A Proteomic Study of *Pseudomonas putida* by Two-Dimensional Gel Electrophoresis: Establishing Quantitative Standards for Intra-Laboratory Results

Kelly Fowlkes

September 2005

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Chemistry.

Approved: _____
Thesis Advisor

_____
Department Head

Department of Chemistry
Rochester Institute of Technology
Rochester, NY 14623-6503

A Proteomic Study of *Pseudomonas putida* by Two-dimensional Gel Electrophoresis: Establishing quantitative standards for intra-laboratory results

I, Kelly Fowlkes, hereby grant permission to the RIT library of the Rochester Institute of Technology to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Date: _____

Signature: _____

Table of Contents

Page

Acknowledgements…… i

Common Abbreviations List…….. ii

List of Figures…… iii

List of Tables……. iv

Abstract…….. 1

Introduction…… 2 – 10

I. Proteomics ……. 2
II. The Basis of Separation in Two-dimensional Gel Electrophoresis…… 4
III. Challenges of Two-dimensional Gel Electrophoresis……7
VI. Specific Aims.……8

Methods and Materials……...…….. 11 – 26

I. Growth conditions…… 11
II. Centrifugation…..…… 16
III. Sonication…….. 16
IV. Determining Protein Concentration……17
V. Treatment with RNAse……17
VI. Rehydration of IPG strips……17
VII. Isoelectric Focusing (IEF)……. 18
VIII. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis.…… 19
IX. Staining the gel……...……21
X. Image Analysis.……22
XI. Analysis with Phoretix 2D Software……...22

Results and Discussion……..… 27 – 52

Conclusion …….. 53 – 64

I. Development of a Standard Operating Procedure…….. 53
II. Reproducibly Determining Mid-log Phase…… 54
III. Changes in Protein Expression Throughout the Log Phase of Growth…. 55
IV. Development of the Universal Standard Gel…….. 58
V. Future Directions …… 59

References ……..... 65 – 67

Glossary……... 68 - 71

Acknowledgements

I would like to take this opportunity to thank the people who made this work possible. To my advisors, Dr. Laura Tubbs and Dr. Paul Craig, thank you for providing me with the insight, guidance, enthusiasm and motivation to complete this project. More importantly, thank you for your never ending support and friendship over the last two year. To Dr. Paul Rosenberg and Dr. Terence Morrill, thank you for believing in me, and in this work, and for driving me to become a better student. Next, I would also like to thank the Department of Chemistry for funding my work. I would also like to thank my colleagues – Maryed Pratts, Nakesha Smith, Nikole Greeson, Archana Pandey, Emmanuel Onyeka, Rachel Pleunther, and Julia Kohn. Each and every one of you has greatly contributed to this project, made me smile every day, and will be greatly missed after I leave RIT.

Lastly, I would like to thank my parents, sisters, and fiancé. Mom and Choch thank you for providing me with the love and support that made me the person I am today, and thank you for the sacrifices you have made that allowed me to excel in all aspects of my life. Bud and Megs, thank you for the interest you have shown in this project. Jess, thank you for all that you do to support me on an everyday basis. Without your constant encouragement, this project would not have been completed.

Common Abbreviations

2DE – Two Dimensional Gel Electrophoresis

CV – Coefficient of Variation

DTT – Dithiothreitol

IEF – Isoelectric Focusing

IPG – Immobilized pH Gradient

PAGE – Polyacrylamide Gel Electrophoresis

pI – Isoelectric Point

P.putida – *Pseudomonas putida*

PpKT2440 – *Pseudomoans putida* KT2440

SDS – Sodium Dodecyl Sulfate

SOP – Standard Operating Procedure

List of Figures

Figure 1: Isoelectric Focusing Schematic – Page 5
Figure 2: Structure of Sodium Dodecyl Sulfate – Page 6
Figure 3: Components of a 2DE System – Page 7
Figure 4: Growth Phases of a Bacterial Lifecycle – Page 9
Figure 5: Four Corners Method of Plate Streaking – Page 11
Figure 6: Sample Placement in an IEF Tray – Page 18
Figure 7: Protean IEF Cell – Page 19
Figure 8: Impurities Commonly Found on Gels – Page 23
Figure 9: Matching and Warping of the Gels – Page 24
Figure 10: Growth Curves of P. putidaKT2440 – Page 27
Figure 11: Average Growth Curve for P. putidaKT2440 – Page 27
Figure 12: Representative Gels from Each Time Point of Log Phase – Page 30
Figure 13: Results of the Intra-set Gel Comparison – Page 32
Figure 14: Representations of the Universal Standard Gel – Page 34
Figure 15: Histogram and Montage Window for Spot 142 – Page 36
Figure 16: Histogram and Montage Window for Spot 75 – Page 37
Figure 17: Histogram and Montage Window for Spot 87 – Page 38
Figure 18: Histogram and Montage Window for Spot 40 – Page 39
Figure 19: Expression Table for a Twofold Change – Page 42
Figure 20: Percent of Spots that Exhibit Fold Changes – Page 44
Figure 21: Coefficients of Variation Compared to the Universal Standard Gel – Page 45
Figure 22: Percent of Matching Spots to the Universal Standard Gel – Page 45
Figure 23: Comparison of the Intra-set and Universal Standard gel Analyses – Page 46

List of Tables

Table 1: Common Volumes of Growth Medium – Page 15

Table 2: Absorbance Readings of P. putidaKT2440 Growth Cultures – Page 26

Table 3: Mid-Log Calculations for P. putidaKT2440 on 5mM Succinic Acid – Page 28

Table 4: Summary of Protein Sample Preparations – Page 29

Table 5: Summary of Expression Changing Spots – Page 43

Table 6: Summary of Universal Standard Gel Comparisons – Page 48

Abstract

Two-dimensional gel electrophoresis (2DE) is the most commonly employed technique in the field of proteomics. The technique is designed to separate thousands proteins in highly complex biological samples. It is largely because of the diversity of the sample that 2DE has encountered problems with reproducibility. This study was designed to address and alleviate many of these problems and also to establish quantitative standards for intra-laboratory results.

A universal standard gel of *Pseudomonas putida* KT2440 was created and tested. The universal standard gel will serve as an intra-laboratory reference to which future experiments will be compared. An acceptable coefficient of variation between future control gels and the universal standard gel was determined to be less than 35%. Another aim of this study is to explore the effects of growth time on protein expression and determine a window of time in which growth can be stopped and protein expression is preserved. Currently there have been no studies relating to this topic. Protein samples were harvested every 30 minutes throughout the log-phase of growth and their protein expression was determined. It was found that within a 30 minute window of time surrounding mid-log phase the protein expression does not change, within statistical limits. This study will serve as the quantitative cornerstone for all future intra-laboratory experiments.

Introduction

I. Proteomics

With current advances in genomics and the complete sequencing of many different genomes, including the human genome, a new field has developed called proteomics. Now that the technological barriers that previously blocked researchers from obtaining genomic data have been overcome, many scientists are focusing on this new field. **[1]** Proteomics can be defined as the "the systematic analysis of the protein expression of healthy and diseased tissues." **[2]** Determining what proteins are activated or expressed when a cell is healthy and then making a comparison to the proteins that are expressed when a cell is under stress or diseased will prove to be invaluable to understanding our body's defense systems at the molecular level. Through proteomics we may unlock the key to understanding disease at a molecular level and then be able to combat the disease more affectively at a genetic level.

Even more impressive than the goals of proteomics are the potential applications and their implications. Various applications include diagnosing disease, drug design, and potential gene therapies. Imagine a person who has recently discovered a small lump under their skin and wants to have it tested to determine if it is cancerous. Instead of undergoing a sizable surgical procedure to remove a large section of the lump, proteomic analysis will only require a few cells to be removed and their proteome analyzed. If these cells are expressing known carcinogenic biomarkers then the diagnosis will be cancer. However, if these known biomarkers are not present then the tumor is benign. Taking this scenario one step further, if the tumor is determined to be cancerous and the patient has to undergo treatment,

proteomics may again play a role. Cancer is a result of uncontrolled or unregulated cell division and proteomic research has the potential to determine what proteins are involved in regulating cell division. If a drug that targets the protein(s) responsible for the uncontrolled division can be designed, it may be able to bind to the protein. This binding will cause the protein to be inactivated and then removed from the cell and the uncontrolled division will cease. Those diseased cells will stop dividing, eventually die, and be replaced by normal healthy cells. Presently this is only a hypothetical situation, however, the technology is present, the research is currently being conducted, and it is quickly becoming reality – all through the understanding and application of proteomics.

The possibilities and promise in this field of research are extremely provocative but accomplishing these goals can prove to be quite tedious due to the variability of the sample. The proteomic signature, which is the specific set of proteins expressed under a particular set of conditions, will change depending on many different variables like environment, age, and stress. The most difficult problem to overcome in proteomics is attempting to separate the vast number of proteins present in a sample. One individual cell contains several thousand different proteins that can all be expressed at different levels and at different times throughout the cell life cycle. The majority of proteomic research employs one of two techniques. Either a multi-dimensional liquid based separation like column chromatography or the alternative of gel electrophoresis. The most commonly used technique is two-dimensional gel electrophoresis (2DE). **[3]** Although this technique is the most commonly used, it presents some inherent problems with reproducibility. This issue

of reproducibility, the development of a standard operating procedure, and the elimination of factors that contribute to variability among samples are the main focus of this project.

II. The Basis of Separation in Two-dimensional Gel Electrophoresis.

2DE involves two dimensions of separation to thoroughly separate the complex mixture of the protein sample. The first phase of separation is isoelectric focusing (IEF) and it separates the proteins based on their charge. The current technology used in isoelectric focusing is immobilized pH gradient strips, or IPG strips. These IPG strips are made by covalently linking acrylamido buffers of different pH and then setting them on a plastic backing. The acrylamido buffer is a derivative of acrylamide and contains both reactive double bonds and buffering groups. The liquid protein sample is taken up into the strip by passive or active rehydration and the proteins are focused. During focusing an electric current is applied to the rehydrated IPG strip which causes the proteins to migrate. (See figure 1.) Each protein will continue migrating until it reaches a point on the strip where the pH gradient equals the isoelectric point (pI) of the protein itself. At this point, the net charge on the protein is zero. **[4]**

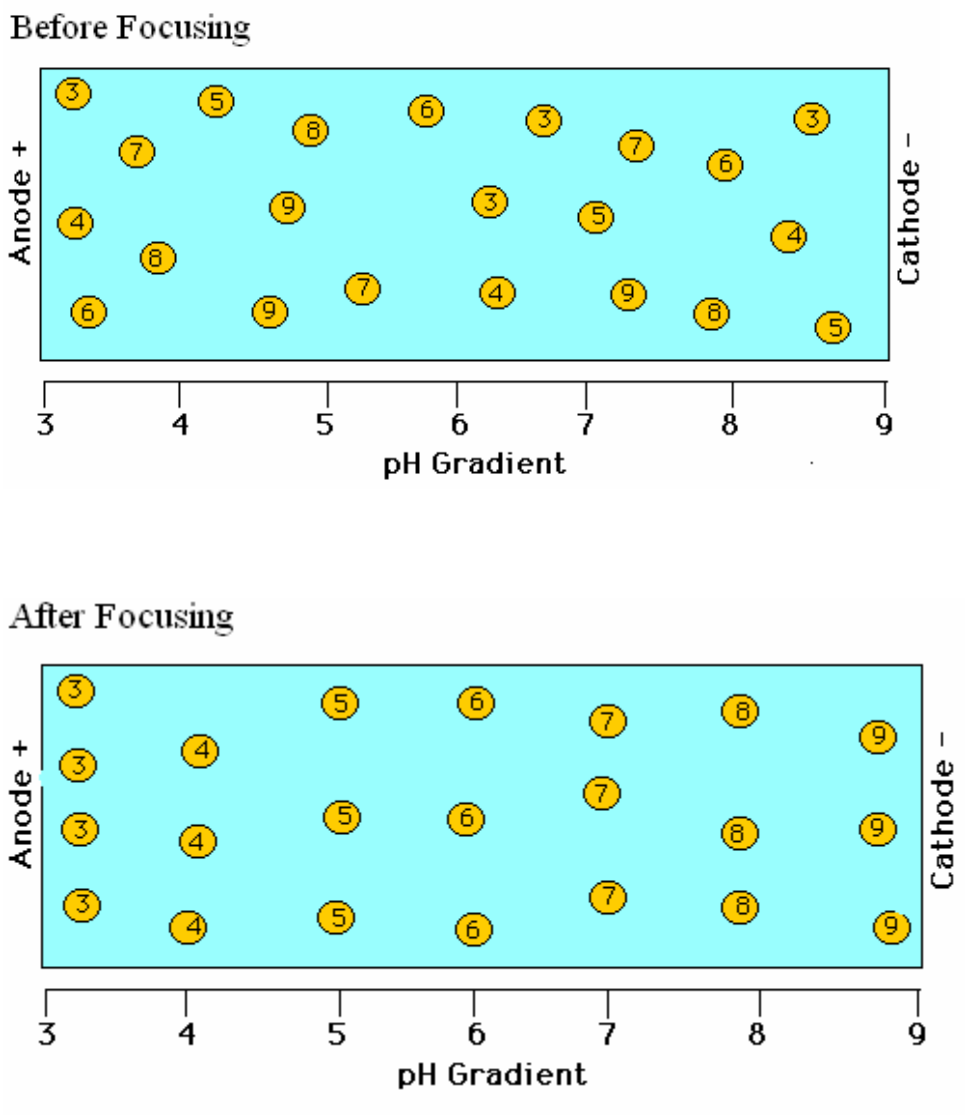


Figure 1: Before focusing the proteins, represented by the circles, are randomly assorted on the IPG strip.**[5]** After focusing the proteins are separated based on their isoelectric points, depicted by the number on each protein.

The second dimension of separation is sodium dodecyl sulfate polyacrylamide gel electrophoresis, or SDS-PAGE. In this phase of separation the proteins are sorted vertically based on their molecular weight. Before the proteins can be separated in the second dimension, they must first be equilibrated in sodium dodecyl sulfate (SDS), see

figure 2. SDS will bind to the protein in a ratio of one SDS molecule for every two amino acids. This binding capability causes the protein to once again become negatively charged in proportion to its molecular weight, which is essential for migration in the second dimension.

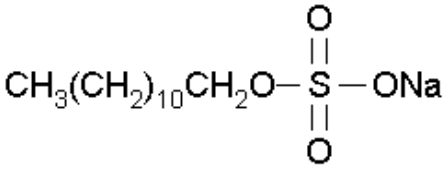


Figure 2: Structure of sodium dodecyl sulfate. **[6]**

After equilibration in SDS, the IPG strip which contains the horizontally focused proteins is placed on top of the acrylamide gel and overlaid with low-melting point agarose to hold the strip in place. The gel cassette is now placed into the electrophoresis cell and the entire electrophoresis cell is filled with a running buffer. (See Figure 3) Once assembly of the electrophoresis system is complete, an electric current is passed through the system causing the charged proteins to migrate according to their molecular weight.

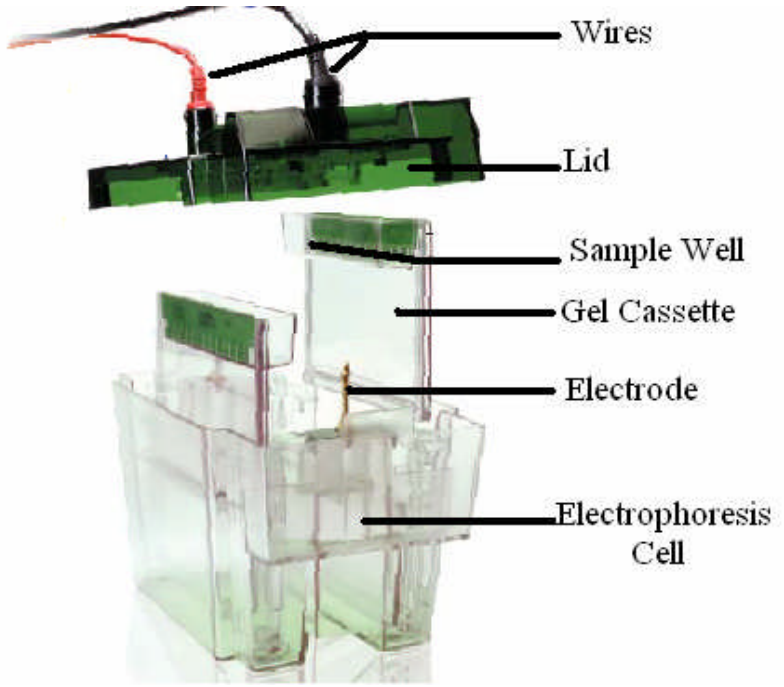


Figure 3: Compontents of a two-dimensional polyacrylamide gel electrophoresis system. **[7]**

III. Challenges of Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis is conceptually sound and current advances have made it more reproducible than earlier applications of the technique. However, there are still well-justified concerns about reproducibility of the technique. Recently, a number of researchers have begun addressing, and publishing, their findings regarding reproducibility and are setting guidelines for statistically acceptable 2DE data. In 1995, the Blombergs et. al. conducted an inter-laboratory study assessing the reproducibility of IPG strips and IEF where they measured the x and y positions of S^{35} labeled yeast proteins. **[8]** As an alternative to IPG strips, Lopez and Patton studied reproducibility of a 2DE separation employing carrier ampholytes to achieve the first dimension of

separation. **[9]** Similar to the Blomberg study, Lopez and Patton were mainly concerned with the positional reproducibility of spots from one gel to another. As a follow-up to the Lopez-Patton study, Voss and Haberl, in 2000, reported quantitative data which suggested that running a very large number of replicate gels may actually decrease the matching efficiency between gels. **[10]** This lead to the suggestion that 3 – 4 replicate gels are sufficient to fulfill quantitative statistical matching requirements. Running more replicate gels than the suggested amount may cause the important variations in individual spot intensity to be overlooked. For example, if an investigator were to run ten replicate gels there would be a fairly wide variation between a specific spot's intensity across these gels. These differences in intensity are mostly due to slight variations in staining of the gels. When the ten replicate gels are then digitally merged to make an averaged gel, most of the spots would appear to be of an average intensity level. Another reason that more replicate gels are not recommended is because most laboratories are not equipped for high-thorough put operations, and only have the capability to simultaneously run a maximum of four replicate gels through SDS-PAGE. More recent studies have addressed reproducibility pertaining to: sample load **[11]**, manual verses automated use of analysis software **[11]**, acceptable coefficients of variation between replicate gels **[11- 13]**, the effects of sample preparation **[14, 15],** reproducibility in the second dimension **[16]**, and variation induced by image analysis and staining. **[17]**

VI. Specific Aims

This study is designed to address any questions associated with cellular growth time and its effects on protein expression relating to gel reproducibility. The second aim

is to develop a ‘universal standard gel’ that can be used as an intra-laboratory reference. The first aim of the study is to quantitatively determine the window of time in which cellular growth can be stopped and the investigator can be assured the protein expression is conserved. *Pseudomonas putida* KT2440 will be grown on succinic acid and samples will be taken at 30 minute intervals throughout the log phase of growth. (Figure 4).

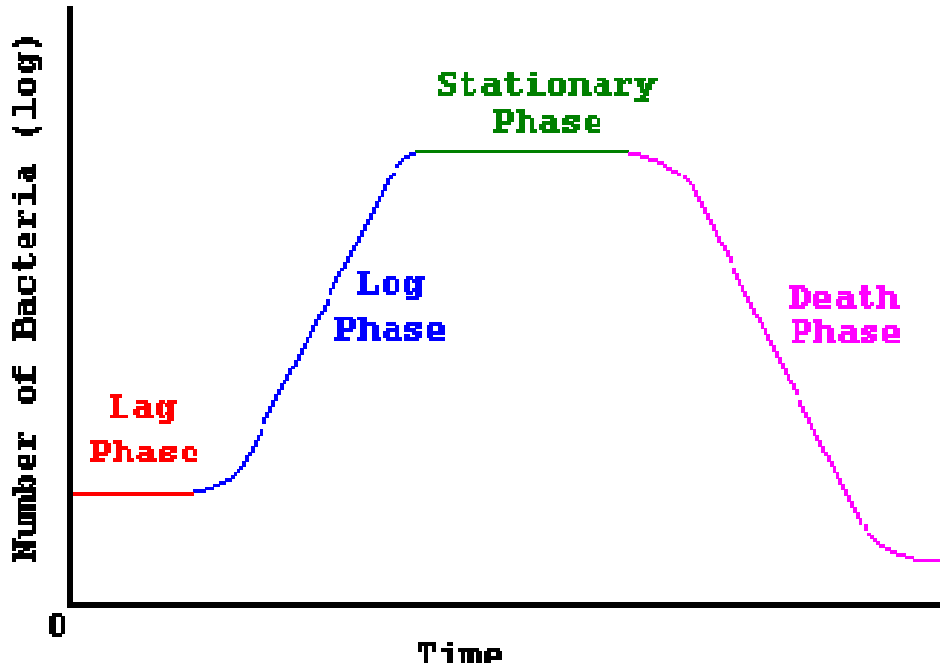


Figure 4: The growth phases of the bacterial life-cycle. **[18]**

Aliquoits of a large cell culture will be taken starting at 5 hours of growth and continuing through 7 hours of growth at half hour intervals. Protein will be isolated from each one of these aliquiots and evaluated using 2DE. The protein expression at each of these 30 minute time points will be compared to one another to determine the necessary time frame in which growth must be stopped to ensure that there are no significant

differences in protein expression. This will be accomplished by using the analysis software Phoretix 2D, from Nonlinear Dynamics, to determine statistical limits of acceptable, and expected, variation between gels. The data generated from this experiment will allow future investigators to disregard or dismiss any gels that fall outside the established statistical limits. This approach is similar to dismissing quantitative data on the basis of a T-test.

The second aim of this work is to produce multiple gels from the protein isolated from two P. putidaKT2440 cell cultures that were collected at the middle of the log (mid-log) phase of growth. These gels will be averaged together, through the use of Phoretix 2D software to create an 'averaged gel'. An averaged gel is an artificial gel that is created from a statistical combination of several normal gels. **[19]** This averaged gel (referred to as the "universal standard gel" from this point forward) will be used as an intra-laboratory reference gel. As in all scientific endeavors it is important to implement and establish standards that can be used as control variables in future experiments. After the universal standard gel is completed future investigators will be able to run their sample gels along with a control gel (P. putidaKT2440 grown on 5mM succinic acid to mid-log phase) that should match the universal standard gel. The protein expression and resolution of this new control gel can be statistically compared to the universal gel standard. If the new control gel matches the universal standard gel, within statistical limits, then the investigator can be assured that the process is working correctly and sample preparation did not effect or influence reproducibility.

Methods and Materials

When conducting scientific research, especially in the field of proteomics, the method used for analysis is slightly altered depending on the question to be answered. The following experimental conditions are the general ones used in this study of *Pseudomonas putida* KT2440. Any necessary additions or descriptions for a specific experiment will be included in the following Results and Discussion section.

I. Growth conditions

A freeze dried sample of *P. putida* KT2440 was obtained from American Type Culture Collection (www.atcc.org; ATCC® catalog # 47054) and rehydrated according to their protocol. From this sample 50% glycerol stocks were made and then snap frozen using a slurry of ethanol and dry-ice. A platinum loop was used to streak the liquid sample on an agar plate which also contains 5mM succinic acid and a variety of other minerals necessary for growth. (Figure 5) The streaked plate was then incubated at 37°C for 24 – 48hours. The goal of streaking plates is to dilute a sample to a single, genetically identical, colony of cells. Each individual colony is considered genetically identical because all of the cells, theoretically, replicated from one founding cell.

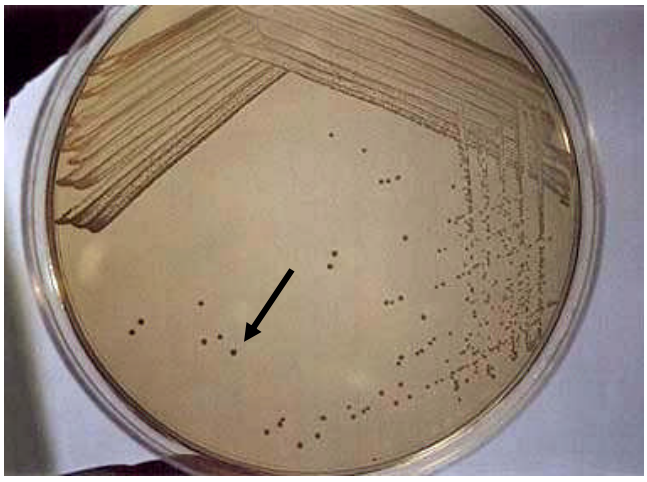


Figure 5: An agar plate that was streaked using the "four corners" method. The arrow indicates a single colony that is composed of genetically identical cells. **[20]**

After obtaining a plate containing single colonies, one of the colonies is picked from the plate using sterile techniques and inoculated in liquid growth medium. All of these liquid cultures are composed of Hutner's Growth Medium. **[21]** The first step in growing the bacterial culture is preparation of the growth medium, which was prepared in the following manner.

Begin by making the 'stock' solutions A, B, C and the Metals 44 Solution according to the recipes below:

Solution A: per liter deionized H_2O:

Na_2HPO_4	141.2 g
KH_2PO_4	136 g

(should give a pH = 7.25)

Solution C: per liter deionized H_2O:

$(NH_4)_2SO_4$	200 g

Metals 44 Solution:

In 800 mL deionized water, dissolve the following:

EDTA	2.50 g
$ZnSO_4 \cdot 7H_20$	10.95 g
$Mn\ SO_4 \cdot 7H_20$	1.54 g
$Fe\ SO_4 \cdot 7H_20$	5.00 g
$Cu\ SO_4 \cdot 5H_20$	392 mg
$Co(NO_3)_2 \cdot 6H_20$	248 mg
$Na_2B_4O_7 \cdot 10H_20$	177 mg

Add a few drops of concentrated H2SO4 to retard the precipitation. Dilute solution with deionized water to a total volume of 1L. The solution should appear clear and lime green in color, and can be stored indefinitely. It

should be stored in a brown glass bottle at room temperature in a closed cabinet.

Solution B: per liter deionized H_2O:

Nitrilotriacetic acid	10 g
MgSO4	14.45 g (or 29.56 g MgSO4 · $7H_20$)
$CaCl_2 \cdot 2H_20$	3.33 g
$(NH_4)_6Mo_7O_{24} \cdot 4H_20$	9.25 mg
Fe $SO_4 \cdot 7H_20$	99 mg
Metals 44	50 mL

It is recommended **[21]** to make up 4L at a time of Solution B with the following procedure:

1. Dissolve 40 g nitilotriacetic acid (NTA) in about 2.5-3 liters of water. (This is easier if the water is hot when adding the NTA. Heating the water to boiling is OK and seems to help.)

   Note: Be careful with NTA, this compound is known to cause hereditary genetic damage.
2. Dissolve 118.24 g $MgSO_4 \cdot 7H_20$ in about 300 mL water. Again, heat may help to dissolve the compound.
3. Dissolve 13.32 g $CaCl_2 \cdot 2H_20$ in 100 mL water
4. Dissolve 37 mg ammonium molybdate and 396 mg $FeSO_4 \cdot 7H_20$ together in 100 mL water.
5. Add the magnesium sulfate solution to the NTA solution slowly while stirring (Use a Teflon stir bar) and be careful to avoid any clouding of the mixture. Once these are mixed, gradually add the calcium chloride solution. Then add the remaining solution. The mixture should now

be a yellow color. Now add 200 mL Metals 44 solution . The solution should now be a deeper yellow color.

6. Bring the volume up to a total volume of 4L, in a large beaker.

At this point have a pH probe and meter inserted into the liquid. Bring the pH up to about 6.7 with 10M NaOH. This must be done slowly, especially towards the end. You will see a transitory precipitate that eventually dissolves upon vigorous stirring. Let the precipitate dissolve before adding more NaOH. If the pH goes over 7, the solution will turn a green/gray color and precipitate will form. It is possible to salvage this solution by adding a drop or two of concentrated sulfuric acid (H_2SO_4) and then again adding NaOH until the pH is 6.7.

All medium component solutions should be autoclaved to ensure sterility. A heavy precipitate will form upon autoclaving. It will dissolve as the medium is stirred, using a magnetic stir bar and stir plate, while cooling. To regain sterility, vacuum-filter through Whatman qualitative filter paper and a low-protein binding sterile filter apparatus (from Corning, catalog #: 25933-200) before storage in order to remove any particle impurities that may be present. After being autoclaved and sterile filtered the solutions can be stored in brown bottles at room temperature for an indefinite amount of time.

These stock solutions are now combined to make the complete Hutner's Growth Medium in any desired amount. An important aspect in combining the stock solutions is to maintain their set ratios. Table 1 contains volumes of growth medium that are commonly prepared and the ratios of the stock solutions. It should be mentioned, that for

all of these experiments succinic acid is used as a carbon source and is added to the growth medium to make the final concentration equal to 5mM of succinic acid.

Table 1: Common growth medium volumes for 5mM succinic acid solutions.

Desired Volume	**700ml**	**500ml**	**200ml**
Solution A	28ml	20ml	8ml
Solution B	14ml	10ml	4ml
Solution C	3.5ml	2.5ml	1ml
Succinate (5M)	7ml	5ml	2ml
Distilled Water	647.5ml	462.5	185ml

Once the Hutner's growth medium has been prepared a single colony is picked from an agar plate with the platinum loop tool, placed into a small volume of growth medium (50 – 100ml), and gently shaken to transfer the cells. This "starter culture" is now incubated overnight at 30°C with moderate agitation (approximately 140 rpm in an orbital shaker). After 24 hours of growth the culture will be completely saturated with cells in the stationary phase and will appear grey, cloudy, and almost opaque. A portion of the starter culture is now transferred to a large volume of new growth medium, in the ratio of 1 part starter culture to 100 parts new medium, and incubated until mid-log phase of the growth curve. At mid-log phase cell growth is stopped and the cells are harvested by centrifugation.

II. Centrifugation

After growth time reaches mid-log phase, the entire cell culture is transferred to tared 250 ml centrifuge tubes. Using the Sorvall® RC 5C Plus centrifuge and the correct rotor (SLA-1500), the cells are spun at 12,000 x g (8,200 rpm) and at 4°C for 40 minutes. This centrifugation step results in a large yellow colored cell pellet with a clear supernatant. The supernatant can be discarded and the centrifuge tubes should be inverted for approximately 10 minutes to remove any remaining liquid. At this point a wet weight of the cell mass should be measured.

III. Sonication

The cell pellet can be removed from the centrifuge tubes by dissolving it in 1 – 3 ml of rehydration buffer [48.1 g of Urea (8M), 4.0 g of CHAPS (4%), 0.2 mL of Ampholytes, and 0.77 g of DTT (50mM) in100 ml ddH_20] so that the volume of rehydration buffer is twice the mass of the cell pellet. The resuspended cell solution is transferred to a 50ml beaker and placed on ice. The sonicator (Ultrasonics Heat Systems Mod W-375, Convertor Mod C3) should be warmed up for at least 10 minutes before use and should be set to %Duty cycle 43 – 48 and output control = 3.5. The sonicator tip is rinsed with doubly-distilled water (ddH_20), then acetone, and again with ddH_20. The sonicator tip is gently lowered into the beaker, taking care not to touch the sides or bottom, and the suspended pellet is sonicated for 3 minutes then left on ice for 10 minutes. This process is repeated three times for a total sonication time of 9 minutes. The final step of sample preparation is to aliquot the sonicated protein sample into sterile 1.5 ml disposable microfuge tubes which are then spun down in a microfuge (Biofuge 13 Mod. 7500464/01) for 5 hours at 13,000rpm and 4°C. Upon completion of the microfuge

step there should be a small black pellet in the bottom of the microfuge tube. Transfer the supernatant, which contains the denatured protein, to new microfuge tubes without disturbing the unwanted pellet. The protein can be stored at 4°C or used immediately.

IV. Determining Protein Concentration

The concentration of the protein is determined by performing the Bradford Protein Assay **[26]** which has been modified by Biorad and can be found on page 30 in the Bio-Rad 2D Electrophoresis for Proteomics: A Methods and Product Manual. **[22]** The concentration of the protein sample must be found to determine the sample volume to be applied to the IPG strips. The general procedure is to create a calibration curve using the standard protein bovine serum albumin (BSA). Then by making dilutions of the sample protein and comparing the absorbance readings to this calibration curve, the investigator can back-calculate the concentration of the sample protein.

V. Treatment with RNAse

Treating with RNAse (Fisher Scientific, cat. #: BP2539250), will remove any RNA still present in the sample, which will lead to better sample purity and less background on the gels. RNAse should be added to the sample in the ratio of 2mg RNAse for 1ml of sample. The protein is microfuged under the previous conditions for 20 minutes then transferred to a clean microfuge tube.

VI. Rehydration of IPG strips **[23]**

Begin by removing the stock solution of rehydration buffer, the protein sample, and the appropriate number of IPG strips from the freezer and/or refrigerator. The rehydration stock solution will show some precipitate, but as it warms all components should re-enter solution. Based on the results of the protein assay, aliquot the protein

sample into new microfuge tubes, one tube per strip, so that each tube contains 200µg of protein. Add the rehydration stock buffer to the 200µg of protein, giving a total volume of 185µl per tube. Mix the sample solution well by pipetting in and out. (A trace amount of bromophenol blue may be added to each sample to help visualize the liquid.) Pipette up the entire 185µl of sample and eject it along the back corner of the lane in the rehydration tray. (See Figure 6)

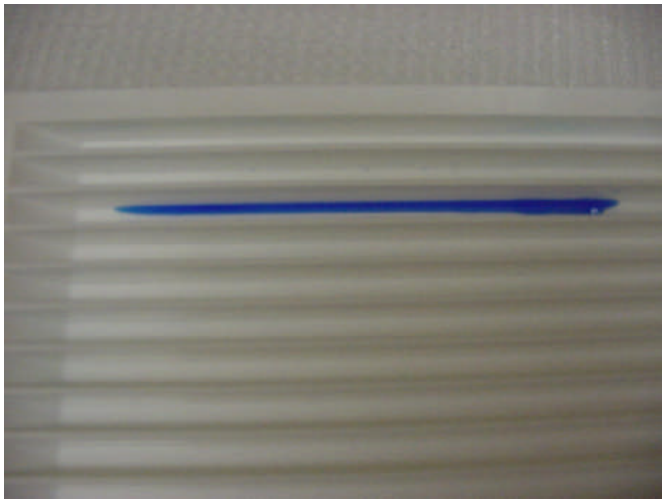


Figure 6: A protein sample that has been dyed with bromophenol blue and correctly loaded into the IEF tray.

Holding one end of the IPG strip with forceps carefully peel away the protective covering. Starting at one end carefully lay the IPG strip, gel side down, over the sample. To ensure even rehydration or sample uptake, make sure no air bubbles are trapped under the strip. Place the plastic cover over the rehydration tray and leave the sample at room temperature for approximately 30 minutes. Cover the IPG strip with 3 – 4 ml of mineral oil to prevent evaporation of the sample and leave for 11 – 16 hours.

VII. Isoelectric Focusing (IEF) **[24]**

Place a clean, dry focusing tray on the lab bench with the + to the bottom left side. Place electrode wicks (BioRad, cat. #: 1654071) over the wire electrodes in the wells to be used (using forceps). Pipette 8µL of nanopure water onto each wick. Using clean forceps, carefully remove 1 IPG strip from the rehydration tray and hold it vertically for about 7 – 8 sec. allowing the mineral oil to drain. Lay one piece of dry filter paper on the bench top and place the strip gel side down on the filter paper. Wet a second piece of filter paper, place it on top of the IPG strip and very gently blot the strip to remove any excess oil or protein sample. This will reduce the horizontal streaking on the gel. Then place the strip gel side down in the focusing tray. Repeat this process for any remaining strips. Cover each strip with 2 – 3 ml of mineral oil, place the lid on the tray and carefully move the focusing tray into the Protean IEF Cell, making sure that polarities are correct. Close the cover and begin running the desired focusing method.



Figure 7: Protean IEF Cell by BioRad. **[24]**

All of the isoelectric focusing conditions were the same for all samples in this experiment. The samples are focused for 16 hours under the following conditions all under linear gradients: Step 1 – 6 minutes at 300V, Step 2 – 4.5 hours at 300V, Step 3 –

5 hours at 4000V and Step 4 – 6.5 hours at 4000V. **[22]** Upon completion of IEF the strips can be stored at -20℃ or used immediately in the second dimension of separation.

VIII. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Pipette out two 10 ml portions of equilibration buffer stock solution (36 g of 6M urea, 10 ml of 20% SDS, 3.3 ml of 1.5M Tris/HCl pH = 8.8, 40 mL of 50% glycerol, in 100 ml ddH20) into 2 separate centrifuge tubes labeled "buffer I" and "buffer II". Add 200 mg of DTT to buffer I and shake to get DTT into solution. Add 3 mL of this solution to each well containing a rehydrated strip, and place the tray on the orbital shaker (40 rpm) for 10 minutes. To complete the preparation of buffer II add 250 mg of iodoacetamide to the tube and shake until it is dissolved. While strip(s) are in orbital shaker, place a vile of stock agarose (0.83 mL of 1.5M Tris buffer (from BioRad), 0.721 g of glycine (from Sigma Aldrich), 0.25 g of low melting point agarose (from Sigma Aldrich), 0.25 mL of 80% SDS (concentration should be 0.1% SDS), trace amount of bromomphenol blue, diluted to a total volume of 50 mL with D.I. water) in the microwave for 30 second increments until it becomes less viscous and slightly warm. Also while the strips are in the shaker, mix 100 mL of 10X Tris/Glycine/SDS Buffer (from Sigma Aldrich) with 900 mL of nano-pure (18μΩ) water in a 1 L graduated cylinder to make the running buffer. Mix thoroughly by covering with parafilm and inverting several times.

When the strips have finished shaking for 10 minutes, slowly pour off equilibration buffer I and add 3 mL of buffer II to each strip, again shake (40 rpm) for 10 minutes. Take pre-cast SDS-PAGE gels out of the refrigerator and out of packaging. Remove the green comb and the white sticker from the gels and rinse grove/wells with

nanopure (18μΩ) water. Remove the IPG strips from the shaker and pour off buffer II and cover them with running buffer and place in orbital shaker for 10 minutes (40 rpm). While strips are shaking, cut a 1 cm wide strip of blotting paper (BioRad, cat. # 1704085) and dry/blot out the well in the pre-cast SDS-PAGE gel. Then remove IPG strips from the shaker and drain the running buffer. With forceps, carefully place IPG strip into pre-cast SDS-PAGE gel slightly above the well with the plastic backing of the IPG strip against the plastic support of the gel cassette. Fill the well to the top with melted agarose and push the IPG strip into the well. Let the gel sit for approximately 5 minutes allowing the agarose to solidify. Repeat this process for any remaining IPG strips.

Rinse the electrophoresis cell with distilled water. Insert the pre-cast SDS-PAGE gels into the electrophoresis cell and pour running buffer into each cell up to the fill line. Make sure the wires on the green top of the electrophoresis cell are completely submerged in the running buffer, which completes the electric circuit. Place the top on the electrophoresis cell and plug the wires into the outlet on the power source. Set the voltage to 200V, set the run time for 50 minutes, and begin the run. Check the gel at least every 15 minutes to see where the dye-line is on the gel to make sure the sample has not run off the gel. If the dye-line runs off of the SDS-PAGE gel some of the proteins may also run off and be lost from the sample.

When SDS-PAGE has been completed the gel can be removed from the plastic casing by placing the well of the gel upside down on the ridge of the electrophoresis cell lid and gently pressing down. The pressure on the lid will cause the sides to crack open and leave the gel on one side of the plastic casing. The side with the gel present can be

submerged in distilled water and with gentle prodding of the gel, using a disposable pipette, the gel can be removed.

IX. Staining the gel

This procedure was edited from the BioRad methods and product manual (page38 - 39. **[22]**

After the gel has been removed from the plastic backing, rinse the gels at least twice for 10 minutes in deionized water with agitation to remove SDS. Add sufficient Bio-Safe Coomassie Blue stain (from BioRad catalog # 161-0787) to cover the gels and allow the gels to soak in the stain for at least 90 minutes. Coomassie blue is an end-point stain, so staining for longer periods will not affect the resolution. Repeat rinsing procedures with deionized water until desired contrast between background and protein spots is achieved. The gels can be left to de-stain in deionized water over night.

X. Image Analysis

After staining and de-staining are completed the gels were imaged using a charged coupled (CCD) camera. The resolution of the images can be improved by adjusting the brightness and contrast settings in any image analysis software program. For this study, the image analysis software of choice is ImageJ, which can be downloaded, free of charge, from the following website: http://rsb.info.nih.gov/ij/ .

To prepare the images for future analysis certain guidelines must be followed. The images must be saved as "tiff files" with the extension .tif or .tiff. All the images involved in a study must also possess the same resolution – either 8 or 16 bit, and all

the images must be greyscale. The final critical component of preparing the images for analysis with the Phoretix 2D software is to standardize the size of the images, which greatly simplifies the process of matching the gels to each other.

XI. Analysis with Phoretix 2D Software

To begin the analysis, all of the processed images for a study must be saved in the same folder. All of the images in this study were analyzed by following the instructions in Tutorial 1 in the Phoretix 2D Manual (2004 version). **[25]** When detecting spots the sensitivity, noise, operator size and background settings were applied to all the gels in the experiment. Then each gel was manually edited to remove any artifact spots at the gel edges or impurities that the software may have considered protein spots. Manual editing was also used to split any obvious spots that contain two different proteins. (See Figure 8)

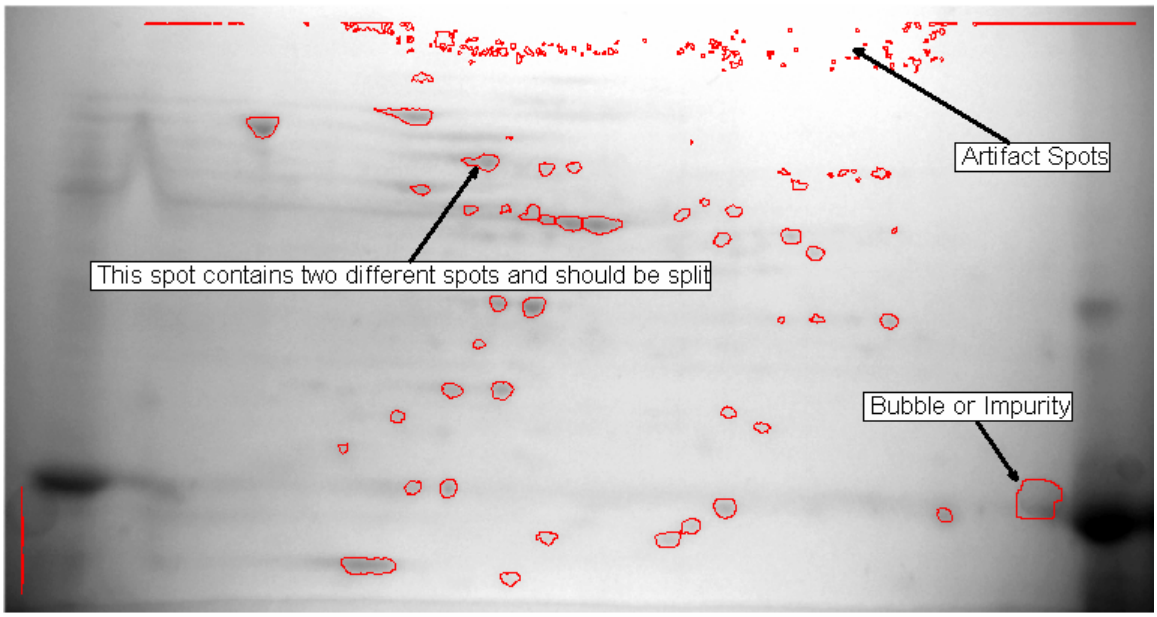


Figure 8: Examples of spots or impurities that need to be manually edited on the gel.

After spot detection and the manual editing process is complete the gels were subjected to background subtraction. The method of background subtraction was "mode of non-spot", which is fully automatic and encloses each spot in a rectangle. This rectangle is then expanded by a set number of pixels, to define the area of examination. The pixels in the defined area that are not part of any spots are then examined and the most frequently occurring pixel intensity is considered the background value of that spot. The next step in the analysis is to perform spot normalization, by the method of "total spot volume normalization". In this method the normalized volume of the spot is calculated by dividing the spot volume by the total volume of all detected spots on the image and then multiplying by a constant factor. For all analyses in this experiment the default settings were used for background subtraction (the set number of pixels to expand by is 45) and spot normalization (multiply by the constant factor of 100).

The last aspect of the gel analysis is matching and warping. To begin matching a reference gel must be chosen. For most aspects of this study the gel containing the most spots was used as the reference gel. Parts of the analysis where this criteria was not used to assign the reference gel are discussed in the following Results and Discussion section. An example of this matching procedure is depicted in figure 9.

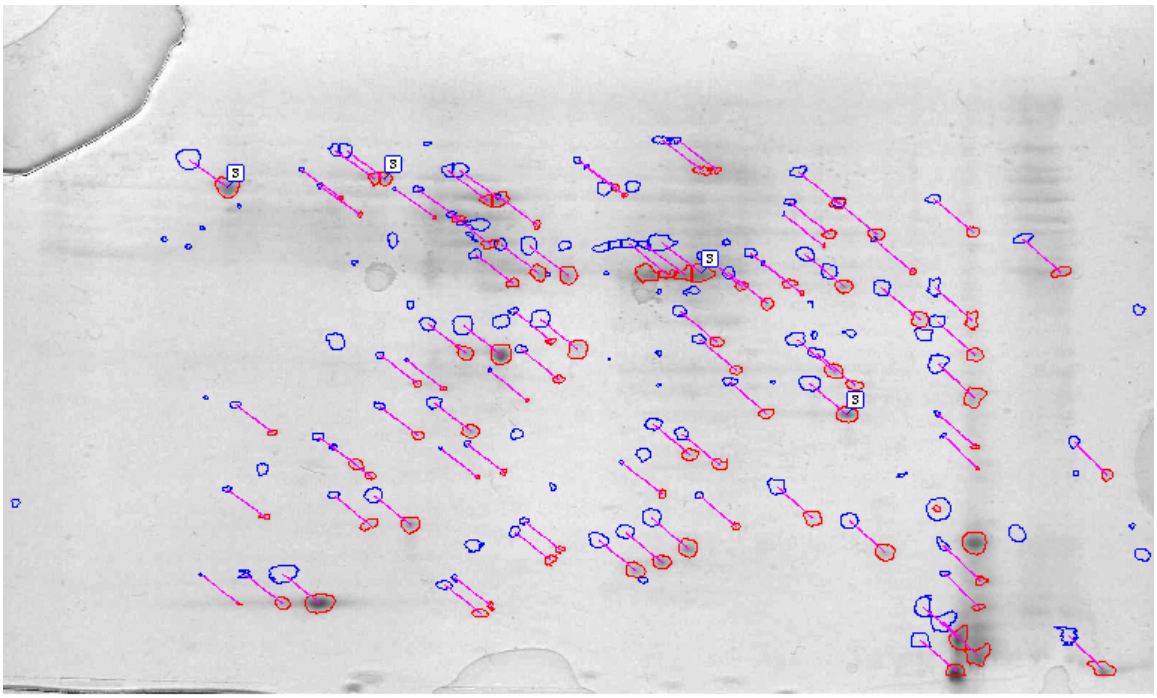


Figure 9: The matching of two gels. The red spots are the reference gel spots and the blue spots are the sample gel spots. The pink lines are the matching vectors that show the trends of matching within the gel.

The spots marked with an 'S' are user defined "seed matches". These seed matches help to establish the general trends on the gel that greatly aid in the matching process. Once these trends in the gel are identified and an unmatched spot is selected, the software will suggest the area of the gel where a matching spot should be found. If a spot is present in that location, it can be assigned as the match to the selected spot. Another feature of the software that assists in spot matching is the "warp to seed matches" option. This will physically overlay any already matched red and blue spots and shift any unmatched spots in the area to an appropriate location based on the trends of matching. This feature allows the user to identify less obvious spot matches. Upon completion of

matching and warping the gel images, the analysis is complete and the statistical data can be examined. This data analysis is discussed in detail in the following Results and Discussion section.

Results and Discussion

The first critical point of this study is reproducibly determining the mid-log phase of growth of P. putidaKT2440 on 5mM succinic acid. Cell cultures were grown in triplicate, as described in the methods and materials section, and the absorbance was measured at 30 minute intervals. (See table 2 and figure 10) The first three trials were performed by Mingyi Yao (a previous graduate student at RIT) and trials four through six were completed by the author.

Growth	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6		
Time (h)	A595nm	A595nm	A595nm	A595nm	A595nm	A595nm	Average	Std. Dev.
0	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.0002
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0000
2	0.005	0.004	0.004	0.001	0.004	0.000	0.003	0.0007
3	0.018	0.013	0.015	0.018	0.009	0.013	0.014	0.0015
4	0.022	0.022	0.022	0.023	0.021	0.018	0.022	0.0004
5	0.143	0.145	0.140	0.154	0.149	0.134	0.144	0.0024
6	0.250	0.244	0.244	0.252	0.252	0.244	0.248	0.0016
7	0.377	0.387	0.409	0.387	0.382	0.377	0.386	0.0050
8	0.420	0.420	0.420	0.432	0.420	0.420	0.422	0.0021
9	0.398	0.432	0.398	0.398	0.409	0.420	0.409	0.0060
10	0.387	0.398	0.387	0.387	0.375	0.409	0.391	0.0037

Table 2: Absorbance readings of P. putidaKT2440 cell cultures to determine mid-log phase.

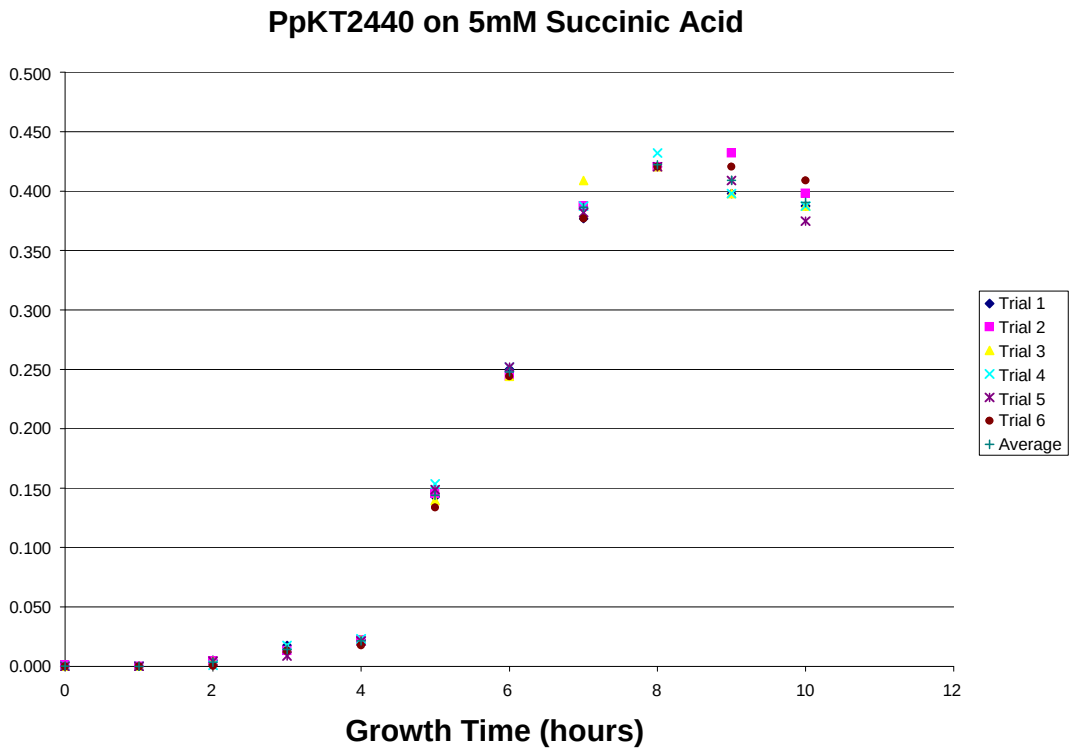


Figure 10: Bacterial growth curve and life cycle of PpKT2440 grown on 5mM succinic acid as a carbon source

To calculate the time it takes to reach mid-log phase, the equation of the line (see Figure 11) was solved for the variable 'x' which represents time.

Original Equation of the line: $y = 0.1199x - 0.4593$

Solved for 'x': $x = (y + 0.4593)/ 0.1199$

Therefore: Time = (Absorbance + 0.4593)/ 0.1199

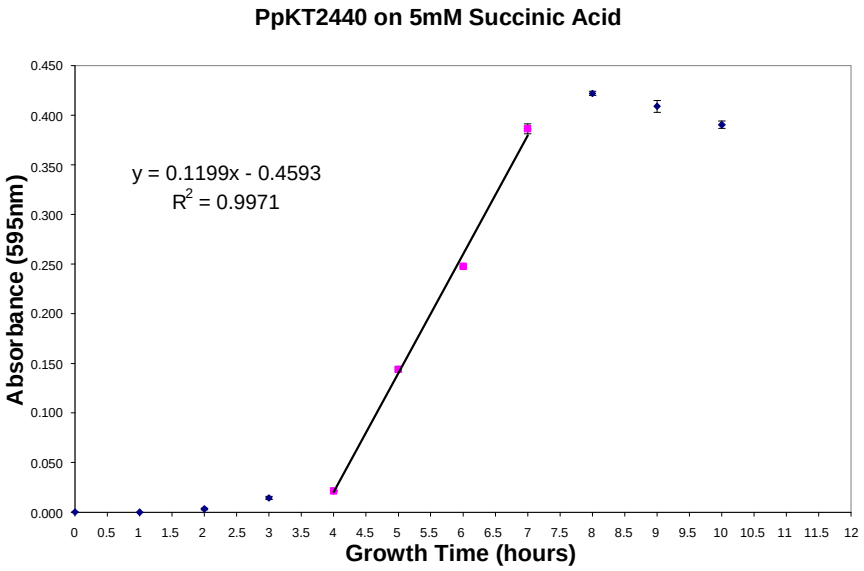


Figure 11: The average of the above six trials was calculated and is plotted above. A linear trend line was drawn through the log phase of growth resulting in the above displayed equation of the line, which will be used to calculate the time it takes the cultures to reach the mid-log phase.

The absorbance at mid-log phase was calculated for each trial by subtracting the lowest absorbance of the curve from the highest absorbance of the curve and dividing by 2. The time it takes to reach mid-log for each trial is summarized below (see Table 3).

	High - Low Abs	Mid-Log Time (h)
Trial 1	0.210	5.58
Trial 2	0.216	5.63
Trial 3	0.210	5.58
Trial 4	0.216	5.63
Trail 5	0.210	5.58
Trial 6	0.210	5.58

Table 3: Summary of mid-log calculations based on trials 1 – 6. Mid-log was found to be 5.6h ± 0.025h.

Based on these six trials, it was found that PpKT2440 on 5mM succinic acid as a carbon source reaches mid-log phase after 5 h 35 min ± 2 minutes. These results confirm that a

very high degree of reproducibility can be achieved when calculating the mid-log phase of growth by monitoring cell culture absorbance readings. However, the true variability is more likely closer to ±10 minutes because of the manual processes that must occur during the associated procedural steps.

Now that the mid-log phase of growth has been quantitatively determined, the first significant aim of this study can be addressed which is to determine the window of time surrounding mid-log phase in which cellular growth may be stopped without compromising the protein expression observed on the gels. To begin exploring this aim a 1 L culture of P. putidaKT2440 was grown on 5mM succinic acid and aliquoits were taken every thirty minutes throughout the log phase of growth. Different culture volumes needed to be harvested at different times of growth due to the exponential nature of cellular replication (see Table 4). Taking different culture volumes ensured that ample protein could be extracted. These aliquoits were spun down to isolate the cell pellet, the protein was extracted, purified, and the concentration of protein was calculated by performing a protein assay, as described in sections II – IV of the previous Methods and Materials section. This entire process was repeated twice (once on 11/10/04 and again on 3/21/05) so that each protein sample was collected in duplicate.

Cells Harvested from 11-20-04				
Volume of Culture Taken	Total Growth Time	A_{595} of Culture	Weight of Cell Pellet	Protein Concentration
250 ml	5 hours	0.112	0.41 g	2.78 µg/µl
250 ml	5.5 hours	0.175	0.45 g	3.33 µg/µl
200 ml	6 hours	0.224	0.34 g	3.49 µg/µl
150 ml	6.5 hours	0.283	0.36 g	3.22 µg/µl
100 ml	7 hours	0.364	0.30 g	3.29 µg/µl
Cells Harvested from 3-21-05				
Volume of Culture Taken	Total Growth Time	A_{595} of Culture	Weight of Cell Pellet	Protein Concentration
250 ml	5 hours	0.106	0.39 g	2.42 µg/µl
250 ml	5.5 hours	0.164	0.37 g	3.05 µg/µl
200 ml	6 hours	0.220	0.30 g	3.52 µg/µl
150 ml	6.5 hours	0.275	0.32 g	2.98 µg/µl
100 ml	7 hours	0.350	0.27 g	2.75 µg/µl

Table 4: Summary of protein samples isolated and prepared from cell cultures.

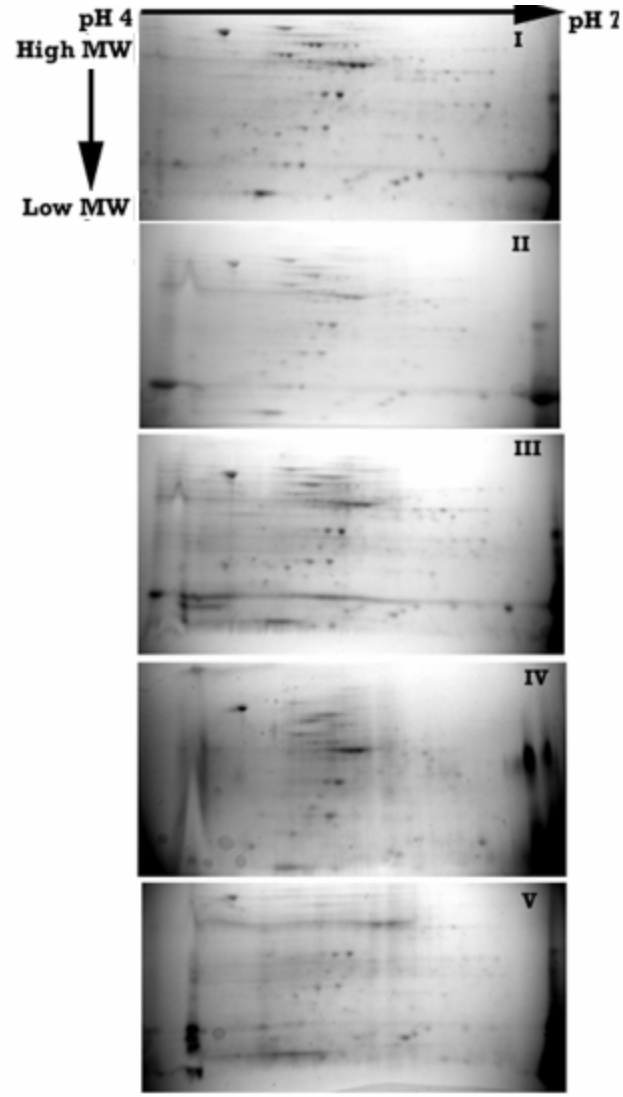


Figure 12: The five displayed gels represent the typical protein expression observed at each time point throughout the mid-log phase of growth. All of the gels have been oriented in the same fashion with a pH of 4 on the left progressing to a pH of 7 on the right. The high molecular weight proteins are located on the top of the gel while the lower molecular weight proteins migrate further down the gel. The top picture (labeled I) is the protein expression observed after 5 hours of growth. Each consecutive gel represents the protein expression observed after the cells have grown for an additional 30 minutes. (II = 5.5 hours of growth; III = 6.0 hours of growth; IV = 6.5 hours of growth; V = 7 hours of growth)

Each isolated protein sample (displayed in Table 4) was used to create a single gel. Theoretically, the gel created from the 5 hour protein sample collected on 3/21/04 should match the gel from the 5 hour protein sample collected on 11/20/05. To compare these two gel sets a method of intra-set pairwise analyses was employed. Each gel was compared to the next gel in the time series; for example the 5 hour gel created from the protein isolated on 3/21/05 was compared to the 5.5 hour gel from 3/21/05. Then the same time points from the 11/20/04 gel set were also compared in the same fashion. The

comparison of gels was based on two factors. The first variable being compared is the percent of spots, out of the total spots on the gel that can be identified or matched on both gels. The second factor was the coefficient of variation (%CV), which is the standard deviation expressed as a percentage of the mean, of the normalized volume of each matched spot on the gel. The normalized volume was chosen as the variable of evaluation because it most directly correlates to the expression of the protein or spot intensity. If the position of the protein spots on the gels being compared match well with each other then the percentage of spots that match within the two should be high. The %CV of the normalized volume can be used to evaluate the intensity of these matched spots and if the %CV is low then the spot intensity, or expression, on each gel is similar. After all the gels in both sets were analyzed based on these two variables the results were cross compared between the two sets to determine an average %CV and % matched, with standard deviation. (see figure 13)

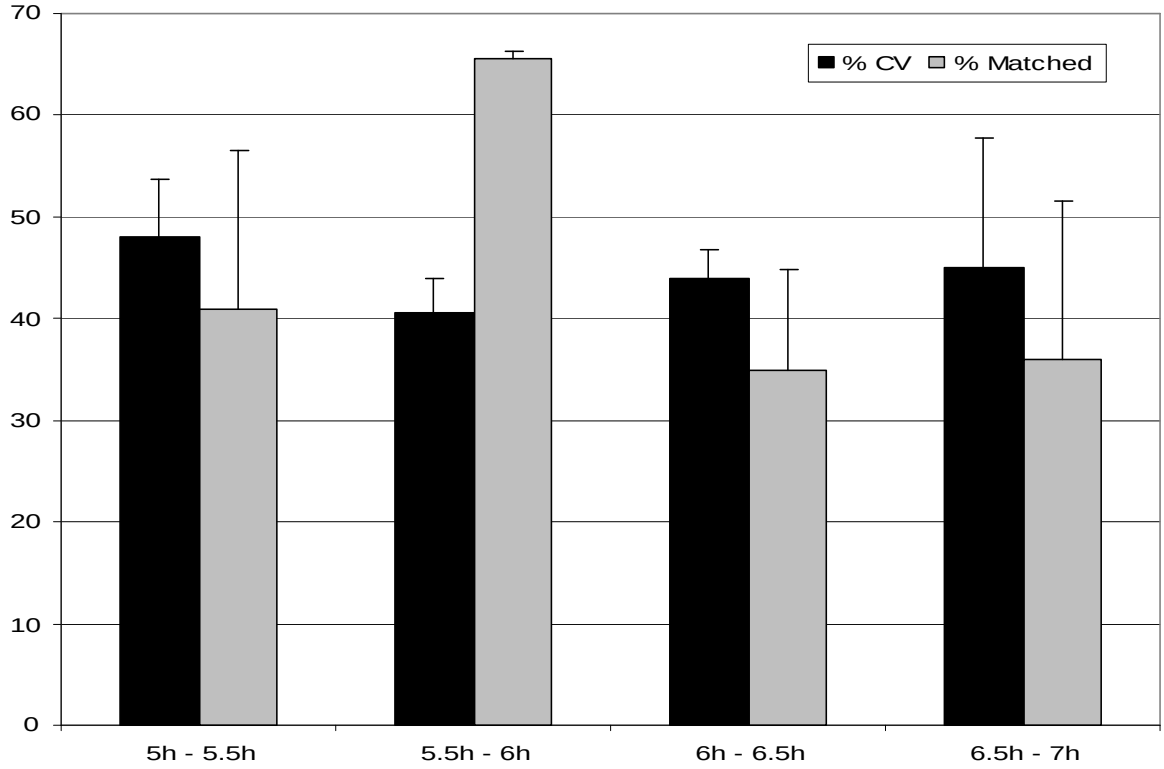


Figure 13: Results of the gel to gel analysis within the two sets. The x-axis is the time points that are being compared and the y-axis is percent.

The %CV between most of the compared gels is fairly similar, ranging from %CV = 40.5 to a %CV = 48%, indicating that within a 30 minute window the protein expression may vary considerably. However, when looking at the % of matched spots between the two gels the 5.5h to 6h gels showed a significant increase (65.5% of the spots on the gels matched) compared to all other gels, indicating that the spots expressed on each gel are in similar locations. This is significant because the mid-log phase falls within this window of time. Another major result of this analysis is that when comparing

the two most extreme gels (5h to 7h) the %CV stays around that same level observed in all of the other gels (%CV = 45%) and the percent of matched spots is also similar to the other comparisons (36% of the spots on the gels matched). With the exception of the percent of matched spots between the 5.5h and 6h gels, the other gel comparisons were quite similar when evaluated on these two criteria and no true conclusions can be drawn at this time.

After performing this intra-set gel analysis it became even more apparent that in order to truly evaluate these gels a standard should be developed that can be used as a universal standard to which all gels, from both sets, will be compared to. These observations helped to solidify the importance of the second aim of this study – the creation of the universal gel standard. To create the universal standard gel a single protein sample was isolated from a cell culture of P. putidaKT2440 grown to mid-log phase on 5mM succinic acid as a carbon source. This protein sample was used to make three replicate gels that were analyzed using Phoretix 2D. We decided to use three gels to make the universal standard gel because previous data has shown that using 3 - 4 gels leads to optimal reproducibility and increasing the number of replicate gels to beyond four gels will decrease the reproducibility and matching between gels. **[11]** Also, in another study conducted by Molloy and Brzezinski a universal standard gel was created from three replicate, or sibling gels, and used for quantitative analysis. **[13]** These gels were analyzed in the normal manner and then used to create an average gel by following tutorial 7 in the Phoretix 2D user manual. Two different versions of the average gel were created. (See figure 14) The first average gel includes spots that were automatically detected by the program and have undergone minimal manual editing. It also includes

spots that are present in 2 out of 3 gels. The second version of the average gel includes only the spots that are present in all 3 of the gels and was subjected to slightly more manual editing that will be further explained.

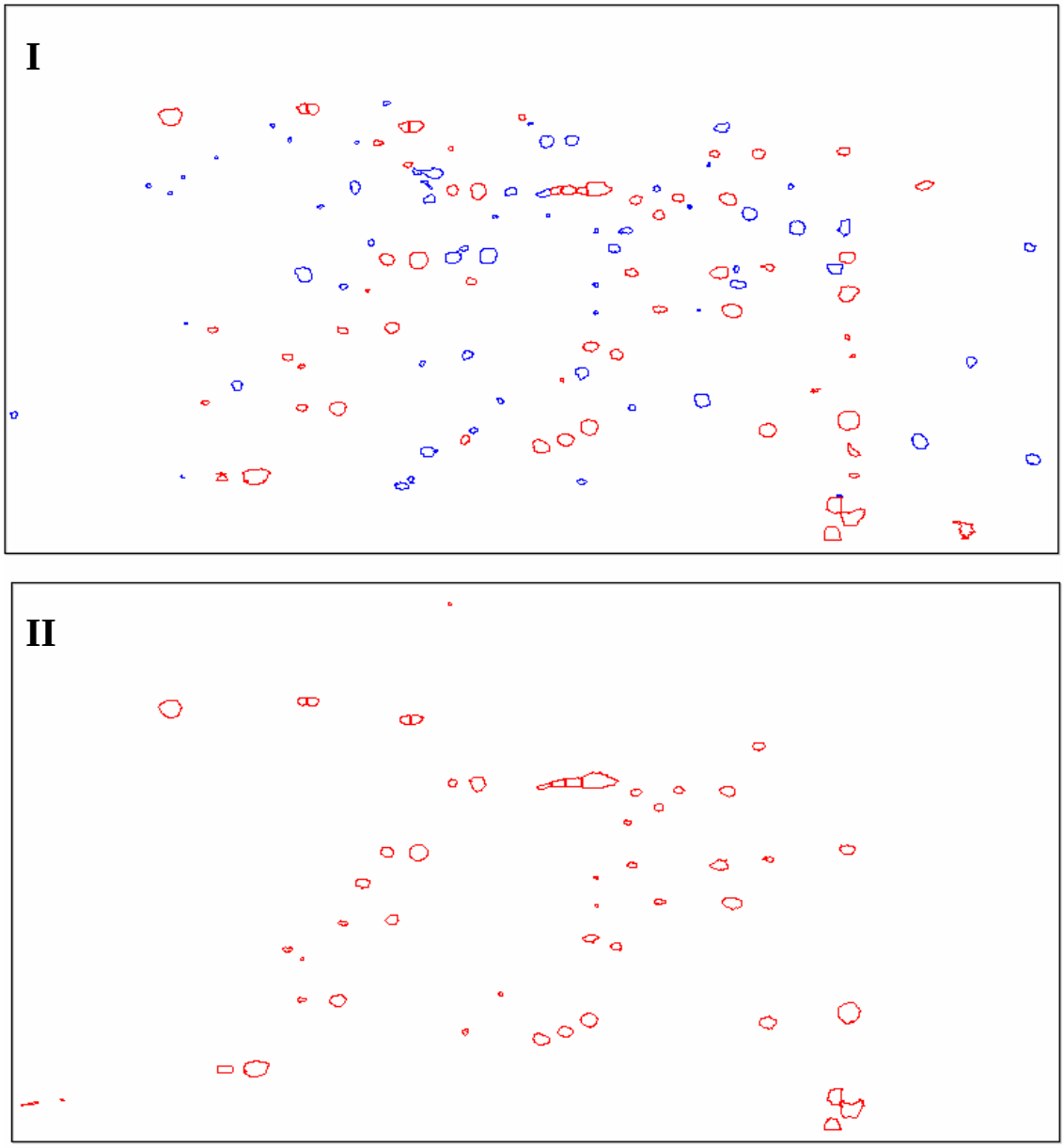


Figure 14: The first representation of the universal standard gel (I) includes spots that are present in 2 out of 3 or 3 out of 3 gels. The blue circles outline the spots that are present in two out of three gels and the red circles indicate spots that are present in all

three gels. The second representation of the universal standard gel (II) shows only spots that are present in all three gels.

Both of the above representations will be useful to future experiments because they set minimum and maximum requirements for matching. The second representation, that only includes spots present in all three gels, displays the minimum number of spots that a future control gel must match. If the control gel contains these spots the investigator can conclude that any changes of protein expression observed in sibling gels are significant and are not caused by variation in sample preparation technique.

Some of the spots that were present in all three gels were removed from the second representation after considering the spot histograms provided by the analysis software. When making the final decision regarding which spots to include in the minimal requirement gel spots were evaluated based on three criteria: normalized volume, area, and circularity of the spot. The normalized volume is the best representation of the protein expression and the area of the spot is another variable that can be used to evaluate protein expression. Based on personal experience, the circularity of the spot helps to determine 'true' protein spots from impurities, bubbles, contamination, or streaks present on the gel. 'True' protein spots are usually circular, whereas the above mentioned impurities are generally not as circular. Phoretix 2D analyzes every spot based on these selected criteria and displays the data in the form of histograms. (See figures 15 – 18)

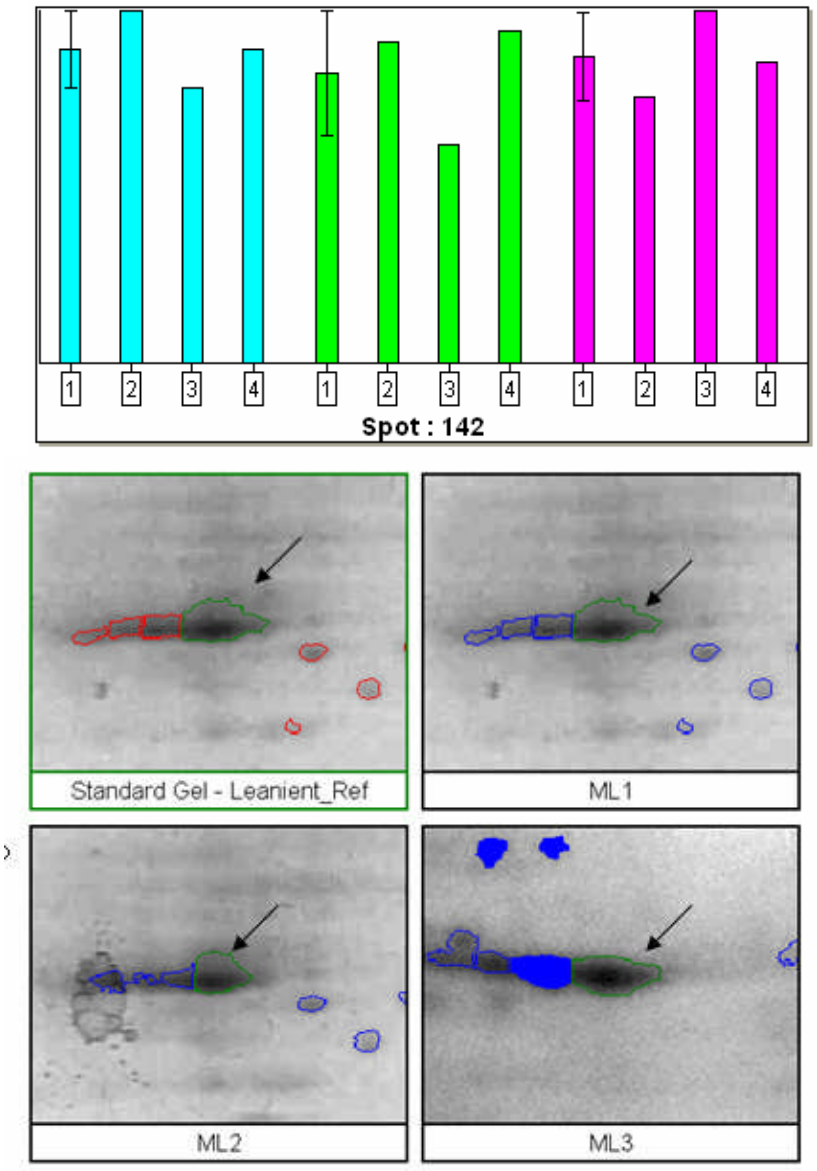


Figure 15: Histograms and montage view for spot 142. On the histogram the blue bars are the normalized volume, the green is the area, and the magenta is the circularity of the spot. The numbers underneath each bar indicate which replicate gel the spot is on. The first bar in every category is the average of the spot measurements (with standard deviation) on the three replicate gels. The montage window shows the spot of interest (indicated by the arrow) on all three replicate gels and the universal standard gel.

Based on the three criteria of evaluation spot number 142 is consistent on all three gels and the error bars present on the histograms indicate low variation of the spot across the replicate gels. Therefore, the spot is included in the universal standard gel. Most of the spots displayed on the minimum requirement version of the standard gel show spot analysis of similar quality to spot number 142. Spots that showed great variability across gels, like spot number 75 shown below, were removed from the standard gel.

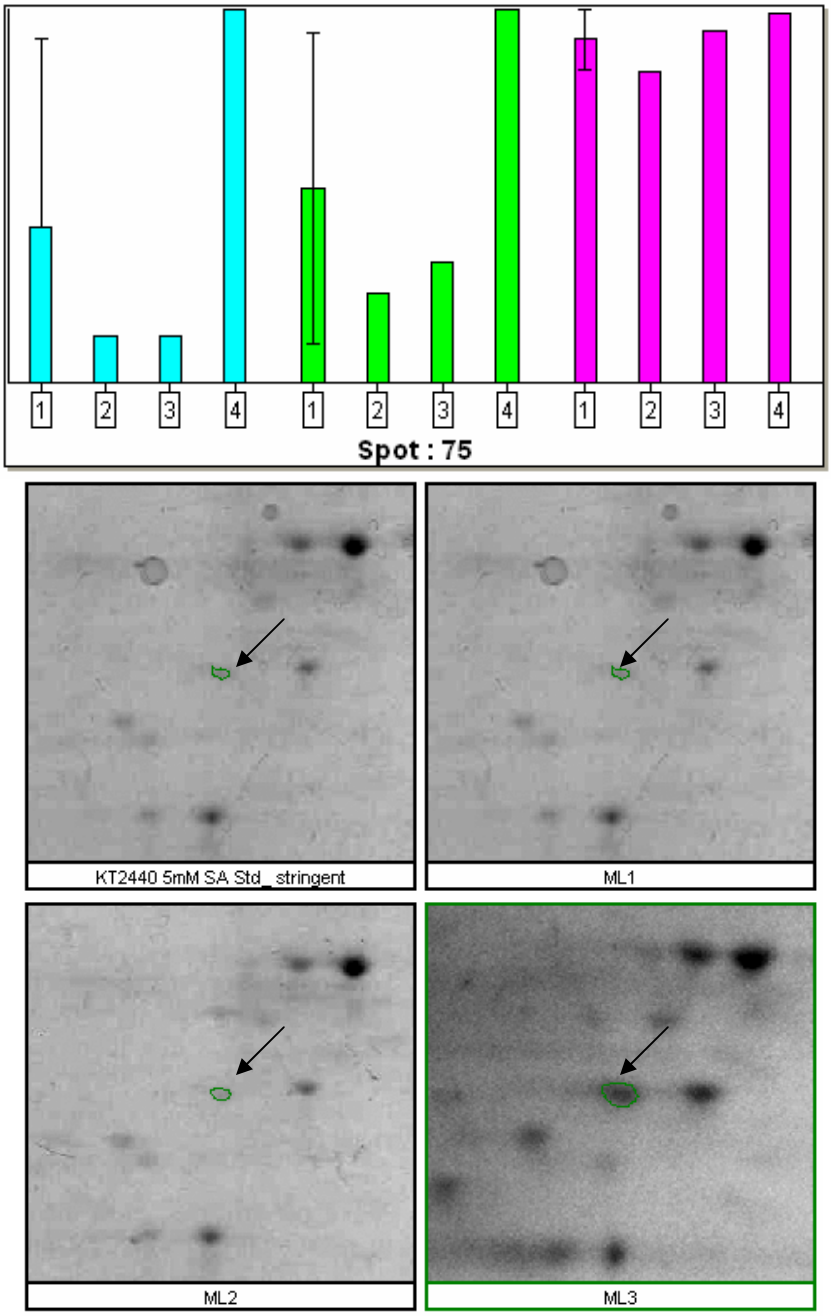


Figure 16: Histograms and montage view for spot 75. On the histogram the blue bars are the normalized volume, the green is the area, and the magenta is the circularity of the spot. The numbers underneath each bar indicate which replicate gel the spot is on. The first bar in every category is the average of the spot measurements (with standard deviation indicated by the error bars) on the three replicate gels. The montage window shows the spot of interest (indicated by the arrow) on all three replicate gels and the universal standard gel.

Spots like number 75 depicted above can easily be identified by the large error bars present on the histogram. These error bars indicate that the spot deviates from gel to gel. Another interesting point can be made in reference to the circularity of the spot. Although the expression levels of the spot changes from gel to gel the circularity of the

spot remains consistent, this indicates that the spot is truly a protein spot and not an impurity or contaminant. Spot number 87, which shows significant deviation of circularity from gel to gel, is a prime example of a contaminant. (See Figure 17) This contaminant is present on two out of three gels and was considered a protein spot by the software. After manual examination of the circularity the spot is removed from the gel.

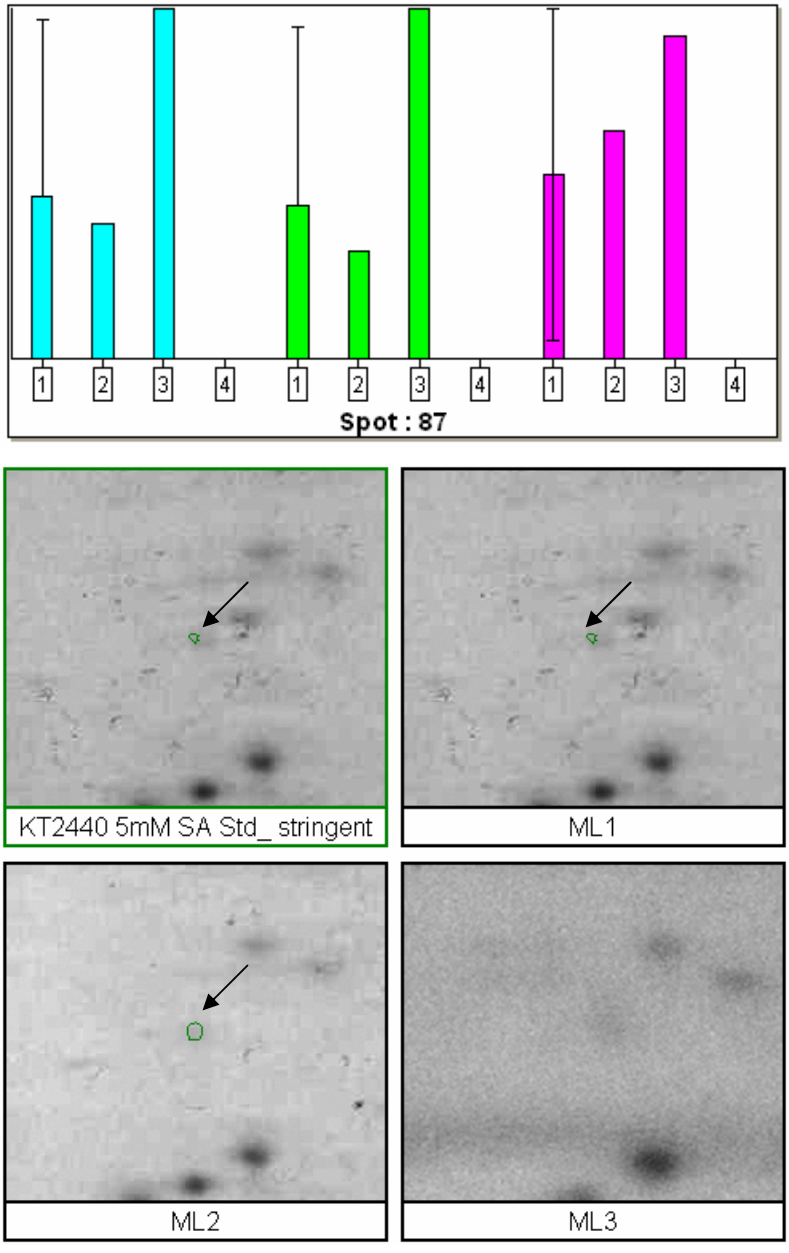


Figure 17: Histograms and montage view for spot 87. On the histogram the blue bars are the normalized volume, the green is the area, and the magenta is the circularity of the spot. The numbers underneath each bar indicate which replicate gel the spot is on. The first bar in every category is the average of the spot measurements (with standard deviation indicated by the error bars) on the three replicate gels. The montage window shows the spot of interest (indicated by the arrow) on all three replicate gels and the universal standard gel.

The first representation of the universal standard gel also took these three criteria into account when deciding what spots to include in the gel, but they were not as thoroughly evaluated. This gel also includes spots that are present in only two out of three gels. Certain spots, like spot number 40, that are present in only two of the three gels should still be present in the universal standard gel because their expression is consistent across a majority of the gels and they are 'true' protein spots, but for whatever reason they are not present on one of the gels. (See figure 18) Again, spots of this nature are easily identified by examining the spot histograms. This type of spot shows low error bars but no bar for one of the gels.

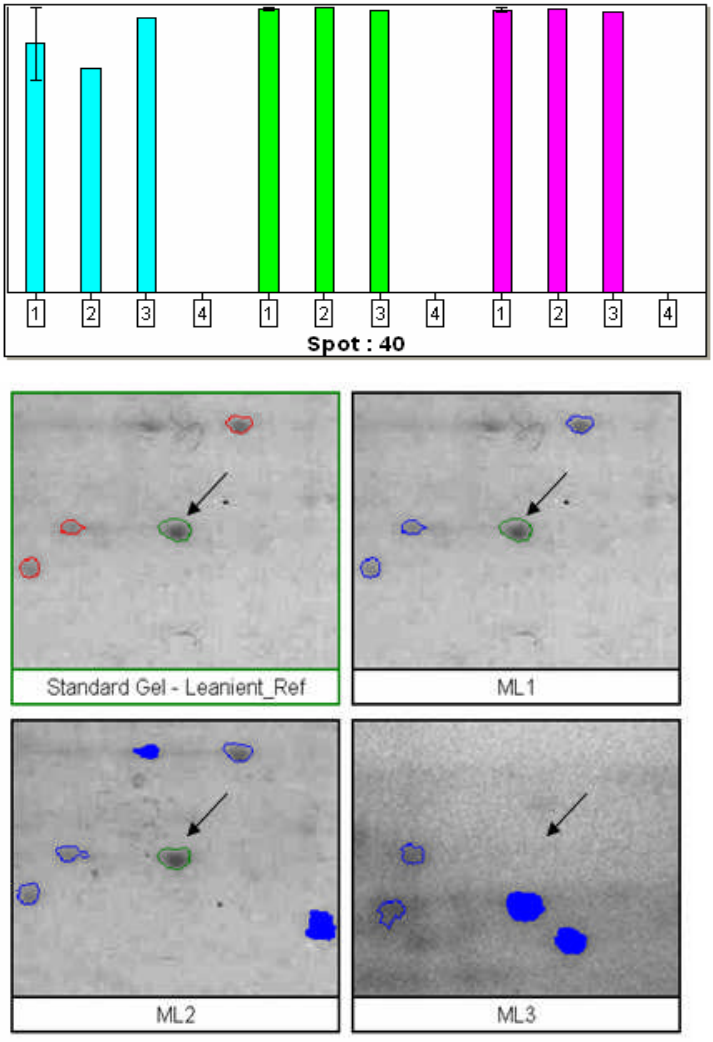


Figure 18: Histogram and montage window of spot number 40. In the histogram, notice that the spot is not present in the fourth bar and the spot is missing from mid-log 3 gel.

Once the universal standard gel was completed it was critical to determine if a control gel would actually match the standard and to establish acceptable statistical limits of matching. To test these aims a co-investigator (Archana Pandey) isolated and purified protein from P. putida KT2440 grown to mid-log phase on 5mM succinic acid and separated the sample through IEF-SDS-PAGE, as previously described. The resulting gel was compared to the universal standard gel and evaluated based on the %CV and the percent of spots that matched within the two gels. It was found that the %CV was 33% and the percent of matching spots were 54% out of the total spots present on Padney's gel.

Previously reported studies found that gels created from identical samples led to %CV values ranging from 19 to 26% and gels from different preparations of the same sample resulted in 24 – 47% variability. **[12]** Based on these previous studies, it can be assumed that future control gels can be quantitatively compared to the universal standard gel. If the future control gels fall within this statistical range for %CV value (24 – 47%) then it can be concluded that the gel agrees with the universal standard gel and that differences in protein expression within the sibling gels are significant and not attributed to technical variability.

Based on these findings the two gel sets that were previously discussed can now be quantitatively compared to the universal gel standard. For this analysis only the second version of the universal standard gel, the minimal requirements gel, will be the basis of comparison. This analysis was based on three criteria; the first two being the coefficient of variation and the percent of spots that could be matched between the gels. The third variable is evaluating intensity changes within the matched protein spots (based

on normalized volume) to determine quantitative changes in the spot intensity and therefore the protein expression. This variable is commonly employed in the field of proteomics. **[11]** Most software packages automatically highlight any protein spots that exhibit a twofold change in their expression between different gels and because of this fact it has become common practice to use this twofold change as a cutoff for determining the significance of a quantitative change. This type of change within a spot is identified without difficulty by using the Phoretix 2D software. After the gels have been analyzed, the program sets the value of every spot in the standard gel equal to 1.00 because it is being compared to itself. It then calculates the increase or decrease in expression of the spots on the other gels and displays this data in an "Expression Table". (See figure 19). This table highlights spots that fall above or below user defined changes of expression.

Ref. Spot	Standard G...		Gel A2		Gel B 11-20-04		Gel C 11-20-04		Gel D 11-20-04		Gel E 11-20-04	
	Volu...	Nor...	Volu...	Nor...	Volu...	Nor...	Volu...	Nor...	Volu...	Nor...	Volu...	Nor...
101	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
9	1.000	1.000	1.940	-1.195	-1.897	-1.091	-3.418	-1.550	1.167	1.895	1.074	-1.415
142	1.000	1.000	1.408	-1.338	-1.276	1.675	-	-	-10.4...	-5.236	1.713	1.386
139	1.000	1.000	2.543	1.250	-2.339	-1.181	-1.558	1.612	-	-	-2.578	-3.438
143	1.000	1.000	1.940	1.030	-	-	-1.560	1.739	-14.4...	-7.234	1.349	1.093
91	1.000	1.000	1.792	1.419	-16.3...	-5.130	-	-	-	-	1.147	1.386
96	1.000	1.000	1.970	1.886	-33.6...	-8.738	-	-	-	-	-	-
53	1.000	1.000	-1.037	-1.946	-3.884	-1.808	-2.836	-1.041	-	-	2.426	1.972
88	1.000	1.000	1.709	-1.523	-2.171	-1.402	-1.204	1.630	-	-	-	-
89	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
55	1.000	1.000	3.137	2.222	-3.577	-1.253	-7.271	-2.008	-	-	-2.240	-2.074
134	1.000	1.000	2.582	1.204	-1.084	1.733	-3.256	-1.366	-5.291	-3.015	-	-
93	1.000	1.000	1.870	1.562	-7.689	-2.284	-	-	-	-	-	-
56	1.000	1.000	1.091	-2.179	-1.100	1.541	-	-	-	-	-	-
75	1.000	1.000	1.646	1.071	-2.825	-1.077	-	-	-	-	-	-
40	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
77	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
137	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
37	1.000	1.000	-1.422	-5.072	1.606	1.814	-1.448	-1.011	-	-	-	-
79	1.000	1.000	1.000	1.000	-	-	-	-	-	-	-	-
42	1.000	1.000	1.398	-1.601	-	-	-1.661	1.375	-	-	-	-
54	1.000	1.000	-	-	-1.736	-1.911	1.386	1.597	1.038	-1.136	-	-
71	1.000	1.000	-	-	-1.389	-1.628	1.280	1.386	-	-	-	-
58	1.000	1.000	-	-	-1.377	-1.613	1.274	1.380	-	-	-	-
61	1.000	1.000	-	-	-1.119	-1.078	-	-	1.106	1.072	-	-
144	1.000	1.000	-	-	-1.087	-1.220	1.507	1.703	-1.743	-2.095	-	-
41	1.000	1.000	-	-	-	-	-	-	-	-	-	-
145	1.000	1.000	-	-	-	-	1.000	1.000	-	-	-	-

Figure 19: This is an example of the expression table generated for a twofold change in expression when compared to the universal standard gel. The numbers displayed in the expression table indicate the spot intensity, relative to the standard gel spot, of the matched spots in each individual replicate gel.

In the expression in figure 19 the first column is the reference spot number and the next column is the standard gel. The remaining gels are from the set that was isolated on 11-20-04. "Gel A2" is from the protein sample that was isolated at 5 hours, "Gel B" is from the 5.5 hour protein sample, "Gel C" is the protein sample from 6 hours and so on. In the columns you will notice spots that are highlighted in yellow and green. If a spot is highlighted in yellow it has shown at least a twofold decrease in protein expression and if it is highlighted in green the spot has shown at least a twofold increase. Boxes that do

not contain any value indicate that that particular spot is not present on the gel. For example, if you look at spot number 134 (indicated by the black box) at five hours of growth time its expression is increased 2.5-fold. Then the expression slowly begins to decrease on the 5.5 hour and 6 hour gels and on the 6.5 hour gel the spot has show a 3-fold decrease in expression. Ultimately, after 7 hours of growth the protein is no longer being expressed which is indicated by the empty box in the Gel E column. Spots that increase or decrease their expression over time are quite interesting and should be noted for later analysis and possibly identified by mass spectrometry.

The gels from both sets of protein were compared to the universal standard gel and the percent of the matched spots that demonstrate a minimum of twofold and a minimum of fivefold change was calculated. If the gel shows a high percentage of spots that exhibit these changes it can be concluded that the protein expression on this gel deviates from the protein expression on the universal standard gel. The isolated protein samples that had similar growth time to the mid-log phase of growth should demonstrate similar protein expression to the universal standard gel, and vise versa. The table below (Table 5) shows the percent of spots in each gel that display either a twofold or a fivefold change in expression. This data is also displayed graphically in figure 20.

	Twofold Change				Fivefold Change			
Growth Time	**21-Mar**	**20-Nov**	**Avg.**	**Std. dev.**	**21-Mar**	**20-Nov**	**Avg.**	**Std. dev.**
5	16.0	14.3	15.2	1.2	4.0	4.8	4.4	0.6
5.5	14.3	13.0	13.7	0.9	3.6	8.7	6.2	3.6
6	9.5	9.5	9.5	0.0	0.0	0.0	0.0	0.0
6.5	27.3	50.0	38.7	16.1	27.3	25.0	26.2	1.6
7	41.7	30.0	35.9	8.3	16.7	10.0	13.4	4.7

Table 5: The percent of spots in each gel that displayed a twofold or a fivefold change in normalized volume when compared to the universal standard gel.

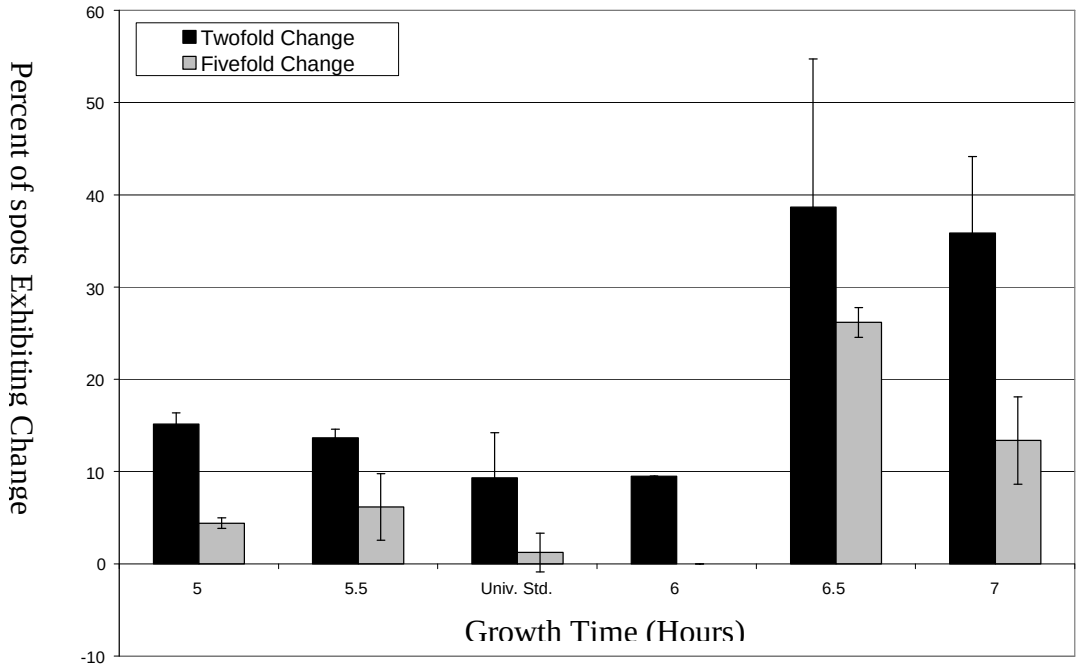


Figure 20: The percent of spots on the gels that exhibit a twofold or fivefold change in their normalized volume.

The three gels that were used to make the universal gel standard were also compared based on this variable of fold-change and the results are shown in figure 20. It is very intriguing that the percent of fold-changing spots on the 5.5 hour gel samples and the 6 hour gel samples both agree, within error, with the percent of fold-changing spots in the three gels that created the universal standard gel. Another crucial observation is that the longer the growth time the more deviant the protein expression on the gel becomes when compared to the universal gel standard. The comparison of the gels based on the %CV and the percent of matched spots showed similar trends in the data as shown in figures 21 and 22.

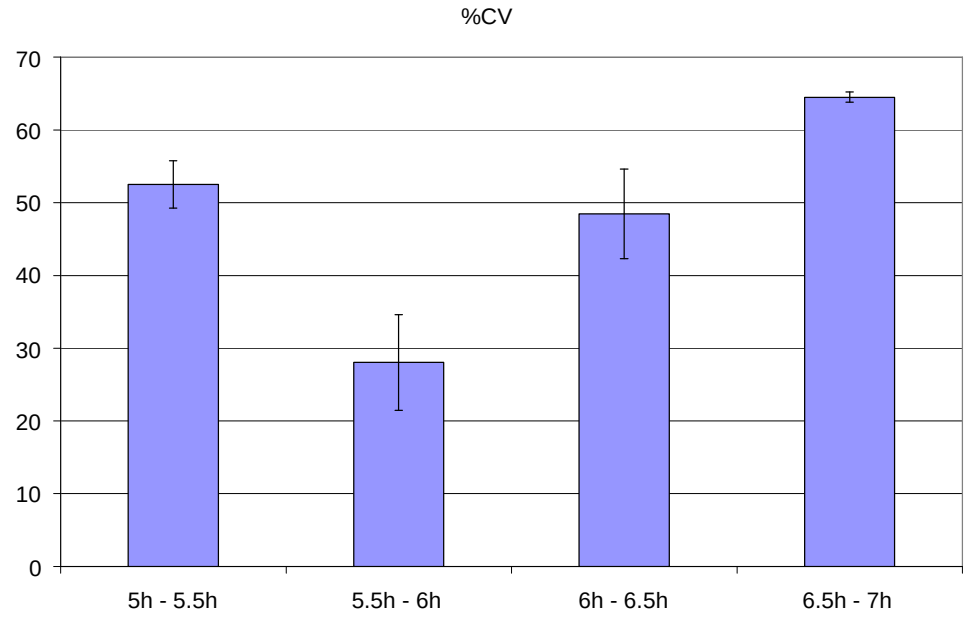


Figure 21: The % CV of the gels compared to the universal standard gel. The x-axis are the two gels that are being compared to the standard gel and the y-axis is the %CV based on the normalized volume.

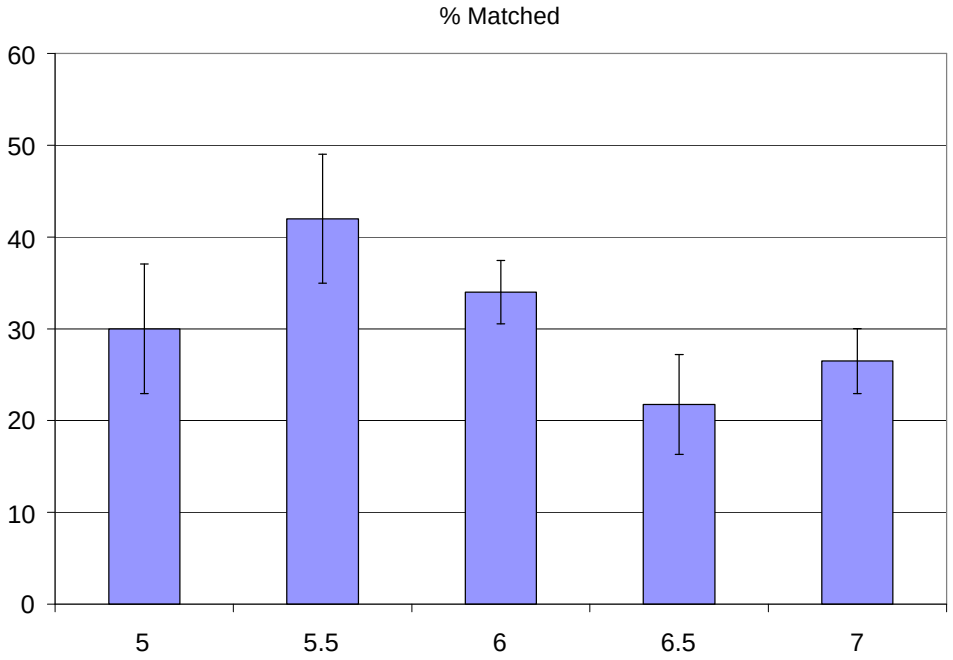


Figure 22: The percent of spots out of total spots on the gel that could be matched to the universal standard gel. The x-axis is what time point (in hours) is being compared and the y-axis is the percent of matched spots.

The reader may notice that in figure 22 the percent of matched spots within the gels are fairly low values. These values are still acceptable because this number is the percent of matched spots out of the total number of spots on the sample gel, not the universal standard gel. In the future it may be beneficial to calculate the percent of matched spots based on the spots present in the universal standard gel. By calculating the variable based on the number of spots in the universal standard gel, the percent of matched spots should be greatly increased. Also, this proposed method of calculation will allow future investigators to establish a value for the percent of matched spots that all replicate gels must agree with. For example, any future gels must match 95% of the spots present on the universal standard gel to be within acceptable statistical limits of matching.

These values of %CV and the percent of matched spots in a gel are different than the values displayed in figure 13, because this experiment uses the universal gel standard as the reference gel in the analysis. In the first analysis, the reference gel was assigned by the software, which automatically makes the gel that contains the most spots the reference gel. This comparison to the universal standard gel is more accurate because it compares three gels instead of two. Both of these variables still agree with the trends observed in the first intra-set comparisons. The percent of matched spots are highest at 5.5 and 6 hours, indicating that the locations of the spots on the gels are similar to each other. The %CV value is again lowest when comparing the gels that had similar growth times to mid-log phase to the universal standard gel.

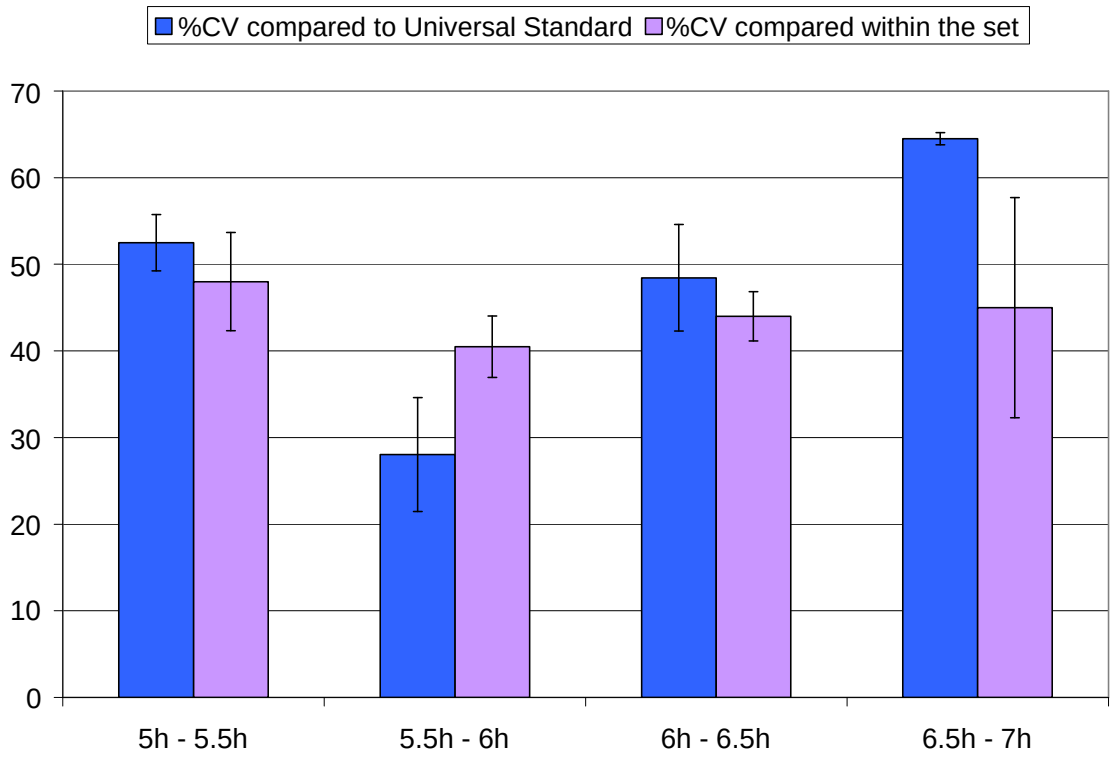


Figure 23: Comparison of the intra-set analysis and the gel sets compared to the universal gel standard.

These data agree with the original trends observed in the intra-set gel comparisons (see figure 23), but go one step further. These new data show that when comparing the 5.5 hour gel and 6 hour gel to the universal standard gel, the coefficient of variation is only 28 ± 6.6%. This value is statistically lower (within error) than the originally calculated value for this timeframe. Another observation relating to this piece of data and the entire study, is that Challapali et. al. **[12]** concluded that protein samples generated from different preparations result in a 24 – 47% coefficient of variation. The observed 28% CV value also agrees with this conclusion.

The three criteria of two-fold or greater spot intensity change, %CV, and percent of matched spots all show some significant correlations trends of the changes in protein expression throughout the log-phase of growth, specifically when compared to a

universal standard gel that represents protein expression at the mid-log phase. (See table 6). To easily interpret the data keep in mind how all the variables correlate to each other and the question that each variable is designed to assess. The twofold change variable is commonly used in proteomic data analysis and is designed to indicate an increase or decrease in the expression of a protein spot across a set of gels. The fivefold change variable is simply an extension of this variable and is designed to point out extreme changes in protein expression. A lower value for the fold-change variables indicates that the protein expression on these gels more closely match the protein expression on the universal gel standard. The percent of matching spots is a similar variable that assesses the location of spots on the gels and the similarity of the 'pattern' of protein expression on the gels. A higher percentage of spots that can be matched between the gels the better the 'patterns' on the gels match each other. The final variable, and perhaps the most conclusive, is the coefficient of variation between gels. Again, this variable is the standard deviation (expressed as a percentage of the mean) of the normalized volume of each matched spot on the gel. The %CV considers all of the expression levels of all of the matched spots on the gels and determines the overall variation of protein expression on the gels. The lower the %CV value the better the protein expression matches the expression at mid-log phase. By evaluating the protein expression based on these three variables a clear relationship of the protein expression throughout the log phase of growth can be established.

Growth Time (hours)	% of spots exhibiting a twofold change	% of spots exhibiting a fivefold change	% of matched spots	% Coefficient of Variation	
5	15.15 ± 1.20	4.4 ± 0.57	30 ± 7.1	**5–5.5h**	52.5 ± 3.3
5.5	13.65 ± 0.92	6.15 ± 3.61	42 ± 7.0	**5.5–6h**	28.05 ± 6.6
6	9.5 ± 0.00	0 ± 0.00	34 ± 3.5	**6–6.5h**	48.45 ± 6.2
6.5	38.65 ± 16.05	26.15 ± 1.63	21.75 ± 5.4	**6.5–7h**	64.5 ± 0.7
7	35.85 ± 8.27	13.35 ± 4.74	26.5 ± 3.5	-	-

Table 6: Summary of results for comparisons made to the universal standard gel. The numbers displayed are the average values between the two gel sets.

The expected trend for these gels, when compared to the protein expression at mid-log phase, is that the protein expression during the 30 minute window of time surrounding mid-log phase (5.5 – 6 hours) will most closely match the universal standard gel. Along the same lines, the protein expression observed at a time point much later than mid-log phase (7 hours) should significantly deviate. By examining the data at the later time points of 6.5 to 7 hours the predicted trend holds true. The assessment of the three variables during this range shows the highest coefficient of variation, the highest percent of spots that display a twofold or fivefold change in expression, and the lowest percent of spots that match the spots on the universal standard gel. Following with this trend, the second 'most deviant' areas of protein expression are the samples that are within an hour of mid-log phase, both between 5 to 5.5 hours and again between 6 and 6.5 hours. The expression during these windows of time does not differ as greatly as the expression at the farthest extremes of the log phase but it does still correlate to the predicted trends. These observations lead to the conclusion that the protein expression is changing significantly throughout the log phase of growth.

These changes in overall expression, as growth time increases, may be because the cells are approaching the stationary phase of growth and they may be expressing very

different proteins as they reach this phase. During the mid-log phase of growth the cells are very healthy and are actively dividing. In the entire culture there is also a large variety of cells. Some cells are 'young' and have just begun dividing while others have already completed many life-cycles and are 'older' cells. This vast variety of cellular life and the proteins that are being expressed during this time is the key reason why this study is focusing on the mid-log phase of the cell life-cycle. It is not that surprising that the proteins expressed in the end of the log phase of growth are greatly different than the protein expression during the diverse mid-log phase because the cells have all undergone many life-cycles and most are slowly approaching the stationary phase and then cell death.

Statistically solidifying the observation that protein expression changes considerably throughout the log phase of growth is a noteworthy accomplishment that also brings about the questions of how quickly the expression is changing and what statistical limits can be applied to validate true changes in expression. An attempt to answer these questions must begin back at the data generated from the gels that were closest to the mid-log phase of interest, between 5.5 and 6 hours. During this time frame the coefficient of variation was the significantly the lowest, within error. The percent of spots that could be matched to the reference gel was the highest, and the intensities in those spots matched the intensity of the spots on the universal standard gel most closely. This can be concluded based on the fact that the fold-change variables were the lowest during this time. The most striking piece of data associated with this time is the extremely low %CV value (28%). Previous studies have determined that an acceptable %CV of two protein samples prepared at different times is 24 – 47% and that technical

variability resulting from the process of two-dimensional electrophoresis results in a coefficient of variation of 20 – 30% for identical samples. **[13]** The %CV value determined between the four gels isolated at 5.5 and 6 hours, which are technically different samples, falls within these ranges of acceptable variability for identical samples when compared to the universal standard gel. This result indicates that the protein expression within this 30 minute window of time surrounding mid-log phase does not show change within established statistical limits. This finding should be evaluated further by testing the growth and expression of P. putidaKT2440 on different carbon sources, but if it holds true it will reward future investigators with more flexibility during some steps of the procedure. These possibilities will be discussed in greater detail in the conclusion section.

Conclusions

Although the technique of 2DE has encountered problems with reproducibility in the past, it is possible to produce quantitative results, achieve very low levels of technical variation, and determine acceptable limits of biological variation between gel samples. 2DE is a very valuable and dynamic technique because it can be applied to an immense number of proteomic studies. These studies are all quite novel because of the questions they are addressing, the organisms they focus on, the growth conditions and treatment of these organisms, the resources each laboratory possesses, and the analysis strategies employed. Due to these vast differences in application it is crucial for individual laboratories to develop robust sample preparation procedures, establish standard operating procedures for protein purification, isolelectric focusing, and sodium dodecyl sulfate polyacrylamide gel electrophoresis. It is also paramount for each laboratory to develop quantitative measures and standards of intra-laboratory reproducibility. Time must be taken to address each one of these issues before any true comparisons can be made between samples that have been subjected to different treatment or experimental conditions. This study has addressed all of the above mentioned factors that can have critical effects on reproducibility.

I. Development of a Standard Operating Procedure

I developed all of the procedures in the methods and materials section of this work into a standard operating procedure (SOP) that serves as the cornerstone of all the research conducted in the RIT Proteomics lab. The development of this SOP has resulted in increased reproducibility between all the individual members of the lab because it has

alleviated any gross technical variation between investigators. Ultimately, by following the standard operating procedure data obtained by one investigator can readily be compared to another investigator's data and they can be confident that changes in protein expression are true and not attributed to technical variation.

II. Reproducibly Determining Mid-log Phase

Another achievement in this study was determining the time it takes a culture of P. putida KT2440 to reach mid-log phase (5 hours and 35 minutes). This piece of data is significant because of the low standard deviation associated with it (± 2minutes). This was based on six replicate trials, three of which were completed by me and the remaining three were completed by a previous investigator. (Mingy Yao, M.S. degree in Chemistry from RIT, 2005) The high level of reproducibility across all six trials (and between investigators) indicates that the established methods of sample preparation and cell culture techniques produce reproducible results. This finding also acts as a check point, fairly early on in the procedure, that will allow future investigators to assess their cell culture and sample preparation technique to determine if they are adequate. When conducting future growth curve studies investigators can simultaneously grow a "control culture" of P. putida KT2440 on 5mM succinic acid and determine the time it takes for this culture to reach mid-log. If the time is fairly close to the mid-log time established in this study then it can be concluded that the mid-log time determined for the sample cultures is accurate.

III. Changes in Protein Expression Throughout the Log Phase of Growth

One specific step involved in the sample preparation of the protein is stopping cellular growth at the desired time point (mid-log phase). This time point was chosen as the point of interest because the cells are thriving in their surroundings and expressing a diverse array of proteins. The stopping of cellular growth is a factor that is commonly overlooked but may contribute significantly to the observed protein expression, because of the exponential nature of cellular growth. Currently, there have been no studies to determine the required timeframe in which growth must be stopped to conserve protein expression.

The initial pairwise comparison of the two sets of proteins isolated throughout the log phase of growth resulted in fairly high coefficients of variation, ranging from 40.5% to 48%. This indicates that within a window of 30 minutes protein expression is changing considerably. One important observation in this analysis was that the percent of spots that were matched between the 5.5 hour and 6 hour gels was significantly higher (65.5%) than the percent of matches between all the other comparisons. This is interesting because mid-log falls within this 30 minute window of time. A possible conclusion, based on these two observations, is that during the 30 minute window of time surrounding mid-log phase most of the same proteins are being expressed (indicated by the high % matching value), but their expression levels are changing (indicated by the moderately high %CV values).

The second analysis of these two sets of gels differs from the first because it uses the universal gel standard as the reference for all comparisons. The data shows trends that are similar to the initially observed trend. The percent of matched spots is the

highest for the 30 minute window of time surrounding mid-log phase and the farther away from the mid-log time point the more the protein expression deviates from the universal standard gel. The variable of fold-change is introduced in this area of the study and measures the changing protein expression. The %CV values did not show any significant trends in the first analysis, but with the universal standard gel as the basis of comparison some significant conclusions can be made.

The first significant conclusion based on this second analysis is the trend of the %CV graph. (See Figure 21) The coefficient of variation increases as the window of time, relative to mid-log, increases. Focusing on just the 30 minute window of time surrounding mid-log, the %CV is the lowest (28.05 ± 6.6 %). This value is also statistically significant and, more importantly, falls within accepted %CV values for this type of analysis based on previously reported studies. Molloy et. al. defined the degree of technical variation associated with 2DE as CV= ~ 23% for duplicate gels, regardless of sample type. **[13]** They also reported that a pairwise comparison of protein gels isolated from E.*coli* cultures showed CV = 24.0 ± 20.0%. The %CV value achieved in this study not only agrees with Molloy's values but the standard deviation associated with it is significantly smaller (6.6% vs. 20.0%). Another study, by Blomberg and Blomberg, determined that a CV = 20 – 28% for all intra-laboratory results is acceptable. **[8]** A final study, by Lee and Choe, analyzed proteins from *Pseudomonas syringae* and used a cut-off value of CV = 30% to validate and determine significant results. **[11]** The proteins isolated within 30 minutes of mid-log exhibit no changes in protein expression based on accepted coefficients of variation.

Another important observation in this analysis is related to the fold-change variables. When the 5.5h gels and the 6h gels are compared to the three gels that were used to make the universal standard gel, the percent of spots that exhibit a twofold change in expression agree within error. (See Figure 20) The same is true for the spots that exhibit a fivefold change. The significance of this result relates back to the high %CV values (40.5 – 48%) obtained in the first pairwise comparison analysis. When the 5.5h and 6h gels were only compared to each other, without the universal standard gel to act as a reference, these fold changing spots may have contributed to the high %CV values. For example, if a spot on the 5.5h gel had an expression value of +1.5 and the same spot on the 6h gel showed a decrease in expression to -1.5, when these spots were compared the total expression change is 3.0. This value is above the twofold change in expression and will result in a high %CV value for that spot. When the same spot is compared to the universal standard gel (whose expression is standardized to ±1.0) the spot on the 5.5h gel now shows a +0.5 increase in expression and the spot on the 6.0h gel shows a -0.5 decrease in expression. When compared in this fashion the absolute value of the change in expression between the two spots would be 1.0, which is far less significant and will show a much lower %CV value. This observation leads to a logical explanation of the differences in %CV values between the first analysis and the analysis that used the universal standard gel as the reference for all comparisons. These conclusions also emphasize the importance of developing a standard gel that will be used as a universal reference across all gel sets and all future experiments.

IV. Development of the Universal Standard Gel

The importance of creating a universal standard gel has already been emphasized and demonstrated by the conclusions drawn in part III above. While creating this gel we had to make a decision about which variables to assess when determining the spots that should be included in the standard gel. The variables chosen in this study were normalized spot volume, area and circularity of the spot. These three variables are a solid basis for determining what spots should be included. We also had to decide how many replicate gels should contribute to the universal standard gel and how many gels must contain a particular spot before we could to add it to the reference. Literature has shown that 3 – 4 replicate gels are sufficient, and that spots contained in 'n-1' gels should be added to the standard.**[11]**

In this study the universal standard gel was created from three replicate gels and was displayed in two different fashions. For statistical analysis one of these representations proved to be more useful. The universal standard gel that contained spots that were present in three out of three gels and was subjected to a more thorough manual analysis should be used as the reference gel for future analyses. When this gel was compared to a gel produced by Archana Pandey (a newer member of the RIT Proteomics lab) the coefficient of variation between her sample gel and the standard was 33% and 54% of the spots on the sample gel were matched to the universal standard gel. This comparison leads to two conclusions or recommendations. The first is that the developed universal standard gel is functional and future investigators should be able to produce gels that will match it within technical variation. The second conclusion establishes standards of intra-laboratory reproducibility based on this comparison and literature

values. When future investigators compare their control gels to the universal standard gel an acceptable %CV value should be less than 35%. This suggested range is more stringent than the acceptable limits established by Challapalli et al who suggested that gels produced from different sample preparations show 24 – 47% CV. **[12]** I believe that the limit for our analysis should be more stringent than this suggested range because of the simplicity of the organism of interest. *Pseudomonas putida* is a bacteria and not nearly as complex as some of the samples involved in that study such as blood serum isolated from humans. The number of proteins expressed by *P. putida* is far less than those present in human blood samples and therefore should lead to less complex samples and higher reproducibility. Another argument for the range of this standard is the %CV value reported by Molloy et al in their study of *E.coli*. They found that inter-experiment gels showed CV = 31.2 ± 19.9%. In conclusion, the universal standard gel has already been an invaluable resource and future experiments should run a control gel that can be compared to this standard and result in %CV values less than 35%.

V. Future Directions

The current and future research conducted in the RIT Proteomics Laboratory is extremely valuable and on the brink of producing cutting-edge results. With the results of this study forming a solid quantitative basis for future analyses, this research will rival most currently published data in the field of proteomics. The long term goal of this research is to develop a screening method for suspect carcinogens and mutagens that will compliment or replace the Ames Test. Currently, the Ames Test is the only method to test mutagenicity without using animal testing and needs to be supplemented because it

produces many false positives and false negatives. This novel screening method will compare the protein expression of *P. putida* grown in the presence and absence of known carcinogens and determine biomarkers of carcinogenicity. Not only will it be able to determine if a chemical or substance is a carcinogen or mutagen, it will be also be able to assign a danger rating to the chemical. If, for example, 10 biomarkers for carcinogenesis are identified, and a chemical shows 8 out of 10 biomarkers it is a high risk carcinogen, but if it only displays 3 out of 10 established biomarkers then the danger rating will be significantly lower and the chemical is less hazardous. The development of this type of screening method may still be years away, but with the quality of research currently being conducted and the enthusiasm of the student researchers and advisors it will quickly become a reality.

The completion of this quantitative study of *Pseudomonas putida* KT2440 by 2DE is a modest step toward the overall goal of this research but will prove to be extremely valuable for future analyses, and can offer some insight to future investigators. This study began by developing a standard operating procedure that is followed by all members of the research laboratory. As our research grows and the science of 2DE advances, this SOP will continually need to be updated. It has already improved the quality of research by alleviating technician variability, and improving overall reproducibility. The process of editing and updating the SOP may be tedious and unexciting but is crucial to the success of new and future members of the lab, especially in the quarter system of RIT (where students often have time for research one quarter but do not the next). If the SOP is current, and regularly maintained, it will be easy for these

students to “pick up where they left off”, which will make their time spent in the lab much more productive.

The next achievement of this analysis was the reproducible growth curves of P. putida KT2440 on 5mM succinic acid. If future investigators grow a control culture (P. putida KT2440 on 5mM succinic acid) simultaneously with the experimental growth curve samples, and the control matches the data presented here, then the experimental growth curves are also accurate. Ultimately, this may decrease the number of trials one must complete to be confident in their results. This point is becoming increasingly relevant as the carbon sources and growth conditions become more complex. Our group has already encountered growth situations, like the growth of P.putida KT2440 on 9-amino acridine, where the mid-log phase was found to be approximately 9 hours and a total growth curve takes 18 hours to reach stationary phase. This process is extremely time consuming and less repetition is greatly desirable. Establishing reproducible growth curves can never be eliminated, but if the investigator had a control culture and at least three replicate sample cultures that show consistent results, then the process may only have to be conducted once instead of multiple times.

Along these same lines, this study concluded that within 30 minutes of mid-log phase protein expression does not change within statistical limits, when compared to the universal standard gel. If a set of replicate growth curves has been completed, the standard deviation of the calculated mid-log time is ±15 minutes, and the control culture matches the established growth curves, then the results are adequate and no further trials need to be conducted. The actual window of growth time in which protein expression is unaltered may be slightly longer than ± 15 minutes, but this recommended range for

stopping growth errs on the side of caution. This recommended range is also more than enough time to complete the necessary steps associated with stopping cell growth. These findings should help eliminate unnecessary trials and save valuable time.

Another insight into improving overall reproducibility relates to the current approach to sample preparation. In this study 'replicate' gels were created from two sets of proteins that were isolated at different times from different cultures. In the future it will be much more valuable, and create more reproducible results, if one large stock of a protein sample is created and 3 – 4 replicate gels were created from that stock, all undergoing IEF and SDS-PAGE at the same time. Simple factors, such as changes in room temperature from day to day or the lots of chemicals present in the lab at a particular time, may significantly affect the reproducibility of the samples. If all of the necessary gels for an entire study originate from the same protein sample and the gels are prepared on the same day it will remove some of these confounding variables. For some of the longer and more involved studies it may not be possible to run all of the gels at that same time. If this is the case, then a large stock of protein should still be created from a single culture and after the concentration of that protein is determined it should be aliquoited into sterile microfuge tubes containing enough protein to run at least three replicate gels. This will prevent repeated freeze-thaw cycle from occurring and preserve the integrity of the protein sample for longer periods of time.

A final suggestion for future analysis is to determine a standard operating procedure relating to the Phoretix 2D analysis software. This SOP should include acceptable limits for spot detection, warping requirements, spot matching requirements, manual editing, and methods of background subtraction and spot normalization. The

Phoretix 2D software is a very powerful tool for quantitative 2DE analyses, but can easily be manipulated by the user to produce the desired results. It may be worthwhile to assign a single person the task of completing all of the quantitative analysis using the software. The gel images that this person was given to analyze should also be coded, so the analyzer would be blinded to the study and their data would not be biased. At the very least, some intra-laboratory guidelines must established that will limit the users influence and avoid these human pitfalls.

The next priority of this research should be to establish a standard gel for protein expression of P. putida KT2440 on a carcinogen. This standard gel will represent the 'stressed' or 'diseased' state of the organism and can be compared to the universal standard gel of the 'healthy' state (P. putida KT2440 grown on 5mM succinic acid). P. putida KT2440 should be grown on a variety of carcinogenic carbon sources and the protein expression analyzed. If a distinct proteomic signature can be determined, for P. putida KT2440 grown in the presence of known carcinogens, and the signature contains distinct spots that are not present in the healthy state gel then these spots should be further analyzed. When determining what spots should be picked for identification by mass spectrometry, the quantitative data relating to that individual spot needs to be considered. Mostly, the spot should show low variation across multiple gels when several variables are evaluated. If some of these protein spots that are unique to the carcinogenic proteomic signature show quantitative data consistent with the values presented in this study, then they should be considered biomarkers. These biomarkers of carcinogenicity will serve as the foundation of the proposed novel screening method.

In conclusion, the proposed screening method may eventually supplement or replace the Ames Test and eliminate many of the false results that are produced. Not only would such a screening method eliminate these false results, it has the potential to assign a danger rating to suspect carcinogens. The development of this screening method must be based on reproducible quantitative data in order to be welcomed and supported in the field of proteomics. By establishing the standards for intra-laboratory reproducibility, presented here, appropriate standards for quantitative data can be readily achieved by a laboratory.

References

1. A.M. Campbell, L.J. Heyer. *Genomics, Proteomics, & Bioinformatics;* Pearson Education, Inc. San Francisco, CA. 2003, p 162.
2. R. Westermeier, T. Naven. *Proteomics in Practice: A Laboratory Manual of Proteome Analysis*. 3rd ed. Wiley-VCH. Verlag-GmbH Weinheim, 2002, p. 2002.
3. 2D Electrophoresis for Proteomics: A methods and product manual. BioRad Laboratories, publication 2641, revision B.
4. D.E. Garrfin. Two-Dimensional gel electrophoresis: an overview. *Trends Anal. Chem. Vol 22 No.*, **2003**, pp. 263 – 272.
5. Isoelectric focusing animation. http://fig.cox.miami.edu/~cmallery/150/protein/proteinsb.htm (accessed 11/2004).
6. Product: L4390. www.sigmaaldrich.com (accessed 2/7/05).
7. http://www.bio-rad.com/LifeScience/pdf/Bulletin_2710.pdf (accessed 2/9/05).
8. Blomberg, A. , Blomberg, L. , Norbeck, J., Fey, S.J., Larsen, M., Roepstorff, P., Degand, H., Bountry, M., Posch, A., Gorg, A. Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. *Electrophoresis* **1995**, 16, pp. 1935–1945.
9. Lopez, M.F., Patton, W.F., Reproducibility of polypeptide spot positions in two-dimensional gels run using carrier ampholytes in the isoelectric focusing dimension. *Electrophoresis* **1997**, 18, pp. 338-343.

10. Voss, T., Haberl, P., Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels: Consequences for comprehensive data analysis. *Electrophoresis* **2000**, 21, pp. 3345-3350.

11. Choe, L. Kelvin, L Qunatitative and qualitative measure of intralaboratory two-dimensional protein gel reproducibility and the effects of sample preparation, sample load, and image analysis. *Electrophoresis*. **2003**, 24. pp. 3500 – 3507.

12. Challapalli, K., Zabel, C., Schuchhardt, J., Kaindl, A., Klose, J., Herzel, H., High reproducibility of large-gel two-dimensional electrophoresis. *Electrophoresis.* **2004**, 17. pp. 3040 – 3047.

13. Molloy, M., Brzezinski, E., McDowell, M., Hang, J., VanBogelen, R., Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics.* **2003**, 3, 1912 – 1919.

14. Klose, J., Kobalz, U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis. **1995**, 16(6), pp.1034-59.

15. Anderson, N.L., Anderson, N.G., Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis.* **1998**, 19, pp. 1853 – 1861.

16. Zahn, X., Desiderio, D. Differences in the spatial and quantitative reproducibility between two second-dimensional gel electrophoresis systems. *Electrophoresis.* **2003,** 24, pp. 1834 – 1846.

17. Dutt, M., Lee, K. The scaled volume as an image analysis variable for detecting changes in protein expression levels by silver stain. *Electrophoresis.* **2001,** 22, pp. 1627 – 1632.

18. Bacterial Growth Curve. http://web.deu.edu.tr/atiksu/toprak/curve.gif (accessed 9-6-05).

19. Phoretix 2D Manual. Nonlinear Dynamics, 2004 version. Tutorial 7 p. 97.

20. The streak plate method of isolation. http://www.austin.cc.tx.us/microbugz/18streakiso.html, (accessed 1/13/05.)

21. Reardon, KF. Kim, KH. Two-dimensional electrophoresis analysis of protein production during growth of *Pseudomonas putida* F1 on toluene, phenol, and their mixture. *Electrophoresis* (02) 23, 2233-2241.

22. 2D Electrophoresis for Proteomics: A methods and product manual. BioRad Laboratories, publication 2641, revision B.

23. ReadyStrip IPG Strip Instruction Manual. BioRad Laboratories, catalog # 163-2099, 4006166 revision G.

24. ReadyStrip IPG Strips and the PROTEAN IEF Cell. BioRad Laboratories. Bulletin 2426 revision D. http://www.biorad.com/LifeScience/pdf/Bulletin_2426.pdf (accessed 1/20/05).

25. Phoretix2D Manual. Nonlinear DynamicsLtd, Durham, NC. 2004 version.

26. Procedure for the original Bradford Protein Assay. http://www.science.smith.edu/departments/Biochem/Biochem_353/Bradford.html (accessed 9-28-05).

Glossary

Active Rehydration: a low electrical current is applied to the IPG strip during sample application. This electrical current helps to pull larger proteins into the strip.

Ames Test: the current non-animal test widely used to determine if a chemical is a mutagen.

Average Gel: an artificial gel that is created from a statistical combination of several normal gels.

Bacterial Colony: a group of cells that was created from multiple divisions of a single cell, and is considered genetically identical.

Biomarker: a biological response to a chemical that gives a measure of exposure and sometimes also of a toxic effect.

Carcinogenic: capable of causing cancer.

Coefficient of Variation (CV): the standard deviation of a sample expressed as a percentage of the mean. (CV = standard deviation/mean x 100)

Column chromatography: A type of chromatography in which the stationary phase is an insoluble material packed into a glass or metal column

Electrophoresis: separation of charged molecules in an electric field across a porous medium

Isoelectric Focusing: the first dimension of separation in two-dimensional gel electrophoresis. Proteins are separated based on their isoelectric point.

Isoelectric Point: pH conditions at which the net charge on a protein is zero.

Mid-log Phase: the middle of the log phase of a bacterial life cycle

Mutagen: a chemical that can cause genetic mutations

Noise: Any unwanted signal, such as in audio and video that adversely affects the quality of the sound or picture

Normalized Volume: the volume of a specific spot divided by the total volume of all the spots on the gel multiplied by 100

Pairwise Comparison: A comparison between the means of two groups.

Passive Rehydration: no electrical current is applied to the IPG strips during sample application and the proteins are allowed to absorb into the dehydrated gel on the strip.

Polyacrylamide Gel Electrophoresis (PAGE): gel electrophoresis using a polymerized polyacrylamide matrix to separate molecules on the basis of size, charge, or both; usually used to separate proteins or sequence nucleic acids.

Proteomics: the systematic analysis of the protein expression of healthy versus disease cells.

Proteomic Signature: the specific proteins expressed by an organism under a particular set of conditions.

Pseudomonas: The bacterial genus *Pseudomonas* includes plant pathogenic bacteria such as *P. syringae* (various pathovars), the opportunistic pathogen *P. aeruginosa*, the ubiquitous soil bacterium *P. putida*, and some species that are known to cause problems in dairy products. However, the actions of this group of bacteria are mainly considered neutral to humans. They are considered to be metabolically diverse agents of spoilage and degredation. In recent times, members of the Pseudomonas have been used as biocontrol agents.

***Pseudomona putida* KT2440:** the organism of interest due to its ability to degrade and metabolize aromatic rings.

Sibling Gel: replicate gels that are created from a single cell culture and processed simultaneously

Sodium Dodecyl Sulfate (SDS): a detergent that becomes an anion in solution and will bind to protein in proportion to the molecular weight.

Sonication: a process that uses high frequency sound waves to break apart cell membranes and releases internal cellular components into solution.

Succinic Acid: a basic carbon source for growth of *P.putida*, and the control carbon source for this study.

T-test: statistical test to determine whether the difference between two sample means is statistically significant

Warping: digitally overlaying the spots on one gel to the spots on a reference gel.