



Proteomic Analysis of K12 and VW *Escherichia coli* Strains Using Two-dimensional Gel Electrophoresis, Maldi-TOF/TOF Mass Spectrometry and Shotgun NanoLC-MS/MS

Kamal F Alblwei*

Department of Chemistry and Biomolecular Science, Macquarie University, Sydney, NSW, Australia

*Corresponding Author: Kamal F Alblwei, Department of Chemistry and Biomolecular Science, Macquarie University, Sydney, NSW, Australia.

Received: April 06, 2021

Published: June 23, 2021

© All rights are reserved by **Kamal F Alblwei**.

Abstract

Escherichia coli is a bacteria that is found in the environment in most warm blooded animals. It is a harmful gram-negative bacteria that has been widely used in research and proteomics owing to its high rate of reproduction. Two strains of *Escherichia coli* were analysed using two different approaches in this experiment, the approaches were 2DE and MALDI TOF/TOF mass spectrometry, and these analysis also included SDS-PAGE with label having 1D shotgun with LC-MS/MS. A quick comparison between the 2DE gels indicated similarities in the molecular weights in the same position or in different positions. A total of 217 proteins were identified using the shotgun method out of which 35 were up regulated, 37 down regulated and 145 remained unchanged. Both the proteins false discovery rate and the peptide false discovery rate for the experiment were zero. It was clear from the results that shotgun proteomics is a more reliable and accurate method when compared to the 2-D method when conducting a comparison in proteomics studies.

Keywords: *Escherichia coli*; VW; MS/MS; nanoLC-MS/MS; MALDI-Tof/Tof; MASCOT Peptide Search; NSAF

Abbreviations

E. coli: *Escherichia coli*; IEF: Isoelectric Focusing; MW: Molecular Weights; BPB: Bromophenol Blue; FDR: False Discovery Rate; NSAF: Normalized Spectral Abundance Factor

Introduction

Escherichia coli

Escherichia coli (*E. coli*) is a bacteria that is found in the environment, foods and intestines of humans and most warm blooded animals. Most strains of *E. coli* are generally harmless while other species however may result in intestinal, respiratory and pneumonitis infections in humans and animals. *E. coli* is widely used in research owing to its physiological characteristics such as convenience of cultivation and high rate of reproduction [1].

Two-dimensional gel electrophoresis

It is a widely used tool used in the analysis of complex protein mixtures extracted either from cells or biological samples. Two-dimensional gel electrophoresis works in two stages owing to the different properties of proteins, these properties are isoelectric focusing and the SDS-polyacrylamide gel electrophoresis. Isoelectric focusing (IEF) separates proteins by taking advantage of the different isoelectric points while SDS-polyacrylamide gel electrophoresis separates proteins by taking advantage of the differing molecular weights (MW) of the proteins [2]. Through these two stages, it is possible to separate a complex mixture containing thousands of proteins thus determining the relative amounts of each protein constituents in the mixture [2].

In order to perform this procedure, a sample under investigation is placed in a gel having a pH gradient; a potential difference is thus applied across the sample resulting in the migration of the protein along the pH gradient until the overall charge becomes zero [3]. The pH of the proteins under investigation determines their location in the gel constituents. Two alternatives can be employed in the creation of the pH gradient they are carrier ampholytes and or immobilized pH gradient gels. Isoelectric focusing is a crucial step in the process of 2-D electrophoresis. In order to get quality data, it is important to attain a low ionic strength before performing IEF analysis. Adjustment of the IEF buffer and electrical profile is key to ensuring that samples that differ in ionic content meet the required low ionic strength in order to get quality data. Slab SDS-PAGE is used in separation of proteins by molecular size, in order to minimize the variations in separation that may occur between individual gels twelve parallel gels were used in separation of the gels at a fixed temperature [4].

Shotgun LC-MS/MS and label free quantitation

Mass spectrometry is an important process in proteomics, it is used to conduct quantitative analysis of protein samples. The emergence of non-gel based proteomics for analysis of expression of proteins has since become available. Isotopes labelled compounds have been used in experimental investigations in shotgun proteomics because of their similarity in terms of their properties when compared to their natural counterparts. It is however important to note that differences still exist in terms of their mass, this has been used as a valuable aspect as a control in mass spectrometry [5].

There are several isotope labels that can be used in shotgun mass spectrometry. The following steps are followed when a label free quantitative method is applied, first the sample is extracted from the source, the extracted sample is then alkylated and enzyme digested using enzymes before being analysed by an appropriate LC-MS/MS or LC/LC-MS/MS. The basis of protein quantification is thus based on two main aspects which are the peptide peak ion intensity and the spectral counting of the proteins identified. A comparison is thus made with different analyses that have already been done [6].

MALDI-TOF/TOF

MALDI-TOF/TOF is a technique used to provide information on molecular weight of compounds in a rapid, accurate and sensitive

manner. Proteins and peptides are just but examples of samples that can be analysed using this technique. MALDI-TOF/TOF is a valuable technique because of its relative high tolerance values towards contaminants. The MALDI-TOF/TOF can be used to analyse molecular weights ranging from 100-150.000 Da, it is however important to note that successful analysis have been conducted for proteins with molecular weights of over 1 kDa [7].

Objectives of the Study

The main objective of this experimental investigation was to identify the main differences that exist between two strains expressed proteins by the use of different proteomic approaches. The 2D gel electrophoresis was done to visually determine the differences that exist between different proteins that were extracted from different strains. Differentially expressed proteins were further identified from the gel by the use of a MALDI-TOF/TOF mass spectrometer. A label free quantitative 1D shotgun was conducted by the use of the SDS-PAGE gel based protein fractionation, this was done in order to come up with a list of differentially expressed proteins that were identified in order to determine the peptide level false discovery rate by the use of the nanoLC-MS/MS linearly on an ion trap mass spectrometer.

Materials and Methods

Escherichia coli protein extracts

A 35 μ L of *E. coli* K12 and VW strain cell lysates prepared protein extract was grown at a temperature of 37°C. The department of Chemistry and Bimolecular science at the university provided the concentration of the proteins approximately at 6 mg/mL.

Two-dimensional gel electrophoresis

A test-tube labelled K12 containing 15 μ l of *E. coli* (concentration 17 mg/ml) protein extracts was provided. To the sample test tube 1 μ L of 10 IPG buffer (ampholytes), 6 μ L of 1 M DTT, 3 μ L of bromophenol blue (BPB) stock and 190 μ L of sample rehydration buffer were added. The mixture was thoroughly mixed by pipetting the samples up and down. The mixture was span in a centrifuge for a duration of 10 minutes, the undissolved particulate were then removed. The IPG strip was then rehydrated in 200 μ L of the sample mixture.

Rehydration was conducted at 50 volts for a duration of 10 hours, this was post ceded by focusing which was done at 200 V for one hour followed by a voltage of 1000 V for the next hour. The

Voltage was further adjusted to 4000 V for a duration of three hours and ultimately increased to 8000 V for a duration of five hours. The voltage was reduced to 500 V for 12 hours on the removal window. The Total Volt hours for the whole exercise was recorded around 48883 Volt Hours.

The IPG strip was then removed from the rehydration tray and re-equilibration done for a duration of 20 minutes using DDT solution which is a mixture of two solutions the 10 mL stock-equilibration buffer and 100 mg dithiothreitol. Further re-equilibration was done for another 20 minutes using a solution of 10 mL stock re-equilibration buffer and 250 mg iodacetamide. The IPG strip was then covered using 1x running buffer.

SDS-PAGE

Into a small tank, a new Bio-Rad criterion precast was placed, the IPG strip was then placed on top of the gel. 0.5% Agarose in a running buffer having traces of bromophenol was used to seal the gel. To the small well, 7 µL of Bio-Rad unstained markers were added. Electrophoresis was started at 60 V and finally increased to 160 V until the BPB line came off the bottom of the gel.

Coomassie staining and imaging

The Colloidal Coomassie staining protocol was used while staining the gel, the Coomassie fix solution was added to the gel and shook for a duration of 60 minutes. The solution was discarded and 40 mL of Coomassie concentrate (containing 10 mL Coomassie diluent) added. Water was used to rinse the gel; the gel was then left on the shaker overnight. The next day the gel was rinsed in water once again and the gel image scanned using a flatbed scanner.

Trypsin in gel digestion

Using the scalpel that was provided, six selected gel spots were cut off from the gel, the cut gel spots were washed in 200 µL of NH₄HCO₃ followed by 200 µL of 50% acetonitrile solution to get rid of the dye. The sampled gels were dehydrated using 50 µL of 100% Acetonitrile solution and then left to dry in the prevailing room temperature. The sampled gels were further reduced using 50 µL of 10 mM NH₄HCO₃ solution, the reduced samples were alkylated with 50 µL of 55 mM iodacetamide in a solution of 50 mM NH₄HCO₃. The samples were then left to incubate in a dark place at the prevailing room temperature for a duration of 25 minutes.

The samples were removed from the dark place after the elapse of the 15 minutes and resuspended in a trypsin mixture containing 12.5 ng/µL of trypsin and 50 mM ammonium bicarbonate. The sample was then rehydrated for a duration of 30 minutes on ice and incubated at 37°C overnight. The sample was made up to a volume of 10 µL by adding 0.1% TFA and *E. coli* extract samples. A Bovine Serum Albumin (BSA) standard protein sample was digested using the same method alongside another BSA control filtered by C18 reverse phase ziptip and used as the internal control experiments for the MALDI-TOF analysis.

MALDI-TOF/TOF

Protein extract samples and BSA controls were spotted onto a MALDI plate after the completion of the digestion of the trypsin using a C18 reversed phase resin-filled ziptip. The following were the characteristics/parameters that were noted for the MALDI-TOF/TOF equipment. The device name was ABI 4800 TOF-TOF, matrix was a-cyano, the precursor acquisition mode was MS and 8 top ions for MS/MS, the device is capable of scanning over 800-4000 mass units and the equipment is able to produce 500 laser shots per spectrum.

Peptide search

Mascot search engine was used to carry out peptide identification, the following parameters were used to conduct the search.

Database	NCBI Non Redundant
Taxonomy	Eukaryotes
Enzyme	Trypsin
Fixed modification	Carbamidomethyl (C)
Variable modification	Oxidation (M)
Instrument	MALDI-TOF/TOF
Peptide tolerance	150 ppm
MS/MS tolerance	150 mmu
Charge state	+1
Mix cleavage tolerance	1

Table

nanoLC-MS/MS (Label free quantitative Shotgun analysis) SDS-PAGE for shotgun experiment

E. coli Cultures for use in the experiment were grown according to the procedure described below. The Streak Plate Starter cultures

of K12 and VW were left to grow overnight on Luria Broth at 37 degrees Celsius while vigorously being shaken. The samples were then sub-cultured in a 2 L conical flask using 500 mL of the fresh media and at a temperature of 37 degrees Celsius. The samples were left in the medium for 24 hours and harvested thereafter for protein extraction.

Harvesting of the *E. coli* from the saturated culture was done by centrifuging the sample in ice at 4 degrees Celsius at a speed of 2500 x g for a duration of 10 minutes. The cell pellets were resuspended in a 1 X Phosphate buffer saline solution having 5 mM DTT, 1 µg/mL of DNase and 1X Protease Inhibitor. In order to achieve cell lysis, four rounds of French press were performed, this was followed by crude lysates centrifugation at a speed of 12,000 x g for a duration of 10 minutes in order to remove cell debris. Methanol Chloroform precipitation was then used to concentrate and purify the extracted proteins.

SDS-PAGE gel analysis

Invitrogen NuPage LDS sample buffer was supplemented with Dithiothreitol (DTT) and used to re-dissolve Aliquots of each *E. coli* protein extract. Approximately 25 µl (60 µg) of each protein sample was loaded onto a 12% linear SDS-PAGE minigel (Biorad), the mixture was electrophoresed for a duration of 90 minutes at a constant voltage of 150 V. Coomassie blue was then used to stain the gel. 16 fractions were cut on each gel lane, the pieces were then finely chopped into smaller pieces. The pieces were then in gel digested using trypsin.

Trypsin in-gel digestion

Using the scalpel that was provided, 16 selected fractions were cut off, the fractions were placed into Eppendorf tube using same the same method as listed on 2.2.3.

LC-MS/MS

Extraction of proteins into a 50 µL of 50% acetonitrile was done after elapse of in-gel digestion, the proteins were then dried down by running the proteins in a vacuum centrifuge. The dried protein samples were then reconstituted in a 10 µL of 1% formic acid, the mixture was then transferred into a 96 well PCR plate (Abgene). A thermo easyLC 1000 nanoflow HPLC system (Thermo, San Jose, CA) started, the system had eight columns of (8 cm x 100 µm I.D) which were prepared as a result of packing 100 Å, 5 µm Zorbax C18

resin at a pressure of 500 psi into columns with integrated electrospray tips made from fused silica, pulled to a 5 µm tip using a laser puller. Using gold electrode and passing through a liquid junction upstream of the column, an electrospray voltage of 1.8 kV was applied. The easyLC1000 auto sampler was used to introduce the samples onto the analytical column. This was followed by elution of the HPLC column eluent directly into the electrospray ionization source of a LTQ-Velos linear ion trap mass spectrometer (Thermo, San Jose, CA). Elution of the peptides was done in a gradient at a flow rate of 400 nL/min using a mixture of buffer A (0.1% formic acid) and buffer B (90% acetonitrile, 0.1% formic acid). Scanning of the spectra was done over the range of 400 - 1500 mass units. Xcalibur software version 2.7 was used for automated peak recognition, dynamic exclusion and daughter ion scanning of the six most intense ions.

Database searching

The readw.exe converter was used to convert raw data files into the mzxml format, the data was then searched against an *E. coli* sequence database which was supplemented with common contaminants using the Xtandem algorithm which was let to run under the GPM-XE interface which was found on the link <http://www.thegpm.org>. The results that was found from the set of 16 gel bands was arranged into a single sheet and exported into an excel table. The outputs that were recorded from the three biological replicates of each of the K12 and VW *E. coli* strains were analysed statistically using the scrappy program. The results of the analysis included the proteins that were identified from each sample, the proteins were then classified as either unchanged or more abundant in the specified samples.

Results and Discussion

Two-dimensional gel electrophoresis comparison between K12 and VW *E. coli* strains

To compare and observe the similarities that exist in protein expression, two 2-D gel electrophoresis images of K12 strain protein extracts were selected. Figure 1A was the gel that our group used while the gel in figure 1B was used by the other group. A quick comparison of the two images indicates a clear protein separation at a satisfactory resolution. The red circles that are named from 1 to 6 are the gel spots that were cut off for trypsin for gel digestion and MS/MS analysis.

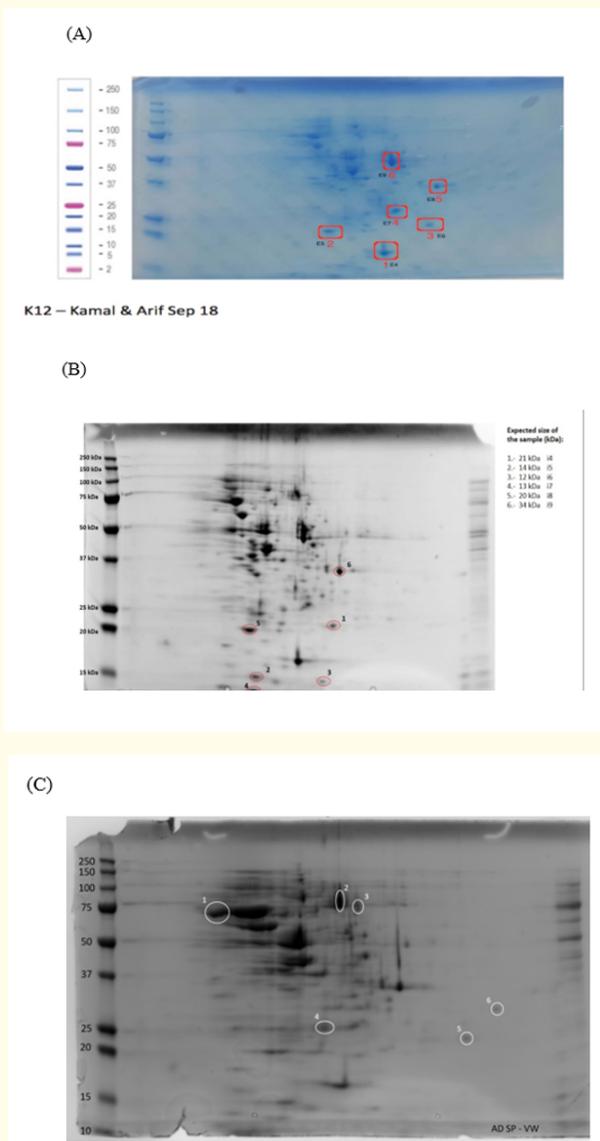


Figure 1: A) The 2-D gel electrophoresis results for our group, it represents two protein strains K12 and VW. MS/MS analysis. B) The 2-D gel electrophoresis results for another group, it represents two protein strains K12 and VW. C) 2-D gel electrophoresis gel image of VW strain. The pH range of the IPG was 3 to 10, the red circles were named from 1 to 6 and represents the gel spots that were cut off for trypsin in-gel digestion and further MS/MS analysis D) 2-D gel electrophoresis gel image of VW strain.

From the two figure 1A and 1B, it is evident that pH range of 6 to 7 and a molecular weight of about 37 kDa contains has numerous proteins compared to any other portion of the gel. A comparison of the two, the results we got as a group versus those garnered by the other group was instrumental in confirming whether correct manipulation of the samples was done and whether this stage of the experiment was a success [8]. The gel spots are as labelled in figure 1C were cut out from trypsin in gel digestion, the gel spots were used for further analysis in the MS/MS analysis

Peptide identification using mascot database

Peptide search was conducted in a Mascot database to detect and identify the differences in the expression between the K12 and VW strains. The gel spots that were cut from the 2-D gel after in-gel digestion with trypsin and let to run on the MALDI -TOF/TOF equipment. The results of the K12 strain extract were recorded in a table 1 below. The results were then compared to the results the other research groups obtained, this was done to determine the similarities and differences between the strains of K12 and VW thus determining the proteins which are characteristic of each other.

Table 2 shows the number of proteins that were analysed, it can be deduced that approximately 6 proteins had 2 or more proteins in common between the two strains.

The MALDI-TOF/TOF analysis employed BSA as the standard control for the experiment, the standard control value is important as it helps in results validation and thus assisting in the recalibration of the equipment which improves the equipment's performance. It is however important to note that none of the BSA controls could be identified in our matrix spots, the absence of the control values was attributed to the technical error from the working group. Possible sources of error in the group could as well been mixed samples, inoculation of wrong/insufficient amount of samples and poor handling of specimen among other errors [9]. The results were however declared satisfactory as the results of other samples were obtained correctly.

Sampling of spots on the 2-D gel was found not to be the optimal way for detecting strain specific proteins according to the results of this experiment. Even though the samples experiments resulted into some data of protein found in a single strain, the method is not

No.	Spot	Protein	Molecular Weight (kDa)	Isoelectric Point	Protein ID
1	E4	nucleoid protein Dps	22,099	5.68	ABE06302.1
2	E5	DNA starvation/protection protein Dps	18,654	5.72	WP_097418011.1
3	E6	Hypothetical protein	45,615	5.16	WP_097433902.1
4	E7	DNA-binding protein, partial	9,206	5.03	PCO95590.1
5	E8	conserved hypothetical protein	26,517	8.89	EFE60772.1
6	E9	tryptophan 2,3-dioxygenase	53,743	5.88	EFK17817.1

Table 1: MASCOT peptide search results for our group.

Proteins	Strain	Search Hits
Nucleoid protein Dps	K12	2
DNA starvation/ protection protein Dps	K12	2
Hypothetical protein	K12	1
DNA-binding protein,	K12	1
Conserved hypothetical protein	K12	1
Tryptophan 2,3-dioxygenase	K12	2
Molecular chaperone DnaK	VW	2
Formate acetyltransferase	VW	1
Formate acetyltransferase	VW	1
NAD(P)H:quinone oxidoreductase	VW	1
Glutaredoxin 2	VW	1
FKBP-type peptidyl-prolyl cis-trans isomerase	VW	1
Chain A, MANGANESE SUPEROXIDE DISMUTASE	K12	1
GNAT family N-acetyltransferase	K12	1
Chain H, 50S ribosomal protein L9	K12	1
Autonomous glycol radical cofactor	K12	1
Superoxide dismutase	K12	1
Alkyl hydroperoxide reductase	K12	1
Glutamate decarboxylase	K12/VW	4
glyceraldehyde-3-phosphate dehydrogenase	K12/VW	3
Chaperonin GroEL	K12/VW	2
Elongation factor Tu	K12	1
Tryptophanase	K12	1
alkyl hydroperoxide reductase subunit C	K12	1
Flagellin (FliC)	VW	1
Glyceraldehyde-3-phosphate dehydrogenase	VW	1
Hypothetical protein AC80_4004	VW	1
Phage protein	VW	1
Phosphatase	VW	1
Hypothetical protein ECEC1864_1415	VW	1
ATP-dependent chaperone protein ClpB	K12	1

Table 2: MASCOT peptide search results for K12 and VW strain protein extract spots cut off from 17 two-dimensional gels.

able to guarantee that the proteins are specific to the strains. Several factors may dictate the quality of results obtained key among these are; the precision with which the gels are cut, the contamination among the samples under investigation, random selection of spots and the resolution of the gel under investigation [10].

The gels were also cut off by different individuals, this means that systematic errors of one person cutting and thus carrying on the same error was avoided. In order to mitigate against such errors and to improve on accuracy of gel electrophoresis, difference gel electrophoresis (DIGE) may be used, in this technique, samples are pre-labeled with two cyanine dyes thus allowing two different samples to be tested on the same gel and ultimately differentiating them using the fluorescence image. The latter allows for better identification of samples and improved comparison between the two images formed [10].

Shotgun proteomics (nano LC-MS/MS) analysis

The figure 2 illustrates the relationship among protein NSAF values of *E. coli* VW and *E. coli* xK12.

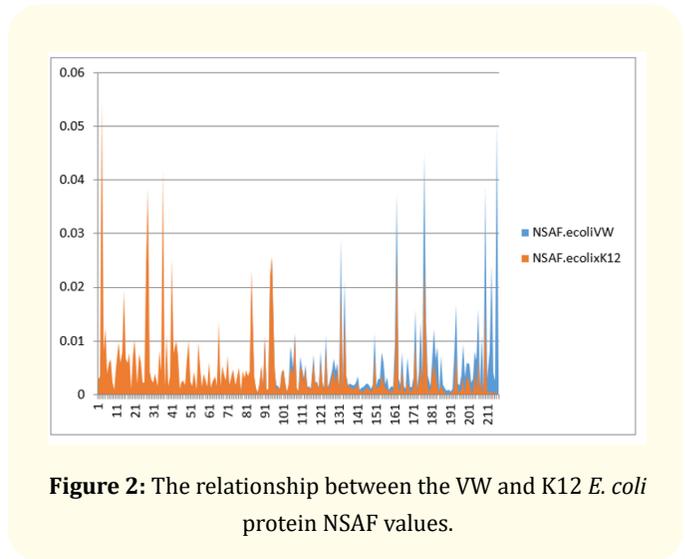


Figure 2: The relationship between the VW and K12 *E. coli* protein NSAF values.

Figure 3A is a plot of VW R1 NSAF and xK12 R1 NSAF values, from the result it is evident that a significant correlation exists between the VW and K12 values. The straight line passing through the cluster of dots is the trendline representing the median NSAF values. A plot of the logarithmic NSAF values of VW and xK12 strains were plotted against each other. The resulting plot signifies the expression of the obtained proteins and the reproducibility of the data that was obtained.

Figure 3B is a clear indication that there is a significant correlation between the VW R2 NSAF and xK12 R2 NSAF values. The straight line passing through the cluster of dots is the trendline representing the median VW R2 NSAF and xK12 R2 NSAF values. A plot of the logarithmic NSAF values of VW and xK12 strains were plotted against each other. The resulting plot signifies the expression of the obtained proteins and the reproducibility of the data that was obtained.

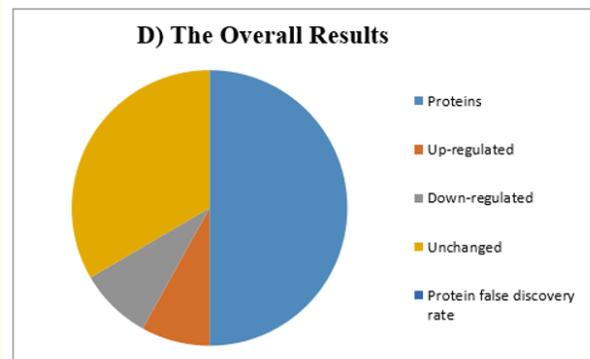


Figure 3: A) The relationship between NSAF values of VW R1 and NSAF values of xK12 R1 B) The relationship between NSAF values of VW R2 and NSAF values of xK12 R2 C) The relationship between NSAF values of VW R3 and NSAF values of xK12 R3 D) The overall results of the shotgun experiment.

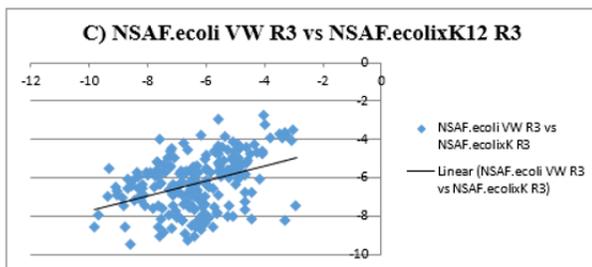
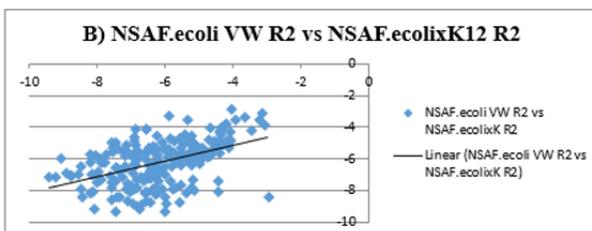
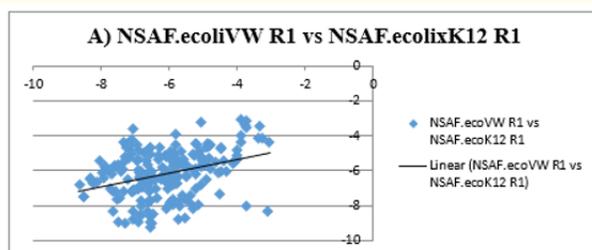


Figure 3C is a clear indication that there is a significant correlation between the VW R3 NSAF and xK12 R3 NSAF values. The straight line passing through the cluster of dots is the trendline representing the median VW R3 NSAF and xK12 R3 NSAF values. A plot of the logarithmic NSAF values of VW and xK12 strains were plotted against each other. The resulting plot signifies the expression of the obtained proteins and the reproducibility of the data that was obtained [11].

Figure 3D is the overall results for the analysis, the pie chart is portioned into 5 different sections representing the five possible outcomes of the analysis. The up-regulated value was represented on the chart with a red colour, whereas the purple colour represented the unchanged value. The protein false discovery rate and also the peptide false discovery rate was zero for the above analysis.

Peptide search for each replicate of K12 and VW strains

The data that was recorded from the mass spectrometer was analysed using Xtandem algorithm which ran on a GPM-XE interface. Three replicates each of K12 and VW were recorded in table 3, the three replicates were averaged for K12 and VW, the average value was 800 and 814 respectively as can be seen in table 3. The

false discovery rate for each replicate were calculated using equation 1 and the results recorded in table 3.

Replicates	K12			VW		
	1	2	3	1	2	3
Protein identified	849	811	742	775	842	825
Protein identifiedd Average	800			814		
Reserved	223	209	186	221	251	0
FDR%	26.24	25.77	25.07	28.52	29.8	0

Table 3: Summarized identified proteins in the 3 replicates of K12 and VW strains and the corresponding calculated FDR % results of the shotgun analysis.

The false discovery rate for each replicate of the *E. coli* strands were calculated from the formula:

$$\text{Protein \% FDR} = \frac{\text{Number of reversed Proteins}}{\text{Number of all proteins}} * 100 \dots\dots\dots \text{Equation 1}$$

The shotgun experiment results indicated that the expressed VW strain were more in the unchanged proteins and was recorded to be 145 as can be seen from table 4. The up-regulated proteins were also recorded to be 35 and were expressed more in the K12 strain with proteins such as Cpn60 chaperonin GroEL. The number of down regulated proteins was recorded to be 37 proteins with

most of them being of the VW strain such as the GTP-binding proteins EFTU_II. The FDR for the unchanged, up regulated and down-regulated proteins were calculated and recorded in table 4. The results of the FDR calculation for the unchanged, up-regulated and down-regulated were found to be 1.38, 0 and 0% respectively. The top ten up/down-regulated proteins were tabulated as identified by the 1D shotgun in table 5 and 6 respectively, the tables also included the statistical t-test values, the ratios and brief descriptions of the proteins.

Normalized spectral abundance (NSAF) was calculated in order to improve the quantitative analysis of the obtained data. The calculated NSAF values were used to determine the relative abundance values and to compare the proteins abundance. The raw data of protein and peptides were filtered to low FDR in order to

	Total of protein identified	Reserved	FDR/%
Unchanged	145	2	1.38
Up-regulated	35	0	0
Down-regulated	37	0	0

Table 4: Summarized unchanged, up and down-regulated proteins of K12 and VW strains resulting from the shotgun analysis and the corresponding calculated FDR%.

	Identifier	statistic.t	ratio	Description
145	gi 16131734	23.37	12.09	Formate dehydrogenase-O, large subunit
85	gi 16129452	16.97	138.99	Glutamate decarboxylase B
21	gi 15829747	13.69	17.95	Lon protease
99	gi 16129913	13.69	10.28	Glyoxalase III and Hsp31 molecular chaperone
22	gi 15832358	12.65	80.09	Outer membrane porin OmpC
172	gi 90111385	12.52	24.84	Fructose-bisphosphate aldolase class I
100	gi 16130088	8.38	8.9	
159	gi 16131968	7.67	2.41	Cpn60 chaperonin GroEL, large subunit of GroESL. groEL (2x), chaperonin GroEL
171	gi 90111100	6.45	6.14	
78	gi 16129202	6.11	4.04	Fused acetaldehyde-CoA dehydrogenase/iron-dependent alcohol dehydrogenase/pyruvate-formate lyase deactivase. AAD_C, C-terminal alcohol dehydrogenase domain of the acetaldehy... PRK13805, bifunctional acetaldehyde-CoA/alcohol dehydrogenaseALDH_F20_ACDH_EutE-like, Coenzyme A acylating aldehyde dehydrogenase (ACDH).

Table 5: Top ten up-regulated proteins.

	Identifier	statistic.t	invratio	Description
211	gi 16130029	-13.19	9.1	galactitol-1-phosphate dehydrogenase, Zn-dependent and NAD(P)-binding
260	gi 16132218	-12.94	7.81	
131	gi 16131218	-12.19	2.93	translation elongation factor EF-Tu 1. EF_Tu, Elongation Factor Tu (EF-Tu) GTP-binding proteinsEFTU_II, Domain II of elongation factor TuPRK00049, elongation factor TuEFTU_III, Domain III of Elongation Factor (EF) Tu
153	gi 16131841	-11.87	2.35	Isocitrate lyase
236	gi 16131188	-9.73	11.81	50S ribosomal subunit protein L24
229	gi 16131121	-9.33	5.16	50S ribosomal subunit protein L13
254	gi 16131813	-7.62	7.28	50S ribosomal subunit protein L11
237	gi 16131194	-7.6	6.63	50S ribosomal subunit protein L22
200	gi 16129178	-7.39	5.93	3-deoxy-D-manno-octulosonate 8-phosphate synthase
213	gi 16130190	-6.83	4.97	UDP-L-Ara4N formyl transferase/UDP-GlcA C-4'-decarboxylase substr

Table 6: Top ten down-regulated proteins.

identify reasonable NSAF measurements. The filtering of the low FDR allows for a statistically high significance between the correlated biological system and the NSAF measurements.

In figure 4, the total proteins NSAF values were plotted, the dark blue circles and open blue circles were used to represent the different protein expressions. The dark blue circles were used to represent proteins which had statistically significant change in protein expression, in this experiment 72 were statistically significant. Skewed circles towards the strain axis represent a more expressed protein. The open light blue circles were 145 proteins, these proteins were statistically insignificant in protein expression. The median NSAF value was determined by getting the best line of fit between the circles (a diagonal line passing through the graph). The log NSAF values for VW were plotted against the K12 NSAF values, the plot is a representation of the expression of the obtained proteins and the reproducibility of the data that was obtained.

The results of the shotgun experiment of the expression of the two strains of VW and K12 of *E. coli* were depicted in figure 5. The up-regulated proteins were more expressed in the K12 *E. coli* strain and was represented in the green coloured region while the VW *E. coli* strain that was predominant in the unchanged proteins were illustrated on the figure in-between up and down regulated protein strains with a black colour.

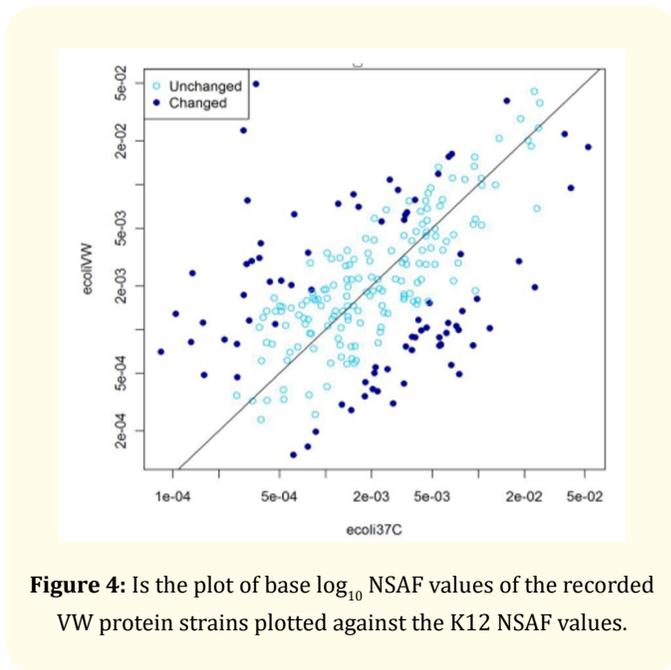


Figure 4: Is the plot of base log₁₀ NSAF values of the recorded VW protein strains plotted against the K12 NSAF values.

Conclusion

Proteomics analysis of the two *Escherichia coli* (*E. coli*) strains were conducted using all the following outlined techniques, to start with the 2-D gel electrophoresis was done, this was quickly followed by a MALDI-TOF/TOF analysis and then the nanoLC-MS/

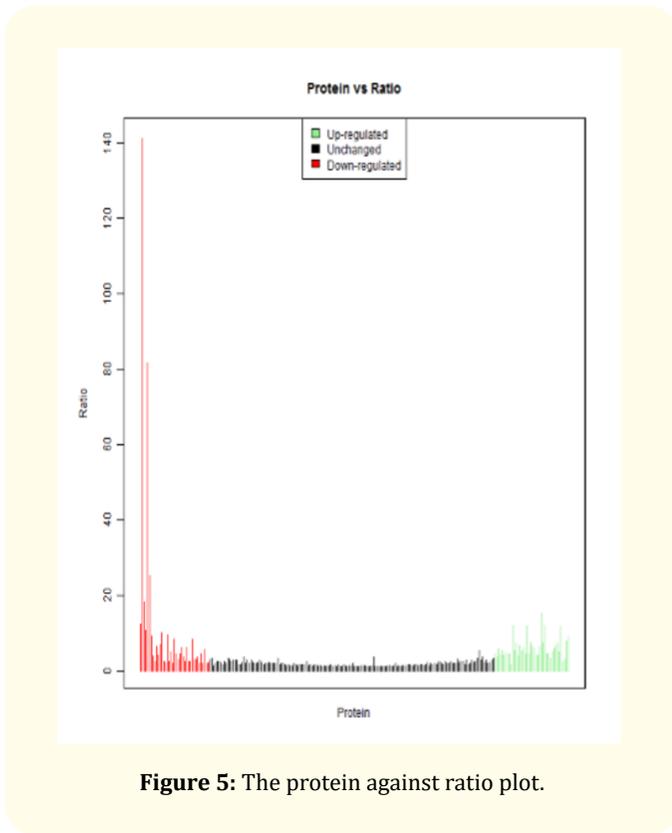


Figure 5: The protein against ratio plot.

MS (shotgun). The experiment focused on comparing and contrasting the two *E. coli* strains K12 and VW in terms of their proteomes.

Using the 2-D gel electrophoresis, it was possible to identify the common proteins between the two strains K12 and VW, the common proteins ranged from two to a maximum of four between some of the strains. The 2-D electrophoresis method is however not an optimal method for identification of proteins owing to its complexity and thus its labour intensity, further the method is prone to error that may be as a result of poor instruments management thus reproducibility of the same results will highly be dependent on the experience of the researcher or laboratory technologist [12]. The dependence of the gel patterns/parameters on the appreciation of the naked eye of the observer results in skewed selection thus mix up of the selected spot especially during the MALDI-TOF/TOF analysis resulting in decreased accuracy.

Shotgun analysis is a more accurate approach, simple and straight forward technique through which many proteins were

identified. The technique was used to determine proteins that were common and or specific to each strain. Information on regulation of the proteins was also easily available through this technique. An average 359 relevant proteins were identified of which 48 were not reliable according to the calculated FDR value [13].

Owing to the limited accuracy of the 2-D gel electrophoresis due to poor manipulation by the operator and risk of incurring error when performing the procedure, the label free quantification shotgun proteomics was preferred for this experimental investigation to deliver the required results with the desired accuracy levels.

Acknowledgement

We thank Prof. Paul Haynes, Thi Huynh, Matt Fitzhenry, Prathiba Ravishankar and Russel Vincent for guide us throughout this study.

Conflict of Interest

There is no conflict of interest according to the author.

Bibliography

1. M Dreyfus and V Heurgué-Hamard. "Termination troubles in *Escherichia coli* K12". *Molecular Microbiology* 79.2 (2017): 288-291.
2. L ZIMMERMAN. "Colonial Differentiation between *Escherichia coli* B and *Escherichia coli* B/r". *Nature* 180.4596 (2017): 1217-1218.
3. J Meier-Kolthoff. "Welcher Coli ist ein Coli Sequenzierung des Typstamms von *Escherichia coli*". *BIOspektrum* 21.4 (2015): 455-455.
4. D Carroll, et al. "Simultaneous Quantitation of Oxidized and Reduced Glutathione via LC-MS/MS in Hematopoietic Tissues". *Free Radical Biology and Medicine* 76.3 (2016): 122-138.
5. M Pang, et al. "Label-free LC-MS/MS shotgun proteomics to investigate the anti-inflammatory effect of rCC16". *Molecular Medicine Reports* 14.5 (2016): 4496-4504.
6. S Jin, et al. "The Effects of Shared Peptides on Protein Quantitation in Label-Free Proteomics by LC/MS/MS". *Journal of Proteome Research* 1.2 (2016): 164-169.
7. Y Miyadera, et al. "Simple LC-MS/MS Methods Using Core-Shell ODS Microparticulate for the Quantitation of Total and

- Free Daptomycin in Human Plasma". *Therapeutic Drug Monitoring* 1.2 (2018): 39-46.
8. S Zhang, *et al.* "Quantification of N-glycosylation site occupancy status based on labeling/label-free strategies with LC-MS/MS". *Talanta* 170.4 (2017): 509-513.
 9. K Bluemlein and M Ralser. "Monitoring protein expression in whole-cell extracts by targeted label- and standard-free LC-MS/MS". *Nature Protocols* 6.6 (2017): 859-869.
 10. C Christin, *et al.* "Data processing pipelines for comprehensive profiling of proteomics samples by label-free LC-MS for biomarker discovery". *Talanta* 83.4 (2016): 1209-1224.
 11. Y Benjamini. "Discovering the false discovery rate". *Journal of the Royal Statistical Society: Series B Statistical Methodology* 72.4 (2016): 405-416.
 12. K Liang and D Nettleton. "Adaptive and dynamic adaptive procedures for false discovery rate control and estimation". *Journal of the Royal Statistical Society: Series B Statistical Methodology* 74.1 (2016): 163-182.
 13. J Habiger. "Adaptive False Discovery Rate Control for Heterogeneous Data". *Statistica Sinica* 37.1 (2018): 140-178.

Volume 5 Issue 7 July 2021

© All rights are reserved by Kamal F Alblwei, *et al.*