. C. Silva *etal. Anal. Chem.* (2004) **77**, 2187-2220.
[3] M. K. Oh *etal. J. Biol. Chem.* (2002) **277**, 13175-13183.
[4] L. Peng *etal. Appl. Microbiol. Biotechnol.* (2003) **61**, 163-178.

720001406EN