PROTEOMICS ANALYSIS OF THE METABOLIC TRANSITION BETWEEN AEROBIC AND ANAEROBIC CONDITIONS IN ESCHERICHIA COLI

by

Mohammed Yahya Refai

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

August 2019

©COPYRIGHT

by

Mohammed Yahya Refai

2019

All Rights Reserved

DEDICATION

I dedicate my dissertation work to my family and many friends. A special feeling of gratitude to my loving father Yahya Refai who supported my journey from the day I started back in 2013, and to my mother, Nadia Almahdi who passed away back in 2008 who was the main reason I pursued graduate studies. To my brothers, Amr, Alaa and Rafan who supported me emotionally, spiritually, and for the love and passion.

I also dedicate this dissertation to the person who sparked lots of ideas and was one day a true mentor who handed me the key to success, Professor Taha Kumosani. I also dedicate this work to Ameen Rezq, my master journey mentor, who supported and shaped my ideas into valid scientific questions.

Lastly, I dedicate this work to my partner and soulmate, Sarah Alsafi, who was the daily safety zone for me when I came home late, thank you, Sarah, for being an awesome wife. To my kids, Yahya and Abdullah, who rode with me through the entire journey.

ACKNOWLEDGEMENTS

I wish to thank my committee members who were generous with their expertise and time. A special thanks to my mentor, Brian Bothner, for the endless supply of optimism and encouragement that gave me much-needed motivation to push through hard times and failures. Brian's patient and constant support in helping me to peruse my dreams will not be forgotten, and I will rise to meet his example in my near future with students. I would like also to thank Valerie Copie for her pushing me to think outside the box and enlighten the way I investigate research questions. Martin Lawrence and Edward Dratz for their big-picture views of my research and thoughtful guidance in teaching me what it means to be a scientist. I would also like to thank Paul Greico and Colin Miller for providing me with their precious reagents for protein tagging which allowed me to conduct critical experiments.

I'm grateful to all the members of Bothner research group, especially Monika Lukaszewska for pushing my scientific knowledge beyond my expectations, and Patricia Mathabe for handing me keys to success. Moreover, to Dana Karmer, Nina Paris, Molly Lukes and Mary Trogden for the willingness to help in my research throughout my time in MSU. I also want to thank Jonathan Hilmer, Ganesh Balasubramanian and Jesse Thomas for their untiring effort to keep the mass spectrometry facility alive and well.

Finally, thanks to the Mass Spectrometry facility former and current members, for all of the mass spectrometry related knowledge I have acquired from you over the course of my tenure at MSU.

TABLE OF CONTENTS

1. INTRODUCTION AND LITERATURE REVIEW 1

Aerobic and Anaerobic Respiration in *E. col.* 2
Morphological Changes Associated with the Anaerobic Life Style 4
Redox Based Signaling in Prokaryotes and Eukaryotes 4
OxyR as a Redox-Sensitive Signaling Molecule 6
GAPDH Regulation of Cellular Function 8
Differential Alkylation to Capture Reversible Cys Oxidative States 11
Research Directions 12
Reference Cited 14

2. MASS SPECTROMETRY-BASED PROTEOMICS DATA ANALYSIS 23

Introduction 23
Mass Spectrometry-Based Proteomics 24
Proteomics-based LCMS Data Acquisition 28
The Basis of Protein Identification 29
Peptide Mass Fingerprint – intact peptide-based method 30
Tandem Mass Spectrometry 31
Peptide Spectra Matching (PSM) Algorithms 34
Mascot (MOWSE) 35
X!Tandem 36
Comet (SEQUEST) 36
Andromeda 38
Global Post-Translational Modification Discovery (G-PTM-D) 38
Label-free Protein Quantification 39
MS- and MS/MS-based Label-free protein quantification 40
The General Proteomic Data Analysis Pipeline 41
Proteomics Packages 41
Conclusion 45
Reference Cited 46

3. PROTEOMICS ANALYSIS OF THE METABOLIC TRANSITION BETWEEN AEROBIC AND ANAEROBIC CONDITIONS IN ESCHERICHIA COLI 52

Manuscript Information Page 52
Contribution of Authors and Co-Authors 53
Abstract 54

TABLE OF CONTENTS CONTINUED

Introduction....55
Materials & Methods....58
Bacterial and Growth Conditions....58
Protein Extraction and Differential Alkylation....59
Two Dimensional SDS-PAGE....60
LC-MS Instrumentation, Data Acquisition, And Data Analysis....60
Results....61
Global Proteome Profiling....64
Differential Alkylation Analysis....72
Discussion and Conclusions....76
Global Proteome Profiling....76
Oxidized Anaerobic Environment....80
Redox Analysis at the Level of Peptides and Potential Redox Signaling Molecule....80
Reference Cited....83

4. CONCLUDING REMARKS....101

REFERENCES CITED....104

APPENDICES....122
APPENDIX I....122
APPENDIX II....164

LIST OF TABLES

Table Page

1.1 *OxyR* transcranial factor in E. coli and regulated cellular functions based on the oxidative state 7

1.2 A Comparison of GAPDH in Different Organisms 9

2.1: Search speed and number of total peptide identifications at 1% Q-value for X!Tandem, OMSSA, and Comet on low-resolution and high-resolution MS2 spectra 37

3.1: Enrichment analysis for the associated KEGG, GOBP, GOMF and GOCC terms 66

3.2 The differential analysis comparison between the two experimental groups 67

3.1S The Identified Proteins in the Two Experimental Groups 124

3.2S The Cys-containing Peptides Identified in the experiment using MaxQuant platform along with sequence window 156

LIST OF FIGURES

Figure Page

Figure 1.1 The three metabolic modes of *E. coli* depending on the availability of electron donors and acceptors 3

Figure 1.2 Oxidative modification on cysteine 5

Figure 1.3 Oxidative thiol modification of GAPDH from cytoprotection to cytotoxicity 10

Figure 1.4 The Differential Alkylation workflow that targets three different Cys oxidative states, SH, SNO and S-SR 12

Figure 2.1: Schematic Diagram of the Quadrupole Time of Flight Mass Spectrometer Q-TOF- MS/MS 26

Figure 2.2: LC-MS data visualization methods 28

Figure 2.3: Schematic representation of the PMF matching process 30

Figure 2.4: The process of tandem mass spectrometry (MS/MS). 33

Figure 2.5 Timeline for the PSM tools 35

Figure 2.6 The general pipeline for proteomics-based mass spectrometry 44

Figure 3.1 Global proteome analysis of the Aerobic and Anaerobic *E. coli* proteins 63

Figure 3.3 Protein Differential alkylation analysis 73

Figure 3.4: Hierarchical cluster analysis for the differentially regulated Cysteine-containing peptide intensities (A) and with their corresponding total protein expression values (B). 75

Figure 3.1S: Growth curve of the E. coli strain MG1655 (K12). 122

LIST OF FIGURES CONTINUED

Figure Page

Figure 3.2S Hierarchical Clustering Analysis for the Differentially Regulated Proteins in the two Experimental groups, Aerobic (Ae) and Anaerobic (An). 123

ABSTRACT

As a facultative anaerobe, *Escherichia coli* has the ability to grow in anaerobic and aerobic environments. Despite detailed characterizations of this model organism in the presence and absence of oxygen, an in-depth understanding of changes to the proteome during transitions from aerobic to anaerobic growth is lacking. This thesis work focuses on elucidating how protein thiol oxidation and reduction change during a facultative anaerobe's transition from aerobic to anaerobic growth conditions, and pathways of thiol-mediated cell signaling. Redox driven changes in cysteine oxidation involved in signaling are referred to as "thiol switches". These modulate diverse biological activities ranging from gene expression and protein synthesis to environmental stress response. Surprisingly, little is known about the role of thiol switches during microbial transitions from aerobic and anaerobic growth conditions. To explore this uncharted territory, a mass-spectrometry (MS)-based proteomics workflow was developed and refined. Following extensive protocol optimization for high-throughput MS data processing, normalization, and pattern matching, the analytical pipeline was fine-tuned for the specific proteome-wide analysis of cysteine chemical modifications in *E. coli*. The approach was based on open-source software and publicly accessible databases, creating a transparent, reproducible, and easily sharable proteomics approach. Herein, the redox state and chemical forms of protein-based thiol switches in *E. coli* were characterized over time as the bacterium reversibly transitioned between aerobic and anaerobic growth conditions. Unexpectedly, differential alkylation analysis of cysteine-containing *E. coli* proteins revealed a higher degree of protein thiol oxidation under anaerobic growth conditions, a result not reported for *E. coli* or any other facultative anaerobe. Our proteome-wide analysis also revealed that cysteine redox potentials vary widely, and several specific *E. coli* proteins contain highly reactive thiols. These findings provide strong evidence for thiol-based signaling in *E. coli* in response to environmental changes such as aerobic to anaerobic growth transitions. Characterization of specific redox switches underlying metabolic changes associated with oxygen availability has uncovered a previously unknown *E. coli* cell signaling mechanism. Since transitioning between aerobic and anaerobic environments is associated with bacterial virulence, this work opens new avenues to target pathogenic facultative anaerobes and to develop novel thiol-based antibacterial therapies.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

Escherichia coli, one of the best-characterized prokaryotes, has served as a model organism for countless biochemical, biological, and biotechnological studies. Since the genome of *E. coli* was first completed in 1997[1], it has been a model for "Omics" approaches, including genomics [2–7], transcriptomics [8–11], proteomics [12–15] and metabolomics [16–20]. This is in part because of the relatively small genome which is relatively straight-forward for genetic mutations introduction. Additionally, *E. coli* can utilize a range of carbon sources, making it a good model for metabolic studies.

E. coli is a facultative anaerobe, which means it can grow in the presence and absence of oxygen. The ability to adapt to such changes is rooted in altered patterns of gene expression connected to oxygen availability [21]. The energetic state of the organism depends on the electron donors and acceptors that are available. In aerobic conditions, oxygen is the terminal electron acceptor in the generation of ATP by oxidative phosphorylation. In the absence of oxygen, formate, and nitrate are used as terminal electron acceptors to generate ATP. *E. coli* uses one of three metabolic modes to support growth needs [22]. In the presence of oxygen, aerobic respiration allows the complete oxidation of the carbon course through the glycolysis pathway, the TCA cycle, and the oxidative phosphorylation; therefore, more ATP is generated. Hence, this is the preferred mode [22]. In the absence of oxygen, metabolism must shift to other modes of energy production, which are less efficient. These fermentative pathways take advantage of a non-cyclic form of TCA cycle operation where glucose molecules are partially oxidized to form acetate, formate, and succinate [23,24].

Aerobic and Anaerobic Respiration in *E. coli*

During the transition from aerobic to anaerobic respiration, the role of the pyruvate dehydrogenase complex (*aceEF, Lpd*) which oxidizes pyruvate to acetyl-CoA and CO_2, is progressively taken over by Pyruvate Formate-Lyase (PFL) which converts pyruvate to formate and acetyl-CoA [25]. Moreover, acetate, succinate, ethanol, and lactate are produced by the mixed acid fermentative pathway [26,27]. The main difference between aerobic and anaerobic respiration in the *E. coli* resides in the enzymatic complexes involved in the oxidative phosphorylation process. An important molecule in this process is Nicotinamide Adenine Dinucleotide (NAD^+) which is the cellular source for the reducing equivalents for the biosynthetic reactions and reduction-oxidation reactions carrying electrons from one molecule to another. NAD^+ (the oxidized form) and NADH (the reduced form) play an essential role in determining the cellular reduction potential, along with the phosphate version, Nicotinamide Adenine Dinucleotide phosphate (NADP), the reduced (NADPH) and the oxidized form ($NADP^+$).

In the presence of oxygen, NADH dehydrogenase and Cytochrome b generate the Proton Motive Force (PMF) across the cytoplasmic membrane, according to the overall reaction, $2\ NADH + 2\ H^+ + O_2 \rightarrow 2\ NAD^+ + 2\ H_2O$, where oxygen is the terminal electron acceptor. In the absence of oxygen, where NADH is not efficiently produced, small molecules such as formate, glycerol-3-phosphate, and lactate donate electrons to nitrate (NO_3^-), nitrite (NO_2^-), trimethylamine oxide (TMAO), dimethyl sulfoxide (DMSO), or fumarate. The two enzymes involved in the transfer of an electron from formate to NO_3^- are formate dehydrogenase and nitrate reductase in an overall reaction as

follows formate + NO_3^- → CO_2 + NO_2^- +H_2O. In both states, PMF is generated to activate the ATPase complex. The ATPase complex then generates ATP to fulfill the cellular need. However, in the aerobic state, complete oxidation of glucose results in the generation of 36 ATP. This contrasts with anaerobic catabolism of glucose, which only produces 36 to 2 ATP is generated in the anaerobic state [28]. Therefore, the presence of oxygen results in a much more energy intensive metabolism in terms of the production of high energy molecules that the preferred metabolic state the organism thrives in (Figure 1.1).

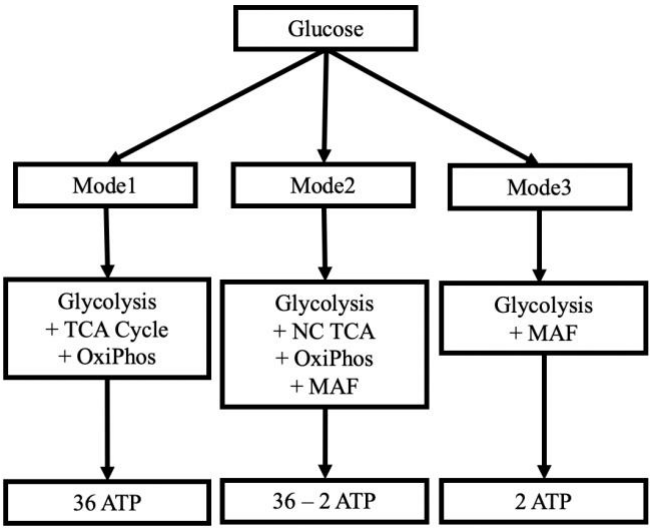


Figure 1.1 The Three Metabolic Modes of *E. coli* Depending on The Availability of Electron Donors and Acceptors. The first mode, the aerobic respiration, is the preferred one for *E. coli* since the glucose molecule is oxidizes completely to generate a net of 36 ATP in the presence of oxygen. The second mode which is referred to as anaerobic respiration. In this mode, incomplete oxidation of the glucose molecule results in less amount of ATP is generated (36 to 2 ATP per one glucose molecule). The third mode is fermentation. In this mode, the glycolysis and MAF pathways are the only functioning ones since no oxygen is present. TCA, Tricarboxylic acid, OxiPhos, Oxidative Phosphorylation, NC TCA, Non-cyclic Tricarboxylic acid cycle, MFA, Mixed Acid Fermentation.

Morphological Changes Associated with the Anaerobic Lifestyle

As many other facultative anaerobes, *E. coli* can grow aerobically and anaerobically. There are morphological changes associated with the growing condition. Murashko & Lin-Chao reported in 2017 changes associated with aerobically growing *E. coli* switched to an anaerobic environment [29]. They found that the anaerobic conditions promote the organism to switch from rod-shaped cells to form a filamentous form. It is noteworthy to mention that the *E. coli* (O157:H7) infectious mode is anaerobic. The infection is initiated in the anaerobic microenvironment in the gastrointestinal tract lumen in human [30]. Likewise, *Shigella flexneri* virulence and cell invasion secretion system proteins are controlled by the anaerobic regulon Fumarate and Nitrate Reductase (FNR) [31]. Therefore, the filamentous shape that is associated with the anaerobic lifestyle is thought to be linked to the virulence of the organism.

Redox Based Signaling in Prokaryotes and Eukaryotes

It has been well documented that ROS leads to cellular damage at the molecular level [32–37]. However, in the last decade, biologists have begun to appreciate that ROS are important mediates of cell signaling [38–43]. Functionally, they act by altering the activity of specific enzymes and small molecules [44]. Therefore, cellular redox potentials set the pace of many processes ranging from gene expression, protein translation, cell signaling, and apoptosis [45]. Proteins most often sense these redox potentials with redox-reactive cysteine residues, where the level of thiol oxidation controls protein activity [46–48]. High ROS levels trigger the rapid oxidation of transcription factors, including *oxyR* in bacteria [49].

Moreover, redox-regulated chaperones, such as Hsp33 and peroxiredoxin, are also affected by the high ROS levels [50,51]. More importantly, high oxidative levels alter the function of redox-sensitive enzymes, like GAPDH, to re-route metabolic fluxes essential for the energy production and redox homeostasis [46], as an example. The types of oxidative thiol modifications that affect redox-sensitive cysteines depend on the local protein environment of the thiol group and the nearby charged residues.

There are two general types of oxidative modifications that occur on cysteine: reversible and irreversible (Figure 1.2). The presence of hydrogen peroxide (H_2O_2)

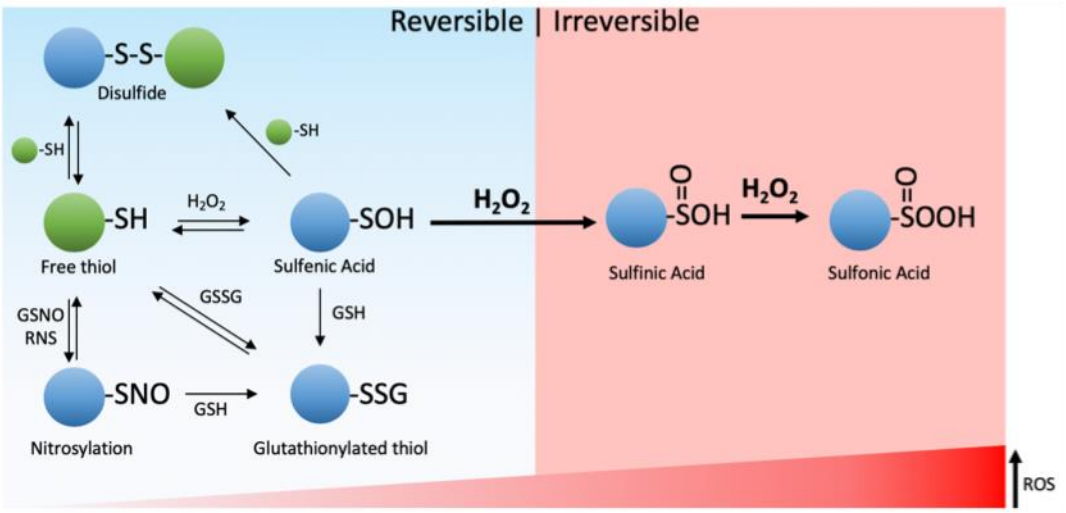


Figure 1.2 Oxidative Modification of Cysteine. Reversible modifications include Disulfide bond formation (-S-S-), Sulfonic acid (-SOH), Glutathionylated thiol (-SSG) and Nitrosylation (-SNO). These reversible modifications depend on the presence of reduced and oxidized glutathione (GSH and GSSG), nitrosoglutathione (GSNO) and hydrogen peroxide (H_2O_2). The irreversible oxidations are Sulfinic acid ($-SO_2H$) and Sulfonic acid ($-SO_3H$) are consequences of overoxidation by the H_2O_2. Adopted from (Chung, Wang, Venkatraman, Murray, & Van Eyk, 2013).

commonly causes Sulfenic (-SOH) acid formation [52,53], which is a revisable oxidative modification. However, the stable -SOH formation has been shown to occur in some redox-

sensitive enzymes like NADH peroxidase and GAPDH. Because of the reactive nature of -SOH, they often lead to further oxidation to form disulfide bonds (S-S) [54]. Furthermore, the presence of reactive nitrogen species (RNS) or nitroglutathione (GSNO) has been shown to lead to S-nitrosylation (-SNO) in a large number of redox-sensitive enzymes [53,55]. Moreover, -SNO regulates a variety of pathways such as G-protein-coupled receptor signaling [56–58], death receptor-mediated apoptosis [59–64], glutamate-dependent neurotransmission [65–68], vesicular trafficking [69–72], stimulation of prostaglandin synthesis [73–75], and the unfolded protein response [76]. Consequently, much effort is focused on understanding the role of -SNO in normal physiology and its contribution to pathophysiology. The second types of the cysteine modifications are the irreversible ones where the -SOH can be overoxidized irreversibly to Sulfinic acid ($-SO_2H$) and subsequently, Sulfonic acids ($-SO_3H$) (Figure 1.2). Moreover, because of the nature of those two irreversible oxidative modifications, they usually play only minor roles in controlling redox-regulated proteins and serve as biomarkers for oxidative protein damage [77].

OxyR as a Redox-Sensitive Signaling Molecule

OxyR is a transcriptional regulator for the expression of antioxidant genes in response to oxidative stress in bacteria [78]. OxyR has a thiol-switch that senses the oxidative stress in *E. coli*. It belongs to the *LysR* family of transcription factors that play a role in the peroxide detoxification pathway and also act as transcriptional activators [78]. According to the recently published review by Hilion and Antelmann (2015) [79], oxyR oxidation and formation of an intermolecular disulfide bond between the two conserved

cysteine residues, Cys199 and Cys208, represses a set of genes that target ROS to maintain a reduced cellular potential [80]. Moreover, -SOH at Cys-199, -SNO at Cys-199 and -SSG at Cys-199 also repress another set of genes including thioredoxin, glutaredoxin, ferrochatase, mannonate hydrolase and gluconate permease in *E. coli* [81]

Table1.1 *OxyR* transcranial factor in E. coli and regulated cellular functions based on the oxidative state. Adopted from Hillion & Antelmann, 2015 [79]

Redox sensor	Organism	Signal	Redox-sensing mechanism	Regulon genes	Regulon functions	Ref.
OxyR	E. coli	H_2O_2	C199*-C208	*ahpCF*	peroxiredoxin	80
				katG	catalase	84
				fur	Fe-uptake repressor	85
				dps	miniferrtin	
				mntH	Mn-uptake	
		H_2O_2	Cys199-SOH	*sufABCDE*	Fe-S cluster assembly	81
		NO	Cys199-SNO	*trxA*	thioredoxin	
		GSSG	Cys199-SSG	*gor*	glutaredoxin	
				dbsG	Cys-OH reduction	
				hemH	Ferrochelatase	
				oxyR	LysR-type regulator	
				fhuF	Fe3+ reductase	
				flu	Antigen 43 outer membrane protein	
				uxuAB	Mannoate hydrolase	
				gntP	Gluconate permease	

(Table1.1). In *Mycobacterium tuberculosis,* OxyS is the transcriptional regulator that senses the presence of peroxide after the oxidation of Cys25, which in turns signals to repress the expression of catalase [82,83]. OxyR and OxyS are considered excellent

examples of redox-sensitive molecules that are involved in redox signaling in response to the oxidative stress in prokaryotes.

GAPDH Regulation of Cellular Function

GAPDH is an enzyme present in all organisms and plays an essential role in glycolysis. GAPDH interconverts 1,3-Bisphosphoglycerate into Glyceraldehyde-3-phosphate and harbors a reactive Cys group in the active site which modulates enzyme activity in response to specific cellular conditions [86]. The reactive Cys residue is highly conserved among the three kingdoms, helping to validate its crucial functional role [86] (Table 1.2). The reactive cysteine in *E. coli* and *Saccharomyces cerevisiae* is Cys150. In higher organisms like *Homo sapiens* and *Arabidopsis thaliana* the catalytic cysteine is Cys152 and Cys156, respectively. It has been assumed that GAPDH operates continuously under physiological conditions; however, this has been brought into question due to the oxidation potential of the catalytic cysteine, which has strict consequences on the activity of this enzyme. It has been reported by Grant *et al*. (1999)[87] and Ralser *et al*., (2007)[88] that the oxidation of the catalytic cysteine in response to oxidative stress leads to enzyme inhibition and therefore rapid re-routing of the glycolytic flux into the pentose phosphate pathway (PPP) to generate more NADPH [87,88]. The redox response toward oxidative stress is expected as a global cellular theme in response to the oxidative stress by the regulation of antioxidant molecules.

In contrast, nitrosylation of the catalytic cysteine residue in GAPDH leads to a completely different outcome. The formation of the nitrosylated GAPDH (SNO-GAPDH) stabilizes Siah1, an E3-ubiquitin-protein ligase, through an unknown mechanism. Siah1

has a high cytosolic turnover rate under reducing conditions resulting in SNO-GAPDH-Siah1 complex binding specific regions of DNA, resulting in apoptosis in neurons and macrophages [86]. The overoxidation of GAPDH causes the enzyme to form inter-disulfide bridges between two GAPDH. This formation is translocated into the nucleus and insoluble aggregates are formed, which in turn leads to cell death [89]. Due to these multiple roles that GAPDH can play, it is characterized as an enzyme with so-called "Moonlighting" functions. These primary and secondary cellular functions that are regulated in response to changes in the environment classify this catalytic cysteine in GAPDH as a thiol switch that mitigates cellular stress (Figure 1.3).

Table 1.2 A Comparison of GAPDH in Different Organisms. The Accession Numbers are for the UniProt entries.

Organism	**Gene name**	**# of Cys**	**Catalytic Cys**	**Accession number (UniProt)**	**Protein MW**
Escherichia coli	gapA	3	150	E0W9Y9_ECOLX	35,532 Da
Saccharomyces cerevisiae	TDH1	2	150	G3P1_YEAST	35,750 Da
Homo sapiens	GAPD	6	152	G3P_HUMAN	44,501 Da
Arabidopsis thaliana	GAPC1	2	156	G3PC1_ARATH	36,914 Da

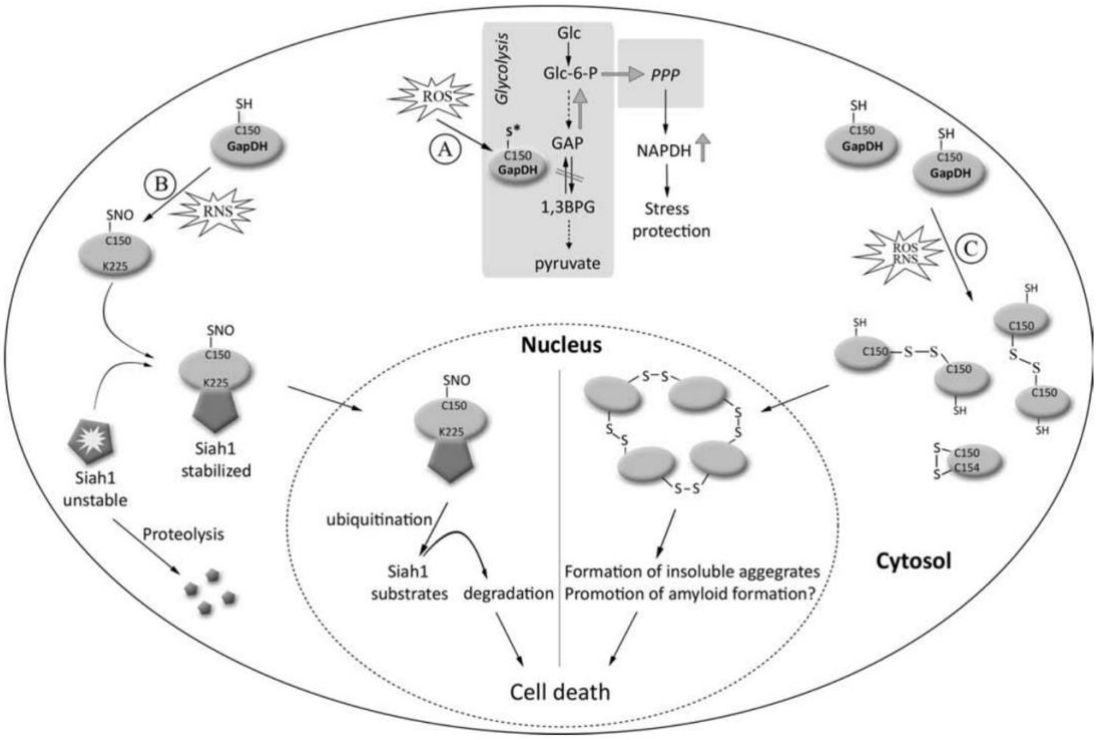


Figure 1.3 Oxidative Thiol Modification of GAPDH from Cytoprotection to Cytotoxicity. A, ROS-mediated inhibition of GAPDH re-routes metabolic flux as immediate response against oxidative damage. Oxidative stress-induced inhibition of GAPDH as a central component of glycolysis leads to the rapid re-routing G6P into PPP, one of the key routes for NADPH production. B, Exposure of GAPDH to RNS causes SNO of Cys150 in GAPDH. Cys150-SNO stimulates the interaction between the E3-ubiquitin ligase Siah1 and GAPDH. Siah1 escorts GAPDH into the nucleus where the GAPDH-Siah1 complex stabilize the E3-ligase. This stabilization increases the rate of degradation for a set of different target proteins and ultimately triggers cell death. C, Exposure of GAPDH to excess ROS leads to the initial disulfide bond formation between two active site cysteines Cys150. This leads to extensive conformational changes and exposes additional cysteines (e.g., Cys281) that undergo further intermolecular disulfide bond formation. This results in the accumulation of insoluble aggregates composed of multiple GAPDH subunits in the cytosol and the nucleus. The aggregates are either intrinsically cytotoxic or interact with other aggregation-prone proteins (e.g., Aß, tau, a-synuclein) to cause cytotoxicity [46]

Differential Alkylation to Capture Reversible Cys Oxidative States

The variety of chemical forms that cysteine can take on in a biological context and their important to catalysis and signaling has inspired the development of chemical means of detecting the different forms. In 2001, differential alkylation was introduced to analyze reversible modifications of cysteine using specific reductants and alkylating agents. Differential alkylation was first introduced in 2001 as a biotin switch for analysis of SNO and has been evolving since [90]. Also, it was also applied to other cysteine isoforms like SOH [91] and SSG [92] and all other chemical reducible cysteine modifications [93]. Regardless of the targeted cysteine modification, the principle remains the same. The initial step involves blocking reduced cysteine residues. The next step involves the addition of modification-specific reductant, followed by the addition of a second alkylation reagent. It is essential, however, to remove excess blocking agent before the addition of the modification-specific reductant. The last part is the addition of a non-specific reducing agent followed by the addition of a third alkylation reagent. The non-reducing agents used in this workflow are dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP). The combination of this workflow with tandem mass spectrometry allows for characterization of redox proteomes [94–96].

A simplified version of the differential alkylation workflow illustrates the elegance of the approach (Figure 1.4). Briefly, the workflow begins with a protein having three different Cys-modifications, SH, SNO, and S-SR. The first step is to introduce the first alkylation reagent N-ethylmaleimide (NEM) to block free thiols. The second step of

the workflow is the addition of an SNO-specific reducing agent, Ascorbate in the presence of Cu^{2+}. The next part is the addition of a second alkylation reagent, NEM-Biotin. The last two steps involve using a non-specific reducing agent like DTT and then eventually the addition of the last alkylation reagent, iodoacetamide (IAM). Lastly, the sample is subjected to the classical shotgun proteomics sample preparation protocol.

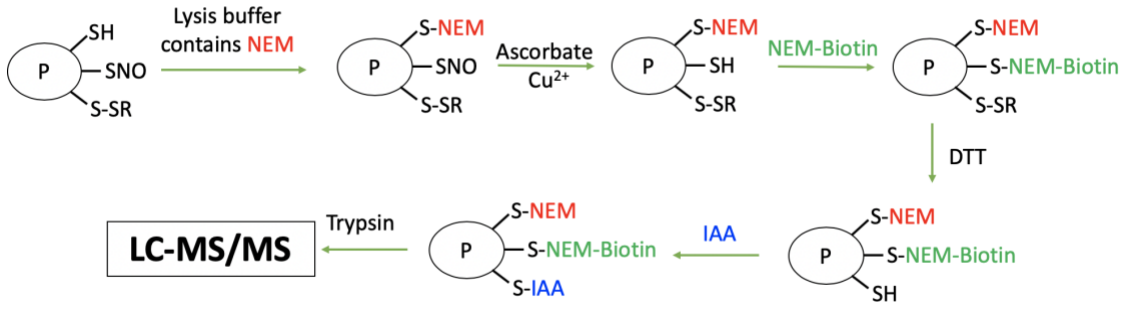


Figure 1.4 The Differential Alkylation workflow that targets three different Cys oxidative states, SH, SNO and S-SR. NEM is used first to block the free thiols SH first. SNO specific reductant (Ascorbate and Cu+2 ions) and non-specific reductant (DTT) are used sequentially to alkylate SNO and S-SR oxidative states using NEM-Biotin and IAM, respectively. Then, the sample is subjected to classical LC-MSMS shotgun proteomics workflow.

Research Directions

My research goal is to expand our understanding about the role of redox signaling in the two relevant conditions, aerobic and anaerobic growing conditions, in a facultative anaerobe. I aimed to investigate the global metabolic pathways regulated in both aerobic and anaerobic conditions and the associated Cys oxidative modifications. I used differential alkylation followed by LC-MSMS shotgun proteomics to achieve this goal.

Chapter two summarizes important avenues in mass spectrometry-based proteomics and describes the data analysis pipelines that I developed, tested, and applied. The chapter begins with the basis of protein identification at the level of MS and MSMS and explores the different open-source algorithms that enable rapid and accurate proteins identification. The chapter concludes with the packages that bundle the identification and quantification of shotgun proteomics in one platform.

Chapter three is the study we conducted to investigate the role of thiol switches in the transition from aerobic and anaerobic. In the study, *E. coli* (K12) was used as a model system. Shotgun proteomics- and thiolomics-based approaches were used in this study to investigate global differences between bacteria aerobic and anaerobic environments. We were able to identify over 1,000 proteins in the two studied groups. The functional annotation enrichment analysis shows higher rates of glucose utilization in the aerobic groups, whereas amino acid degradation and TCA cycle and glyoxylate associated enzymes are enriched in the anaerobic group. The shotgun based thiolomics data analysis along with the gel-based thiolomics data suggests a high degree of oxidation at the level of Cys-containing proteins is associated with anaerobic metabolism. The analysis at the level of Cys-containing peptides shows different patterns associated either to the nature of disulfide forming Cys residues or functionally active Cys-residues. We reported tnaA as potential redox switch that shows differential behavior at the level of tnaA Cys-containing peptides.

Reference Cited

1. Blattner FR. The Complete Genome Sequence of Escherichia coli K-12. *Science (80-)*. 1997;277(5331):1453-1462. doi:10.1126/science.277.5331.1453

2. Feist AM, Palsson BØ. The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol*. 2008;26(6):659-667. doi:10.1038/nbt1401

3. Barrick JE, Yu DS, Yoon SH, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*. 2009;461(7268):1243-1247. doi:10.1038/nature08480

4. Richmond C, Glasner JD, Mau R, Jin H, Blattner FR. Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Res*. 1999;27(19):3821-3835. doi:10.1093/nar/27.19.3821

5. Chang D-E, Smalley DJ, Tucker DL, et al. Noninvasive measurement of anatomic structure and intraluminal oxygenation in the gastrointestinal tract of living mice with spatial and spectral EPR imaging. *PNAS*. 2004;96(8):4586-4591. doi:10.1073/pnas.96.8.4586

6. Kolisnychenko V, Plunkett G, Herring CD, et al. Engineering a reduced Escherichia coli genome. *Genome Res*. 2002;12(4):640-647. doi:10.1101/gr.217202

7. Chen S, Lesnik EA, Hall TA, et al. A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. *Biosystems*. 2002;65(2-3):157-177. doi:10.1016/S0303-2647(02)00013-8

8. Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol*. 2015;23:133-140. doi:10.1016/J.MIB.2014.11.011

9. Hain T, Hossain H, Chatterjee SS, et al. Temporal transcriptomic analysis of the Listeria monocytogenes EGD-e σB regulon. *BMC Microbiol*. 2008;8(1):20. doi:10.1186/1471-2180-8-20

10. Chueca B, Pagán R, García-Gonzalo D. Transcriptomic analysis of Escherichia coli MG1655 cells exposed to pulsed electric fields. *Innov Food Sci Emerg Technol*. 2015;29:78-86. doi:10.1016/J.IFSET.2014.09.003

11. Nie L, Wu G, Culley DE, Scholten JCM, Zhang W. Integrative Analysis of Transcriptomic and Proteomic Data: Challenges, Solutions and Applications. *Crit Rev Biotechnol*. 2007;27(2):63-75. doi:10.1080/07388550701334212

12. Weber A, Kögl SA, Jung K. Time-dependent proteome alterations under osmotic

stress during aerobic and anaerobic growth in Escherichia coli. *J Bacteriol*. 2006;188(20):7165-7175. doi:10.1128/JB.00508-06

13. Vertommen D, Ruiz N, Leverrier P, Silhavy TJ, Collet J-F. Characterization of the role of the Escherichia coli periplasmic chaperone SurA using differential proteomics. *Proteomics*. 2009;9(9):2432-2443. doi:10.1002/pmic.200800794

14. Schmidt A, Kochanowski K, Vedelaar S, et al. The quantitative and condition-dependent Escherichia coli proteome. *Nat Biotechnol*. 2016;34(1):104-110. doi:10.1038/nbt.3418

15. Lai E-M, Nair U, Phadke ND, Maddock JR. Proteomic screening and identification of differentially distributed membrane proteins in Escherichia coli. *Mol Microbiol*. 2004;52(4):1029-1044. doi:10.1111/j.1365-2958.2004.04040.x

16. Rabinowitz JD, Kimball E. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Anal Chem*. 2007;79(16):6167-6173. doi:10.1021/ac070470c

17. Nobeli I, Ponstingl H, Krissinel EB, Thornton JM. A Structure-based Anatomy of the E. coli Metabolome. *J Mol Biol*. 2003;334(4):697-719. doi:10.1016/J.JMB.2003.10.008

18. Heinemann J, Noon B, Mohigmi MJ, Mazurie A, Dickensheets DL, Bothner B. Real-Time Digitization of Metabolomics Patterns from a Living System Using Mass Spectrometry. *J Am Soc Mass Spectrom*. 2014;25(10):1755-1762. doi:10.1007/s13361-014-0922-z

19. Yuan J, Doucette CD, Fowler WU, et al. Metabolomics-driven quantitative analysis of ammonia assimilation in E. coli. *Mol Syst Biol*. 2009;5(1):302. doi:10.1038/msb.2009.60

20. Tweeddale H, Notley-McRobb L, Ferenci T. Effect of Slow Growth on Metabolism of Escherichia coli, as Revealed by Global Metabolite Pool ("Metabolome") Analysis. *J Bacteriol*. 1998;180(19):5109-5116. https://jb.asm.org/content/180/19/5109.short. Accessed April 16, 2019.

21. Partridge JD, Scott C, Tang Y, Poole RK, Green J. Escherichia coli transcriptome dynamics during the transition from anaerobic to aerobic conditions. *J Biol Chem*. 2006;281(38):27806-27815. doi:10.1074/jbc.M603450200

22. Guest JR, Green, J. Irvine AS, Spiro S. *Regulation of Gene Expression in Escherichia Coli*. Boston, MA: Springer US; 1996. doi:10.1007/978-1-4684-8601-8

23. Gennis RB S V. *Respiration, Escherichia Coli and Salmonella Cellular and Molecular Biology*. 2nd ed. (Neidhardt FC, ed.). Washington, D.C: ASM Press;

1996.

24. Unden G, Bongaerts J. Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. *Biochim Biophys Acta*. 1997;1320(3):217-234. http://www.ncbi.nlm.nih.gov/pubmed/9230919. Accessed May 19, 2019.

25. Pecher A, Blaschkowski HP, Knappe K, Böck A. Expression of pyruvate formate-lyase of <i>Escherichia coli</i> from the cloned structural gene. *Arch Microbiol*. 1982;132(4):365-371. doi:10.1007/BF00413390

26. Dawes EA, Foster SM. The formation of ethanol in Escherichia coli. *Biochim Biophys Acta*. 1956;22(2):253-265. http://www.ncbi.nlm.nih.gov/pubmed/13382840. Accessed November 9, 2016.

27. Gupta S, Clark DP. Escherichia coli derivatives lacking both alcohol dehydrogenase and phosphotransacetylase grow anaerobically by lactate fermentation. *J Bacteriol*. 1989;171(7):3650-3655. http://www.ncbi.nlm.nih.gov/pubmed/2661531. Accessed November 9, 2016.

28. Stryer L. *Biochemistry*. 4th ed. (W. H. Freeman and Company., ed.). New York-Basingstoke; 1995.

29. Murashko ON, Lin-Chao S. Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. doi:10.1073/pnas.1703731114

30. Jones SA, Chowdhury FZ, Fabich AJ, et al. Respiration of Escherichia coli in the mouse intestine. *Infect Immun*. 2007;75(10):4891-4899. doi:10.1128/IAI.00484-07

31. Marteyn B, West NP, Browning DF, et al. Modulation of Shigella virulence in response to available oxygen in vivo. *Nature*. 2010;465(7296):355-358. doi:10.1038/nature08970

32. Davies K. Oxidative Stress, Antioxidant Defenses, and Damage Removal, Repair, and Replacement Systems. *IUBMB Life*. 2001;50(4):279-289. doi:10.1080/713803728

33. Davies KJ. Oxidative stress: the paradox of aerobic life. *Biochem Soc Symp*. 1995;61:1-31. doi:10.1042/BSS0610001

34. Riley PA. Free radicals in biology: oxidative stress and the effects of ionizing radiation. *Int J Radiat Biol*. 1994;65(1):27-33. http://www.ncbi.nlm.nih.gov/pubmed/7905906. Accessed May 8, 2019.

35. Sahnoun Z, Jamoussi K, Zeghal KM. [Free radicals and antioxidants: human physiology, pathology and therapeutic aspects]. *Therapie*. 1997;52(4):251-270.

http://www.ncbi.nlm.nih.gov/pubmed/9437876. Accessed May 8, 2019.

36. Sakac V, Sakac M. [Free oxygen radicals and kidney diseases--part I]. *Med Pregl*. 2000;53(9-10):463-74. http://www.ncbi.nlm.nih.gov/pubmed/11320727. Accessed May 8, 2019.

37. Devasagayam TPA, Tilak JC, Boloor KK, Sane KS, Ghaskadbi SS, Lele RD. Free radicals and antioxidants in human health: current status and future prospects. *J Assoc Physicians India*. 2004;52:794-804. http://www.ncbi.nlm.nih.gov/pubmed/15909857. Accessed May 8, 2019.

38. Finkel T. Signal transduction by reactive oxygen species. *J Cell Biol*. 2011;194(1):7-15. doi:10.1083/jcb.201102095

39. Turkan I. ROS and RNS: key signalling molecules in plants. *J Exp Bot*. 2018;69(14):3313-3315. doi:10.1093/jxb/ery198

40. Bigarella CL, Liang R, Ghaffari S. Stem cells and the impact of ROS signaling. *Development*. 2014;141(22):4206-4218. doi:10.1242/dev.107086

41. Zhang J, Wang X, Vikash V, et al. ROS and ROS-Mediated Cellular Signaling. *Oxid Med Cell Longev*. 2016;2016:1-18. doi:10.1155/2016/4350965

42. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol*. 2014;24(10):R453-62. doi:10.1016/j.cub.2014.03.034

43. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8(10):813-824. doi:10.1038/nrm2256

44. Knoefler D, Leichert LIO, Thamsen M, et al. About the dangers, costs and benefits of living an aerobic lifestyle. *Biochem Soc Trans*. 2014;42(4):917-921. doi:10.1042/BST20140108

45. Halliwell, B.; Gutteridge J, ed. Free radicals in biology and medicine. In: 4th ed. Oxford: Oxford University Press; 2007.

46. Brandes N, Schmitt S, Jakob U. Thiol-based redox switches in eukaryotic proteins. *Antioxid Redox Signal*. 2009;11(5):997-1014. doi:10.1089/ars.2008.2285

47. Klomsiri C, Karplus PA, Poole LB. Cysteine-based redox switches in enzymes. *Antioxid Redox Signal*. 2011;14(6):1065-1077. doi:10.1089/ars.2010.3376

48. Paulsen CE, Carroll KS. Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chem Biol*. 2010;5(1):47-62. doi:10.1021/cb900258z

49. Antelmann H, Helmann JD. Thiol-based redox switches and gene regulation. *Antioxid Redox Signal*. 2011;14(6):1049-1063. doi:10.1089/ars.2010.3400

50. Jang HH, Lee KO, Chi YH, et al. Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell*. 2004;117(5):625-635. doi:10.1016/j.cell.2004.05.002

51. Winter J, Ilbert M, Graf PCF, Ozcelik D, Jakob U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell*. 2008;135(4):691-701. doi:10.1016/j.cell.2008.09.024

52. Dalle-Donne I, Rossi R, Colombo G, Giustarini D, Milzani A. Protein S-glutathionylation: a regulatory device from bacteria to humans. *Trends Biochem Sci*. 2009;34(2):85-96. doi:10.1016/j.tibs.2008.11.002

53. Burgoyne JR, Eaton P. A rapid approach for the detection, quantification, and discovery of novel sulfenic acid or S-nitrosothiol modified proteins using a biotin-switch method. *Methods Enzymol*. 2010;473:281-303. doi:10.1016/S0076-6879(10)73015-9

54. Yeh JI, Claiborne A, Hol WG. Structure of the native cysteine-sulfenic acid redox center of enterococcal NADH peroxidase refined at 2.8 A resolution. *Biochemistry*. 1996;35(31):9951-9957. doi:10.1021/bi961037s

55. Foster MW, Hess DT, Stamler JS. Protein S-nitrosylation in health and disease: a current perspective. *Trends Mol Med*. 2009;15(9):391-404. doi:10.1016/j.molmed.2009.06.007

56. Kokkola T, Savinainen JR, Mönkkönen KS, Retamal MD, Laitinen JT. S-nitrosothiols modulate G protein-coupled receptor signaling in a reversible and highly receptor-specific manner. *BMC Cell Biol*. 2005;6(1):21. doi:10.1186/1471-2121-6-21

57. Nozik-Grayck E, Whalen EJ, Stamler JS, McMahon TJ, Chitano P, Piantadosi CA. S-nitrosoglutathione inhibits alpha1-adrenergic receptor-mediated vasoconstriction and ligand binding in pulmonary artery. *Am J Physiol Lung Cell Mol Physiol*. 2006;290(1):L136-43. doi:10.1152/ajplung.00230.2005

58. Whalen EJ, Foster MW, Matsumoto A, et al. Regulation of beta-adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2. *Cell*. 2007;129(3):511-522. doi:10.1016/j.cell.2007.02.046

59. Benhar M, Forrester MT, Hess DT, Stamler JS. Regulated protein denitrosylation by cytosolic and mitochondrial thioredoxins. *Science*. 2008;320(5879):1050-1054. doi:10.1126/science.1158265

60. Chanvorachote P, Nimmannit U, Wang L, et al. Nitric oxide negatively regulates

Fas CD95-induced apoptosis through inhibition of ubiquitin-proteasome-mediated degradation of FLICE inhibitory protein. *J Biol Chem*. 2005;280(51):42044-42050. doi:10.1074/jbc.M510080200

61. Hoffmann J, Haendeler J, Zeiher AM, Dimmeler S. TNFalpha and oxLDL reduce protein S-nitrosylation in endothelial cells. *J Biol Chem*. 2001;276(44):41383-41387. doi:10.1074/jbc.M107566200

62. Kim J-E, Tannenbaum SR. S-Nitrosation regulates the activation of endogenous procaspase-9 in HT-29 human colon carcinoma cells. *J Biol Chem*. 2004;279(11):9758-9764. doi:10.1074/jbc.M312722200

63. Kim YM, Talanian R V, Billiar TR. Nitric oxide inhibits apoptosis by preventing increases in caspase-3-like activity via two distinct mechanisms. *J Biol Chem*. 1997;272(49):31138-31148. http://www.ncbi.nlm.nih.gov/pubmed/9388267. Accessed March 30, 2016.

64. Mannick JB, Hausladen A, Liu L, et al. Fas-induced caspase denitrosylation. *Science*. 1999;284(5414):651-654. http://www.ncbi.nlm.nih.gov/pubmed/10213689. Accessed March 30, 2016.

65. Huang Y, Man H-Y, Sekine-Aizawa Y, et al. S-nitrosylation of N-ethylmaleimide sensitive factor mediates surface expression of AMPA receptors. *Neuron*. 2005;46(4):533-540. doi:10.1016/j.neuron.2005.03.028

66. Lipton SA, Choi YB, Pan ZH, et al. A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. *Nature*. 1993;364(6438):626-632. doi:10.1038/364626a0

67. Mustafa AK, Kumar M, Selvakumar B, et al. Nitric oxide S-nitrosylates serine racemase, mediating feedback inhibition of D-serine formation. *Proc Natl Acad Sci U S A*. 2007;104(8):2950-2955. doi:10.1073/pnas.0611620104

68. Takahashi H, Shin Y, Cho S-J, et al. Hypoxia enhances S-nitrosylation-mediated NMDA receptor inhibition via a thiol oxygen sensor motif. *Neuron*. 2007;53(1):53-64. doi:10.1016/j.neuron.2006.11.023

69. Lowenstein CJ. Nitric oxide regulation of protein trafficking in the cardiovascular system. *Cardiovasc Res*. 2007;75(2):240-246. doi:10.1016/j.cardiores.2007.03.024

70. Matsushita K, Morrell CN, Cambien B, et al. Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. *Cell*. 2003;115(2):139-150. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2846406&tool=pmcentrez&rendertype=abstract. Accessed March 30, 2016.

71. Morrell CN, Matsushita K, Chiles K, et al. Regulation of platelet granule exocytosis by S-nitrosylation. *Proc Natl Acad Sci U S A*. 2005;102(10):3782-3787.

doi:10.1073/pnas.0408310102

72. Wang G, Moniri NH, Ozawa K, Stamler JS, Daaka Y. Nitric oxide regulates endocytosis by S-nitrosylation of dynamin. *Proc Natl Acad Sci U S A*. 2006;103(5):1295-1300. doi:10.1073/pnas.0508354103

73. Kim SF, Huri DA, Snyder SH. Inducible nitric oxide synthase binds, S-nitrosylates, and activates cyclooxygenase-2. *Science*. 2005;310(5756):1966-1970. doi:10.1126/science.1119407

74. Xu L, Han C, Lim K, Wu T. Activation of cytosolic phospholipase A2alpha through nitric oxide-induced S-nitrosylation. Involvement of inducible nitric-oxide synthase and cyclooxygenase-2. *J Biol Chem*. 2008;283(6):3077-3087. doi:10.1074/jbc.M705709200

75. Tian J, Kim SF, Hester L, Snyder SH. S-nitrosylation/activation of COX-2 mediates NMDA neurotoxicity. *Proc Natl Acad Sci U S A*. 2008;105(30):10537-10540. doi:10.1073/pnas.0804852105

76. Uehara T, Nakamura T, Yao D, et al. S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature*. 2006;441(7092):513-517. doi:10.1038/nature04782

77. Dalle-Donne I, Scaloni A, Giustarini D, et al. Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. *Mass Spectrom Rev*. 2005;24(1):55-99. doi:10.1002/mas.20006

78. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol*. 1990;58(3):157-161. doi:10.1007/BF00548927

79. Hillion M, Antelmann H. Thiol-based redox switches in prokaryotes. *Biol Chem*. 2015;396(5):415-444. doi:10.1515/hsz-2015-0102

80. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie Van Leeuwenhoek*. 1990;58(3):157-161. http://www.ncbi.nlm.nih.gov/pubmed/2256675. Accessed May 20, 2019.

81. Kim SO, Merchant K, Nudelman R, et al. OxyR: a molecular code for redox-related signaling. *Cell*. 2002;109(3):383-396. http://www.ncbi.nlm.nih.gov/pubmed/12015987. Accessed May 20, 2019.

82. Li Y, He Z-G. The Mycobacterial LysR-Type Regulator OxyS Responds to Oxidative Stress and Negatively Regulates Expression of the Catalase-Peroxidase Gene. Cornelis P, ed. *PLoS One*. 2012;7(1):e30186. doi:10.1371/journal.pone.0030186

83. Domenech P, Honoré N, Heym B, Cole ST. Role of OxyS of Mycobacterium

tuberculosis in oxidative stress: overexpression confers increased sensitivity to organic hydroperoxides. *Microbes Infect*. 2001;3(9):713-721. http://www.ncbi.nlm.nih.gov/pubmed/11489419. Accessed April 28, 2019.

84. Zheng C, Xu J, Li J, et al. Two Spx Regulators Modulate Stress Tolerance and Virulence in Streptococcus suis Serotype 2. Lemos JA, ed. *PLoS One*. 2014;9(9):e108197. doi:10.1371/journal.pone.0108197

85. Choi H, Kim S, Mukhopadhyay P, et al. Structural basis of the redox switch in the OxyR transcription factor. *Cell*. 2001;105(1):103-113. http://www.ncbi.nlm.nih.gov/pubmed/11301006. Accessed May 20, 2019.

86. Hildebrandt T, Knuesting J, Berndt C, Morgan B, Scheibe R. Cytosolic thiol switches regulating basic cellular functions: GAPDH as an information hub? *Biol Chem*. 2015;396(5):523-537. doi:10.1515/hsz-2014-0295

87. Grant CM, Quinn KA, Dawes IW. Differential protein S-thiolation of glyceraldehyde-3-phosphate dehydrogenase isoenzymes influences sensitivity to oxidative stress. *Mol Cell Biol*. 1999;19(4):2650-2656. http://www.ncbi.nlm.nih.gov/pubmed/10082531. Accessed April 28, 2019.

88. Ralser M, Wamelink MM, Kowald A, et al. Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. *J Biol*. 2007;6(4):10. doi:10.1186/jbiol61

89. Sampathkumar Krishnan ‡, Eva Y. Chi §, Stephen J. Wood ⦀, et al. Oxidative Dimer Formation Is the Critical Rate-Limiting Step for Parkinson's Disease α-Synuclein Fibrillogenesis†. 2002. doi:10.1021/BI026528T

90. Wojdyla K, Rogowska-Wrzesinska A. Differential alkylation-based redox proteomics - Lessons learnt. *Redox Biol*. 2015;6:240-252. doi:10.1016/j.redox.2015.08.005

91. Jaffrey SR, Snyder SH. The biotin switch method for the detection of S-nitrosylated proteins. *Sci STKE*. 2001;2001(86):pl1. doi:10.1126/stke.2001.86.pl1

92. Shelton MD, Chock PB, Mieyal JJ. Glutaredoxin: Role in Reversible Protein *S* -Glutathionylation and Regulation of Redox Signal Transduction and Protein Translocation. *Antioxid Redox Signal*. 2005;7(3-4):348-366. doi:10.1089/ars.2005.7.348

93. García-Santamarina S, Boronat S, Espadas G, Ayté J, Molina H, Hidalgo E. The oxidized thiol proteome in fission yeast—Optimization of an ICAT-based method to identify H2O2-oxidized proteins. *J Proteomics*. 2011;74(11):2476-2486. doi:10.1016/J.JPROT.2011.05.030

94. Wojdyla K, Williamson J, Roepstorff P, Rogowska-Wrzesinska A. The SNO/SOH

TMT strategy for combinatorial analysis of reversible cysteine oxidations. *J Proteomics*. 2015;113:415-434. doi:10.1016/J.JPROT.2014.10.015

95. Qu Z, Meng F, Bomgarden RD, et al. Proteomic Quantification and Site-Mapping of *S* -Nitrosylated Proteins Using Isobaric iodoTMT Reagents. *J Proteome Res*. 2014;13(7):3200-3211. doi:10.1021/pr401179v

96. Fares A, Nespoulous C, Rossignol M, Peltier J-B. Simultaneous Identification and Quantification of Nitrosylation Sites by Combination of Biotin Switch and ICAT Labeling. In: Humana Press, Totowa, NJ; 2014:609-620. doi:10.1007/978-1-62703-631-3_41

CHAPTER TWO

MASS SPECTROMETRY-BASED PROTEOMICS DATA ANALYSIS

Introduction

Proteomics is the study of protein interactions, localization, abundance, and post-translational modifications (PTMs) on a system-wide level. The term proteomics was introduced in 1997 [97]. The proteome is the entire set of proteins that are produced by a living system and can refer to a cellular compartment, cell, tissue, or even an organism as is the case for the Human Proteome. The study of proteins within the context of the proteome includes, but is not limited to, the change of protein abundance in response to changes in the environment, the interaction of proteins with one another, protein PTMs, the involvement of a protein in a specific metabolic pathway, and protein localization. Proteomics is now an integral part of systems biology research providing insights into gene expression, protein regulation, and protein function. Importantly, because proteins are functional and catalytic biomolecules, proteomics is a step closer to the action compared to transcriptomics, yet does not suffer from the molecular identification problem that is prevalent in metabolomics (the study of small molecules). By integrating "Omics" studies such as Genomics, Transcriptomics, Proteomics, and Metabolomics systems, biology is becoming more comprehensive and powerful [98–104]. In this chapter, I will briefly cover a standard proteomics workflow and then overview different strategies employed in mass spectrometry-based proteomics data analysis and highlight different algorithms and the basis behind them.

Mass Spectrometry-Based Proteomics

Over the past 20 years, proteomics has been a driving force for the development of mass spectrometry (MS) technology that can facilitate the rapid and accurate identification of the thousands of proteins in complex biological samples. While the instrumentation often gets the bulk of the attention, there are many steps involved in a proteomics workflow. The overall process involves sample preparation, liquid chromatography for sample separation, and the mass measurement of ionized peptides. The first step, sample preparation, is critical to the generation of high-quality data and is an important part of the workflow. It is worth mentioning that there is no universal method for biological sample preparation. Therefore, developing a method that takes into account the nature of the biological sample and the protein fraction of interest. In this step, a researcher needs to address the goals of the experiment such as: the chemical nature of targeted proteins (soluble or insoluble fraction), the buffer needed for protein extraction, and the cellular localization of the targeted proteins (cytosolic, nuclear or extracellular). Following extraction, the collected proteins are subjected to proteolysis to generate peptides. Trypsin is the preferred enzyme because of the high specificity and ease of use. Second, liquid chromatography is used to reduce the complexity of the collected peptides by spreading out the mixture before MS analysis. Reverse-phase (RP) or Hydrophilic Interaction (HILIC) chromatography can be used to separate the peptides in order to perform the mass spectrometry measurement of the separated peptides sequentially. Lastly, mass spectrometry allows the accurate and rapid identification of thousands of peptides in one sample. There are several different mass analyzers and

instrument configurations that are used in mass spectrometry-based proteomics. These include the Triple Quadrupole mass analyzer (QQQ), Orbitrap mass analyzer, and Quadrupole Time of Flight mass analyzer (Q-TOF). The proteomics work that is described here was performed using Q-TOF style instruments.

The Q-TOF mass spectrometer consists of three main parts; source, analyzer, and detector (Figure 2.1). In order for the process to work, ionized molecules must be generated which are then directed into the instrument and subsequently measured. Ions are produced in the source, which can be either Electrospray Ionization (ESI) or Matrix-Assisted Laser Desorption/Ionization (MALDI). The most commonly used source for proteomics is ESI. When coupled with the LC system, the ESI source allows the liquid to be directly infused into high vacuum instruments and using high voltage and sheath gas, produce a spray of very small droplets known as Taylor cone [105]. To facilitate the ionization process, the source temperature is manipulated to rapidly evaporate solvent droplets containing the analytes. In proteomics, the analytes are peptides. The peptides become charged and travel through the instrument under high vacuum system to reach the second part of the instrument, the analyzer (analytical quadrupole).

The function of the quadrupole is to allow only ions with the selected m/z ratio to reach the detector while diverting other ions [106]. This is achieved by modulating the quadrupole Radio Frequency (RF) and Direct Current (DC) potential [106]. At specific RF and DC, only ions of specific m/z can pass through the quadrupole to reach the second part of the analyzer, TOF. The TOF is a tube under a high vacuum consisting of a pulser, reflectron and detector [107]. Basically, the ions that passed through the quadrupole pulsed through the tube depending on their kinetic energy [108]. Every ion travels through the tube

based on its m/z, where smaller ones fly faster than bigger ones. The reflectron uses a constant electrostatic field to reflect the features toward the detector, which in turn generates signals that are converted to quantitative values in the acquisition software.

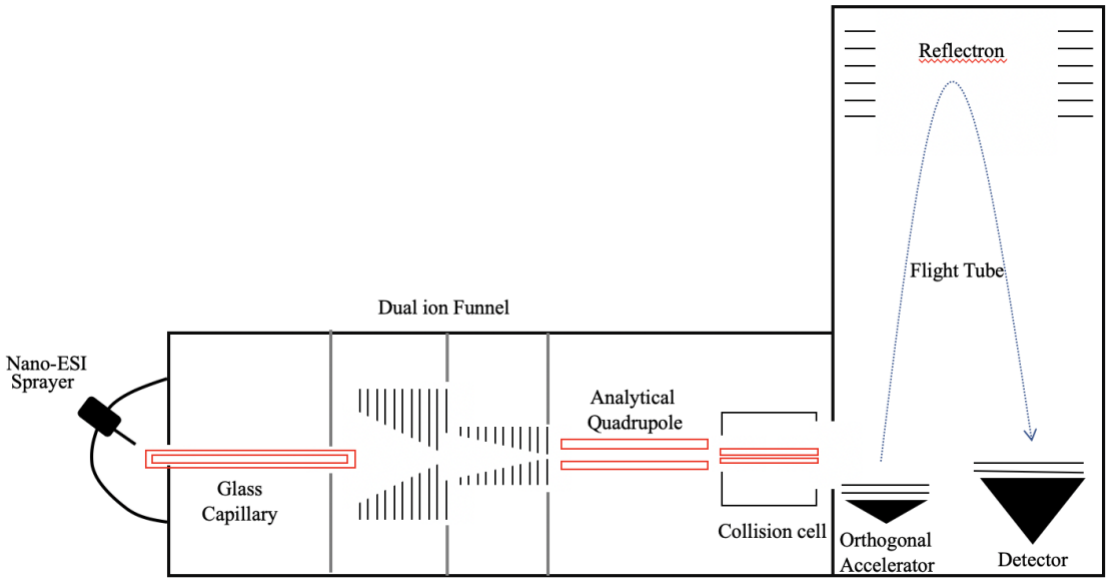


Figure 2.1: Schematic Diagram of the Quadrupole Time of Flight Mass Spectrometer Q-TOF- MS/MS. Peptides enter the system through the Nano-ESI source and are subjected to high voltage. The resulted ionized peptides fly inside the glass capillary and then filtered through the dual ion funnel. The analytical quadrupole allows only a selected range of m/z ions to pass MS. According to the acquisition method, which determined the peptide required for fragmentation, the analytical quadrupole isolate that ion and eventually subjected to fragmentation in the collision cell MS/MS. The Orthogonal Accelerator pushes the intact ions MS and fragment ions MS/MS to fly into the TOF. The smaller the m/z the shorter time it takes to reach the detector and vice versa for the large m/z. Finally, ions reach the detector where they produce a signal which is recorded and displayed by the acquisition software.

The visualization of mass spectrometry raw data is one of the quality control steps that allows quick and relative comparisons between biological replicates and

experimental samples. There are two ways to investigate mass spectrometry raw data. The first method involves using vendor-specific tools such as DataAnalysis® (Bruker) and MassHunter® (Agilent). The data is plotted according to their elution profile as x-axis and the intensity of the eluted analytes as the y-axis (Figure 2.2A). A second method is to plot the elution profile as x-axis against the m/z as y-axis for every feature in the sample. The second method involves plotting a third dimension in which the intensity of every feature is displayed as gray to black scale (Figure 2.2B). This method allows the visualization of every m/z intensity regardless of the elution profile and also can show more details about the quality of the sample preparation in terms of the presence of contaminants and polymers [109].

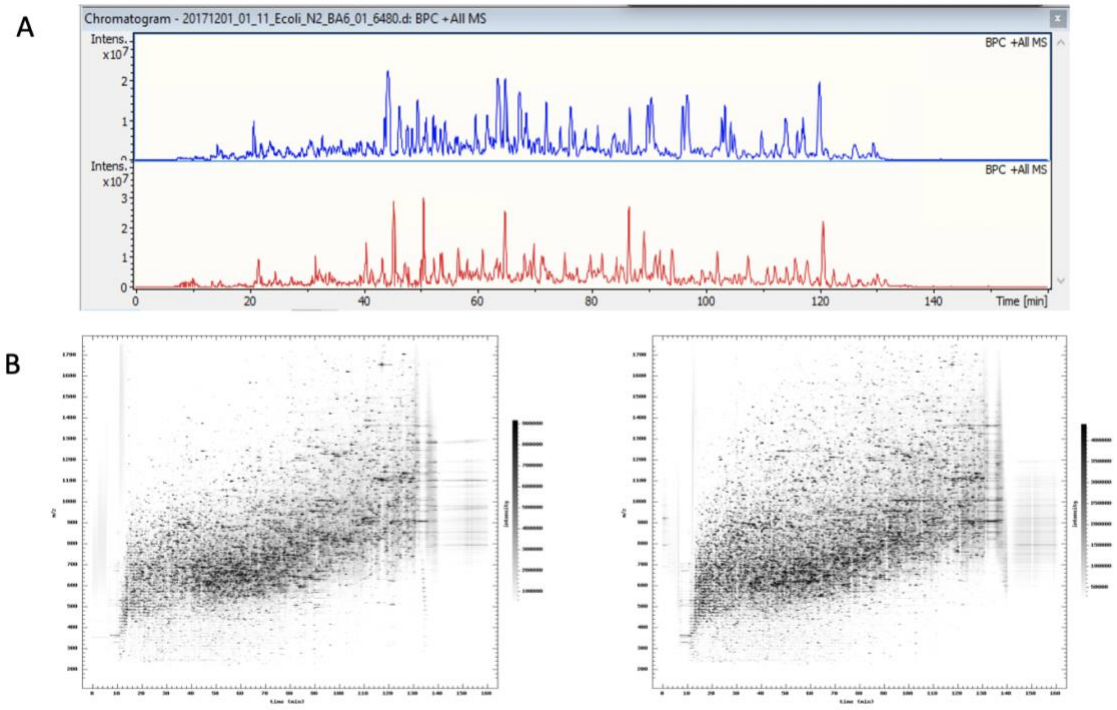


Figure 2.2: **LC-MS data visualization methods**. A, Basic Peak Chromatogram (BPC) for aerobic (bottom) and anaerobic (top) shotgun runs on the LC system. Reverse Phase (RP) chromatography is preformed to separate the digested peptides over a course of 5-30% MeCN. This method shows peptides elution profile over the course of the LC gradient run time. B, is the peptide elution map over three axis dimensions, retention time (bottom), m/z ratio (left) and the peptide intensity (the gradient scale on the right).

Proteomics-based LCMS Data Acquisition

In mass spectrometry-based proteomics, the biomolecules of interest are peptides. Most likely, tryptic peptides. Trypsin is the most used protease in the field due to the high specificity and reproducibility, although other proteases may be used[110]. Tryptic peptides are considered good candidates for ESI-LC-MS analysis. This is because trypsin cleaves at arginine and lysine residues, assuring that a basic amino acid is present in each peptide.

Proteomics is exclusively conducted in positive mode, meaning that positive ions are measured. Because of their basicity, tryptic peptides tend to be multiply protonated when analyzed by electrospray ionization (ESI) forming +2, +3 or +4 charged species[111]. Multiply charged peptides generate better fragmentation data than singly charged species; therefore, trypsin has become the gold standard for protein digestion in proteomics-based shotgun proteomics [112]. The mass analyzer, then, selects these peptides and subject them to tandem analysis in which every peptide is subjected to inert gas to dissociate into fragments. In proteomics, collision-induced dissociation (CID) and higher-energy collision dissociation (HCD) are used [113]. The acquisition software collects peptides and their fragments for later on analysis. The number of peptides and their fragments acquired in every sample could reach to >100,000 peptides.

The recent advances in instrumentation and high throughput mass spectrometry have increased the demand for computational power for the analysis of raw data. A single shotgun proteomics LCMS experiment can produce a gigabyte of data. Depending on the number of samples in a project, data conversion and analysis can take hours to days. Hence, specialized tools and data analysis pipelines are continually being developed and improved to efficiently and adequately analyze the raw data.

The Basis of Protein Identification

The first step in the process of proteomics data analysis is protein identification, where computational power is needed. Proteomics-based mass spectrometry involves two strategies in this step. The first strategy is peptide mass fingerprinting (PMF) [114–118]

and the second strategy is sequencing information derived from tandem mass spectrometry (MS/MS) [119].

Peptide Mass Fingerprint – intact peptide-based method

Peptide Mass Fingerprint (PMF) consists of matching the theoretical masses generated from *de novo* protein digestion to experimental data collected in the form of a list of peptide masses. Essentially, there are two types of information needed to confidently identify a protein: theoretical *de-novo* for the predicted peptide masses and the experimentally collected feature masses. The theoretical *de-novo* masses for all peptides are compared to the experimentally collected features at the level of intact peptide mass (MS) and then converted into a protein sequence map (Figure 2.3).

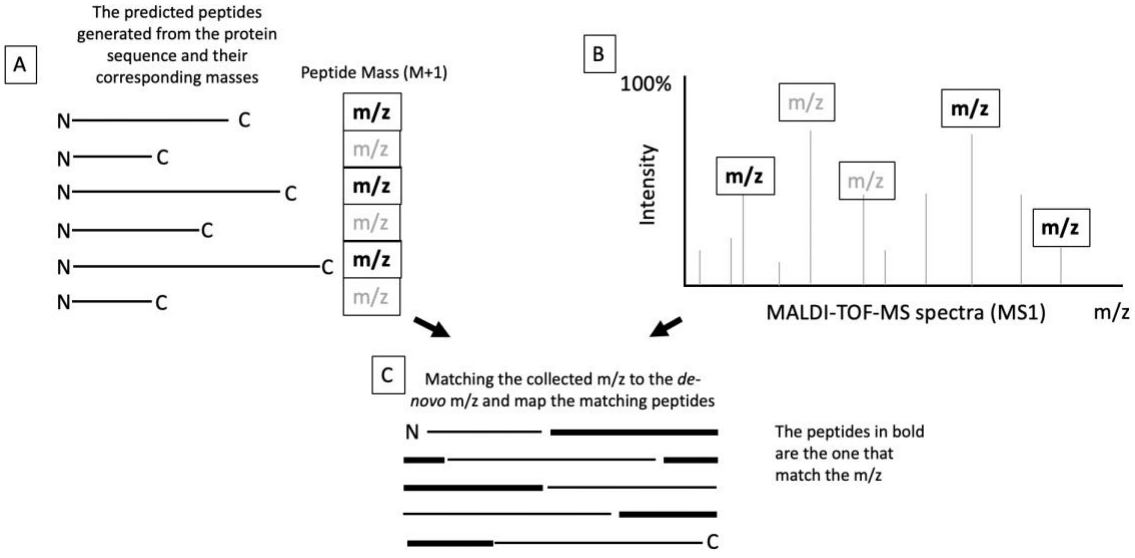


Figure 2.3: Schematic representation of the PMF matching process. A, the list of *de-novo* generated peptides and the predicted mass to charge ratios m/z that are generated theoretically. B, The MALDI-TOF-MS spectrum. C, The matching process at the level of MS and the generation of a protein sequence map. The peptides in bold are the ones that match the theoretical m/z to the experimental data.

The approach is most often used with Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry-Time of Flight data (MALDI-TOF-MS) because these instruments do not typically have MS/MS capability. The advantage of this approach is that it is fast and straightforward to interpret. However, there are limitations to this approach. Protein identification is made by matching the observed mass features to a list of expected peptides based on an *in silico* protease digestion of the proteins believed to be in the sample. This could be an entire proteome or a single sequence. Because peptides with different sequences can have the same mass, this approach is not typically used for complex mixture, but rather for samples that have been partially purified. Often not every mass feature is matched to the *in-silico* peptides masses. This could be because the experimental data may include chemically modified peptides, as

well as peptides that are generated from contaminant proteins such as keratins. Because only the m/z of intact peptides are used to match with the expected peptides, false positives are a serious concern. This must be considered by search algorithms and leads to lower confidence in identifications.

Tandem Mass Spectrometry

The second strategy for protein identification by mass spectrometry uses tandem mass spectrometry (MS/MS) (Figure 2.4). In this strategy, the intact mass and the masses of fragments generated by collision-induced disassociation (CID) are used to identify peptides. Because of the repeating nature of the peptide backbone and its inherent chemical property, CID induces predictable fragmentation patterns. This allows a predicted list of fragment masses to be generated for any given peptide sequence. Using intact and fragment masses, peptide and protein identifications can be made with high statistical confidence.

Complex samples such as the peptides generated from a whole-cell lysate are best analyzed by this technique. Due to many peptides generated by trypsin in such a sample (> 10,000 for a typical bacterial cell), chromatography to fractionate the sample is critical. High-Pressure Liquid Chromatography (HPLC) or now more commonly Ultra High-Pressure Liquid Chromatography (UPLC) is used before MS analysis. This step is essential to reduce the complexity entering the instrument at any given one time. Usually, RP chromatography is conducted in the separation step where more polar peptides elute early and the less polar peptides elute late when a gradient of increasing organic solvent is used.

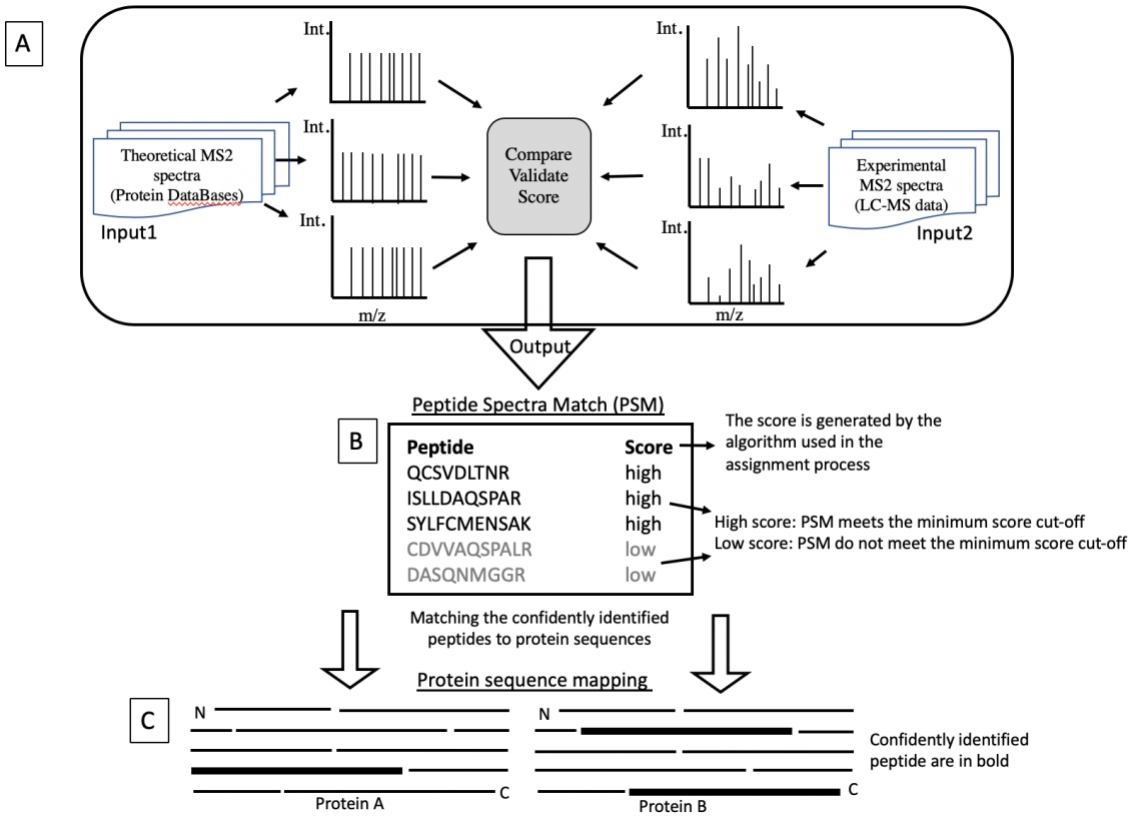


Figure 2.4: The process of tandem mass spectrometry (MS/MS). A, the two input types of information are on the top. 1, the theoretical *de-novo* information for all peptides generated by the database search engine. 2, the experimental mass spectrometry raw data. Both are then processed by the database search engine. The database search engine compares, validates and scores the inputs and generates outputs. B, the output contains a list of PSMs along with their scores. The scores represent the quality of the PSM, where the score that meets the cut-off is considered a significant PSM and the scores that do not meet the minimum cut-off is considered an insignificant PSM. C, the list of PSM is mapped by the database search engine to the matching protein sequences.

In a typical experiment, thousands of peptides are analyzed along with their MS/MS fragments. Therefore, performing *de-novo* sequencing on each peptide is not a

feasible option. As a result, numerous computational approaches and tools have been developed for the assignment of peptide sequence using MS, MS/MS, and gene sequence information [120]. These approaches involve generating *in silico* spectra using gene sequences, which are translated into protein sequences. The expected and observed data is compared at the level of MS and MS/MS to make identifications. Every peptide is unique in terms of the amino acid sequence; hence, a confident identity of a peptide can be determined, and a specific level of confidence in the identity can be predicted [121]. The output information mainly consists of a list of Peptide Spectrum Matches (PSMs) and a list of the identified proteins and quantification values. The validation process depends on the more than one aspect; one of which is the presence of the MS/MS fragment in the theoretical reference spectrum and the second is the intensity of these fragment in respect to the background noise [121].

Peptide Spectra Matching (PSM) Algorithms

In the last 20 years, many open-source tools have become available for researchers to use. Several research groups around the world are focused on developing and continually improving software packages to overcome the complexity of the raw mass spectrometry data and improve the confidence of protein identification. Figure 2.5 shows a timeline for the development of the algorithms described in the chapter. This timeline only shows the key important PSM algorithms that significantly impacted the way protein is being identified today. Although we are going to cover only open-source PSMs algorithms, there are many commercially available tools for researchers to use.

Before we introduce the PSM algorithms, we need to acknowledge the added benefits for the implementation of these algorithms to the mass spectrometry proteomics data analysis pipeline. Often the most important factor in method development is algorithm efficiency because this can reduce the computational time required for analysis. Other desirable features include the capability to analyze multiple samples that contain thousands of peptides as a single sample to enhance the depth of coverage.

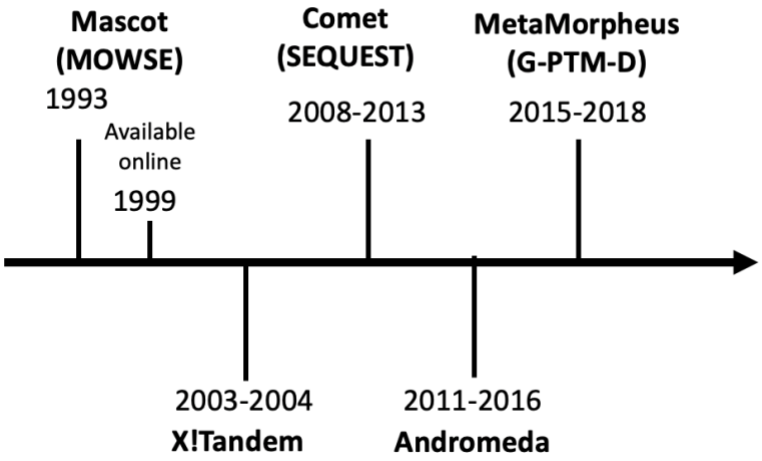


Figure 2.5 Timeline for the PSM tools. MOWSE was first reported back in 1993 and became available online by the end of 1999. After that, X!Tandem was reported in 2003. Then, a new version of SEQUEST, Comet, was reported in 2013. Andromeda was reported in 2011 and then updated in 2016. The last addition to the PSM tools is G-PTM-D which was introduced in the 2015 and updated in 2018.

Mascot (MOWSE) The first database search algorithm for protein identification using mass spectrometry was MOlecular Weight SEarch (MOWSE). It was published 26 years ago by Pappin and Bleasby, in 1993 from Darsebury Laboratory, Cheshire – UK.

Although several research groups developed independently their own peptide mass fingerprint tools, MOWSE was the first to take account on the non-uniform distribution of peptide sizes, which results from digestion by an enzyme [122]. In 1999, MOWSE became available for the public use under Mascot Science online server (www.matrixscience.com).

X!Tandem X!Tandem was published by Craig and Beavis in 2003 and 2004. While this was one of the first algorithms to be introduced, it is still used extensively. X!Tandem can reduce the search time required by several orders of magnitude compared to the classical one-step methods for matching proteins to mass spectra. Akhtar *et al.*, reported a search time comparison for Crux and OMSSA in comparison to X!Tandem for the same data set containing 7,850 peptides. X!Tandem came out the fastest (23 seconds). Crux came out slower by 3.8-fold in comparison to X!Tandem and OMSSA came out the slowest by 5.3-fold in comparison to X!Tandem [123].

Efficiency is achieved by dividing the process into multiple steps; the first step identifies all candidate protein sequences that match the set of MS/MS data. The protein candidates are then used to build a model which is defined by detailed analysis of the set of MS/MS for evidence of incomplete enzymatic hydrolysis, non-specific hydrolysis and chemical modifications of the amino acid resulting either from sample handling (e.g. deamidation of asparagine and glutamine, oxidation of methionine) or PTMs (e.g. N-terminal acetylation) [124,125].

Comet (SEQUEST) Another tool that has gained much attention from the proteomics community is Comet, published by Eng et al., 2013. Comet is a direct

descendant of the University of Washington's academic version of SEQUEST [126]. Comet uses the fast cross-correlation algorithm that was first described by Eng, *et al.*, 2008, with a new implementation that the spectral pre-processing in the fast-correlation algorithm eliminates the need for creating and sorting theoretical spectra. This allows the cross-correlation to be calculated by a simple sum of pre-processed peak intensities at each calculated fragment ion mass [127]. Table 1 displays the differences of Comet, OMSSA and X!Tandem in run time and total identified number of PSMs for the same sample that was run on two different instruments, low-resolution data (Ion-trap MS/MS) and high-resolution data (Orbitrap MS/MS), respectively. The LC-MS/MS files of human embryonic stem cells samples were analyzed, both of which were obtained from the Stem Cell Omics Repository [128]. Comet is a powerful PSM tool for the analysis of high- and low-resolution mass spectrometry raw data, which makes it one of the first options for any mass spectrometry user. The more peptides an algorithm can identify confidently, the more proteins that can be confirmed in a biological sample, which is one of the goals of any PSM tool.

Table 2.1: Search speed and number of total peptide identifications at 1% Q-value for X!Tandem, OMSSA, and Comet on low-resolution and high-resolution MS2 spectra. This table is obtained from [127].

	Low-resolution Ion-trap MS2		High-resolution Orbitrap MS2	
	# IDs	Runtime (s)	# IDs	Runtime (s)
X!Tandem	7,329	46	8,003	19
OMSSA	7,383	112	7,179	24
Comet	8,013	82	8,376	84

Andromeda Since the publication of Mascot back in the 1990s, it has become the "gold standard" in proteomics data analysis. The only constraint that is still facing researches is that it is a server-based interface that limits the access and robustness. Hence, Cox and Mann from the Max Blanck Institute of Biochemistry, Germany, started working on developing a tool that handles big experiments style mass spectrometry data offline. Andromeda is the name of the algorithm they developed and made available in 2011. This algorithm has become available as a stand-alone and also implemented in the MaxQuant platform for mass spectrometry data analysis [129]. Andromeda shares the same probability-based approach used by Mascot to match sequences from a database to tandem spectra. However, it is capable of handling searches for multiple modifications of the same peptide. Because shotgun proteomics samples are incredibly complex, two peptides are frequently "co-isolated" and fragmented at the same time. While many algorithms have addressed the issue of "co-isolated" peptides in the past, there was no clear-cut method to achieve the identification of both peptides [130–133]. With the implementation of Andromeda within the MaxQuant platform, it is now possible to deconvolute the fragmentation data and confidently identify more than one peptide in a fragmentation spectrum [129].

Global Post-Translational Modification Discovery (G-PTM-D) The goal of classical shotgun proteomics is to identify as many PSMs as possible to cover most of the proteome. This is possible if the generated peptides are not modified. However, a considerable number of peptides are prone to be modified either during processes within the cell or during sample preparation. It is not practical to search even a moderately complex proteomics mass spectrometry data file for more than five variable

modifications at the same time. Doing so results in a high false discovery rate (FDR); higher than 5% which is not accepted. Therefore, a new strategy that was implemented in an algorithm that was first published by Shortreed from the University of Wisconsin in 2015. The standard strategy to identify PTMs is a variable modification search, where a particular PTM is allowed to be present in any amino acid in any peptide in the experiment (like phosphorylation on serine, threonine, or tyrosine), which generates enormously large search space that burden supercomputers. Therefore, a two-pass search strategy that was suggested by Craig and Beavis (2003) [124] and then by Gupta, *et al*., (2011) [134] to reduce the search space was novel but may introduce bias when employed with target-decoy FDR. Now, the strategy that was introduced by Global-PTM-Discovery (G-PTM-D) addressed these problems by implementing an alteration of the classical variable modification approach to consider only previously curated PTMs at specific amino acid residue positions [135]. This residue position specificity generates only a modest increase in database search space. In 2017, the same research group published the graphical user interface version of MetaMorpheus that implements G-PTM-D. The three stages search includes 1) first-round database search with wide precursor tolerance (e.g., □1000 Da), 2) addition of discovered potential modifications to the database, and 3) second-round database search with narrow precursor tolerance (e.g., □10 PPM) [136].

Label-free Protein Quantification

One of the primary goals of conducting a proteomics-based mass spectrometry experiment is to generate an in-depth survey of the proteins present in a sample. However, it is now acutely aware to most scientists that a long list of proteins on its own

is unlikely to answer the questions being asked. Therefore, many database search engines have incorporated a quantification step, where MS and/or MS/MS information are used to generate quantitative values that are associated with the protein levels in the biological samples. It is assumed that the measured signal has a linear dependence on the amount of materials in the biological samples [137]. Quantitative proteomics represents a significant expansion in the protein identification process, enabling the comparison of single protein expression across different experimental treatments [138]. In fact, it has become essential for every mass spectrometry database analysis platform to have a quantitative step in their pipeline. The following section discusses label-free quantification strategies.

MS- and MS/MS-based Label-free protein quantification

In general, there are two approaches to label-free protein quantification in mass spectrometry-based proteomics data analysis. One is based on MS signal intensity, and the second method is based on MS/MS, which is referred to as spectral counting. The first method is useful when applied to high-resolution mass spectrometry data. Briefly, MS signal intensity is transformed into Label-Free Quantification (LFQ) values [139] by either summing selected discriminatory peptides intensity or summing observed peptide intensities with to respect the protein molecular weight and report Intensity Based Absolute Quantification (iBAQ) values [140]. LFQ quantification addressed one of the significant challenges in quantitative shotgun proteomics, which was sample to sample normalization. This may occur from up-front sample separation and results in data variability across the injected materials within fractions [139]. Although LFQ is the preferred method in protein quantification, a considerable number of researchers in the

field use the MS2 method for protein quantification [141–147]. The reasons behind this are access to high-resolution instrumentation, and the ability to compare MS/MS quantification values across different experimental setups [148]. The most popular methods for MS/MS relative quantification are Normalized Spectral Index (SI_N) [149], Exponentially Modified Protein Abundance Index (emPAI) [150] and Normalized Spectral Abundance Index (NSAF) [151]. Briefly, SI is calculated as the sum of total fragment intensities over all spectra assigned to all peptides identified for a protein [149]. The emPAI is defined by the ratio of the number of experimentally observed peptides to the number of potentially observed peptides for a protein [150]. The NSAF for a given protein is defined as the number of spectra identified for the same protein normalized by the protein length and the sum of such ratios for all proteins identified in the experiment. [151].

The General Proteomic Data Analysis Pipeline

Soon after collecting the mass spectrometry raw data, the choice of the database search engine is yours. The choice depends on the questions the research addresses. For example, what protein is present in the sample? Are there any co-purified proteins in the sample? How pure is the protein? What sort of PTM is present in the processed sample? Are there differences between treatment X and treatment Y? What protein is in high abundance relative to another set of samples that undergoes specific perturbations? To answer any of the questions above, a researcher needs to process the raw data in a general pipeline. A typical workflow includes inputs such as 1) the mass spectrometry raw data in universal file extension mzML, mgf or mzXML, 2) the referenced database that contains the entire proteome sequences in FASTA file format. This could be downloaded from any

of the online database servers like NCBI, UniProt, or Swissprot. The second step is to operate the database search engine by choosing experiment conditions like the protease used in the generation of the peptides, potential modification (like methionine oxidation and N-terminal acetylation). After that, comparisons, validations, and score generations take place under the hood. The output consists of a list of identified proteins, and their quantitative values and validation scores, but more comprehensive questions need further post-mass spectrometry data analysis. In addition to that, a list of all potential PTMs is reported along with their localization scores to confirm the presence/absence of such PTM. Figure 2.6 shows the general pipeline for in-gel protein identification/shotgun proteomics.

Proteomics Packages

Recently, research labs that developed database search engines have provided a one-stop-shop for both database searching and statistical data analysis (e.g.. MaxQuant, SearchGUI and MetaMorpheus). MaxQuant is one of such packages where researchers plug in the raw files and specify the experiment parameters and obtain an extensive list of information [152]. The MaxQuant platform is considered one of the best in performance stability, and for the ability to re-calibrate raw data. In long gradient LC experiments, two factors affect the quality of the collected data, one of which is the fact that mass measurement calibration shifts with time which may affect the precursor error assignment associated with that. Furthermore, the second factor is instrument performance, which varies with these long gradient LC experiment. These variabilities could impact the database search engine performance. Therefore, the re-calibration step implemented

within the MaxQuant platform eliminates losing essential information [152]. It is also noteworthy that MaxQuant report LFQ values, as mentioned previously.

A second "shop" is SearchGUI/Peptideshaker. What is unique about this package is the ability to use multiple database search engine and compile the outputs into one, which improves the identification quality since every search engine uses different PSM method [153]. Also, the fact that SearchGUI/Peptideshaker is a Java-based software gives it the ability to run on any operating system. In terms of protein quantification, SearchGUI/Peptideshaker only reports MS/MS quantification values, like emPAI.

Recently, MetaMorpheus was published by Lloyd lab from the University of Wisconsin-Madison in 2018 as a mass spectrometry data analysis tool that uses G-PTM-D for protein identification and PTM detection. Like MaxQuant, MetaMorpheus is able to re-calibrate mass spectrometry raw data. Unlike MaxQuant, MetaMorpheus is able to identify more PTMs reliably in a unique way, as discussed above [154].

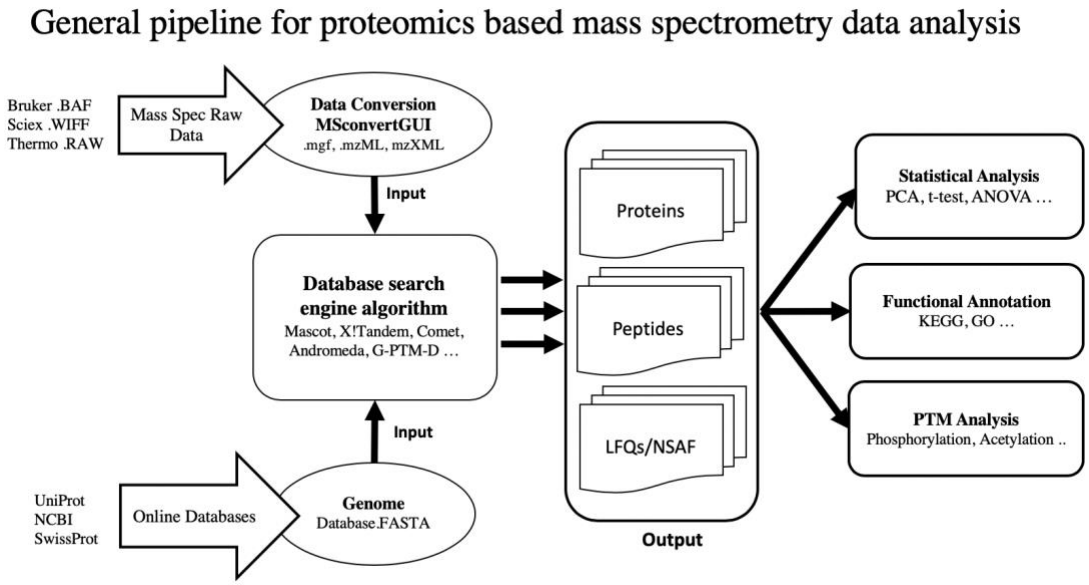


Figure 2.6 The general pipeline for proteomics-based mass spectrometry. The two inputs are the left side; 1: The mass spectrometry raw data generated from any instrument such Bruker .BAF, Sciex .WIFF or Thermo .RAW. These raw data need to be converted to a universal raw data type such Mascot Generic File (MGF) or any of the two extensible markup languages, mzML or mzXML. The second input is the reference genome sequence in FASTA file format. This file could be obtained from any of the online data bases such Uniprot, NCBI or Swissprot. The database search engine algorithm performs the PSM matching process and generate lists of identified proteins and peptides and the quantitative values associated to them. Later, the researcher can perform any of the downstream statistical analysis like PCA and t-tests. Functional Annotations also could be investigated like KEGG and GO terms. PTM analysis also is one of the downstream analysis that is also generated as an output.

Conclusion

The proteomics data analysis tools are continuously evolving to keep up with the complexity of collected raw data and improve insights into global proteome analysis. PSM tools, both older versions, and recently released software strive to identify and quantify peptides with accuracy and confidence. As is the nature with software, daily improvements are published and reported; claiming less computational power and improved identifications. Yet, every day, a new tool reports a new strategy that improves the peptide identification that requires less computational power. Protein quantification translates a long list of proteins into a meaningful piece of information that allows functional pathway annotation through Gene Ontology (GO) databases, and multiple omics integrations. The “one-stop-shop” packages represent the easiest and fastest tools for researchers with a less computational and statistical background. To conclude, proteomics data analysis tools are critical aspects of proteomics research that every researcher needs to be knowledgeable about.

Reference Cited

97. James P. Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys*. 1997;30(4):279-331. http://www.ncbi.nlm.nih.gov/pubmed/9634650. Accessed January 11, 2019.

98. Feng J, Zhang Q, Zhou Y, et al. Integration of Proteomics and Metabolomics Revealed Metabolite–Protein Networks in ACTH-Secreting Pituitary Adenoma. *Front Endocrinol (Lausanne)*. 2018;9:678. doi:10.3389/fendo.2018.00678

99. Simithy J, Sidoli S, Garcia BA. Integrating Proteomics and Targeted Metabolomics to Understand Global Changes in Histone Modifications. *Proteomics*. 2018;18(18):1700309. doi:10.1002/pmic.201700309

100. Smoyer W, Klein J. Integrating Proteomics and Metabolomics to Understand Pediatric Glomerular Disease Pathophysiology and Prognosis. http://grantome.com/grant/NIH/R01-DK110077-02. Accessed January 11, 2019.

101. Li X, Mu L, Hu X. Integrating proteomics, metabolomics and typical analysis to investigate the uptake and oxidative stress of graphene oxide and polycyclic aromatic hydrocarbons. *Environ Sci Nano*. 2018;5(1):115-129. doi:10.1039/C7EN00803A

102. Perna F, Berman SH, Soni RK, et al. Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. *Cancer Cell*. 2017;32(4):506-519.e5. doi:10.1016/j.ccell.2017.09.004

103. Gene Regulation, Transcriptomics and Proteomics. https://www.caister.com/hsp/abstracts/borrelia/04.html. Accessed January 11, 2019.

104. Haider S, Pal R. Integrated analysis of transcriptomic and proteomic data. *Curr Genomics*. 2013;14(2):91-110. doi:10.2174/1389202911314020003

105. Taylor G. Disintegration of water drops in an electric field. *Proc R Soc London Ser A Math Phys Sci*. 1964;280(1382):383-397. doi:10.1098/rspa.1964.0151

106. Edmond de H, Vincent S. *Mass Spectrometry: Principles and Applications*. (John Wiley & Sons L, ed.). Toronto; 2003.

107. Siuzdak G. *The Expading Role of Mass Spectrometry in Biotechnology*. 2nd ed. San Diego, California: MCC Press; 2006.

108. B.A. Mamyrin, V.I. Karataev, D.V. Shmikk VAZ. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *JETP*. 1972;37(1):45. http://www.jetp.ac.ru/cgi-bin/e/index/e/37/1/p45?a=list. Accessed May 11, 2019.

109. Deutsch EW, Mendoza L, Shteynberg D, et al. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*. 2010;10(6):1150-1159. doi:10.1002/pmic.200900375

110. Zhang Y, Fonslow BR, Shan B, Baek M, Yates JR. Protein Analysis by Shotgun / Bottom-up Proteomics. 2013:1-101. doi:10.1021/cr3003533.Protein

111. Loo JA, Udseth HR, Smith RD. Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Anal Biochem*. 1989;179(2):404-412. doi:10.1016/0003-2697(89)90153-X

112. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR, III. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev*. 2013;113(4):2343-2394. doi:10.1021/cr3003533

113. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004;101(26):9528-9533. doi:10.1073/pnas.0402700101

114. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2

115. Matthiesen R, Trelle MB, Højrup P, Bunkenborg J, Jensen ON. VEMS 3.0: Algorithms and Computational Tools for Tandem Mass Spectrometry Based Identification of Post-translational Modifications in Proteins. *J Proteome Res*. 2005;4(6):2338-2347. doi:10.1021/pr050264q

116. Matthiesen R, Bunkenborg J, Stensballe A, Jensen ON, Welinder KG, Bauw G. Database-independent, database-dependent, and extended interpretation of peptide mass spectra in VEMS V2.0. *Proteomics*. 2004;4(9):2583-2593. doi:10.1002/pmic.200300792

117. Matthiesen R, Lundsgaard M, Welinder KG, Bauw G. Interpreting peptide mass spectra by VEMS. *Bioinformatics*. 2003;19(6):792-793. doi:10.1093/bioinformatics/btg059

118. Listgarten J, Emili A. Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry. *Mol Cell Proteomics*. 2005;4(4):419-434. doi:10.1074/mcp.R500005-MCP200

119. Jimmy KE, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994;5(11):976-989. doi:10.1016/1044-0305(94)80016-2

120. Nesvizhskii AI, Aebersold R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discov Today*.

2004;9(4):173-181. doi:10.1016/S1359-6446(03)02978-7

121. Hunt DF, Yates JR, Shabanowitz J, Winston S, Hauer CR. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 1986;83(17):6233-6237. http://www.ncbi.nlm.nih.gov/pubmed/3462691. Accessed January 13, 2019.

122. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*. 1993;3(6):327-332. http://www.ncbi.nlm.nih.gov/pubmed/15335725. Accessed January 13, 2019.

123. Akhtar MN, Southey BR, Andrén PE, Sweedler J V, Rodriguez-Zas SL. Evaluation of database search programs for accurate detection of neuropeptides in tandem mass spectrometry experiments. *J Proteome Res*. 2012;11(12):6044-6055. doi:10.1021/pr3007123

124. Craig R, Beavis RC. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun Mass Spectrom*. 2003;17(20):2310-2316. doi:10.1002/rcm.1198

125. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004;20(9):1466-1467. doi:10.1093/bioinformatics/bth092

126. Eng JK, Fischer B, Grossmann J, MacCoss MJ. A Fast SEQUEST Cross Correlation Algorithm. *J Proteome Res*. 2008;7(10):4598-4602. doi:10.1021/pr800420s

127. Eng JK, Jahan TA, Hoopmann MR. Comet: An open-source MS/MS sequence database search tool. *Proteomics*. 2013;13(1):22-24. doi:10.1002/pmic.201200439

128. Phanstiel DH, Brumbaugh J, Wenger CD, et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat Methods*. 2011;8(10):821-827. doi:10.1038/nmeth.1699

129. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res*. 2011;10(4):1794-1805. doi:10.1021/pr101065j

130. Zhang N, Li X, Ye M, Pan S, Schwikowski B, Aebersold R. ProbIDtree: An automated software program capable of identifying multiple peptides from a single collision-induced dissociation spectrum collected by a tandem mass spectrometer. *Proteomics*. 2005;5(16):4096-4106. doi:10.1002/pmic.200401260

131. Bern M, Finney G, Hoopmann MR, Merrihew G, Toth MJ, MacCoss MJ. Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem*. 2010;82(3):833-841. doi:10.1021/ac901801b

132. Wang J, Pérez-Santiago J, Katz JE, Mallick P, Bandeira N. Peptide Identification from Mixture Tandem Mass Spectra. *Mol Cell Proteomics*. 2010;9(7):1476-1485. doi:10.1074/mcp.M000136-MCP201

133. Houel S, Abernathy R, Renganathan K, Meyer-Arendt K, Ahn NG, Old WM. Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies. *J Proteome Res*. 2010;9(8):4152-4160. doi:10.1021/pr1003856

134. Gupta N, Bandeira N, Keich U, Pevzner PA. Target-Decoy Approach and False Discovery Rate: When Things May Go Wrong. *J Am Soc Mass Spectrom*. 2011;22(7):1111-1120. doi:10.1007/s13361-011-0139-3

135. Shortreed MR, Wenger CD, Frey BL, et al. Global Identification of Protein Post-translational Modifications in a Single-Pass Database Search. *J Proteome Res*. 2015;14(11):4714-4720. doi:10.1021/acs.jproteome.5b00599

136. Qiyao L, Shortreed MR, Wenger CD, et al. Global Post-Translational Modification Discovery. *J Proteome Res*. 2017;16(4):1383-1390. doi:10.1021/acs.jproteome.6b00034

137. Zhang G, Ueberheide BM, Waldemarson S, et al. Protein quantitation using mass spectrometry. *Methods Mol Biol*. 2010;673:211-222. doi:10.1007/978-1-60761-842-3_13

138. Wang M, You J, Bemis KG, Tegeler TJ, Brown DPG. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Briefings Funct Genomics Proteomics*. 2008;7(5):329-339. doi:10.1093/bfgp/eln031

139. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13(9):2513-2526. doi:10.1074/mcp.M113.031591

140. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098

141. Pisanu S, Ghisaura S, Pagnozzi D, et al. The sheep milk fat globule membrane proteome. *J Proteomics*. 2011;74(3):350-358. doi:10.1016/J.JPROT.2010.11.011

142. Zybailov BL, Florens L, Washburn MP. Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Mol Biosyst*. 2007;3(5):354. doi:10.1039/b701483j

143. Arike L, Peil L. Spectral Counting Label-Free Proteomics. In: Humana Press, New York, NY; 2014:213-222. doi:10.1007/978-1-4939-0685-7_14

144. Paoletti AC, Parmely TJ, Tomomori-Sato C, et al. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci U S A*. 2006;103(50):18928-18933. doi:10.1073/pnas.0606379103

145. Dheilly NM, Raftos DA, Haynes PA, Smith LC, Nair S V. Shotgun proteomics of coelomic fluid from the purple sea urchin, Strongylocentrotus purpuratus. *Dev Comp Immunol*. 2013;40(1):35-50. doi:10.1016/J.DCI.2013.01.007

146. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

147. Timmins-Schiffman E, Coffey WD, Hua W, Nunn BL, Dickinson GH, Roberts SB. Shotgun proteomics reveals physiological response to ocean acidification in Crassostrea gigas. *BMC Genomics*. 2014;15(1):951. doi:10.1186/1471-2164-15-951

148. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem*. 2007;389(4):1017-1031. doi:10.1007/s00216-007-1486-6

149. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

150. Ishihama Y, Oda Y, Tabata T, et al. Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein. *Mol Cell Proteomics*. 2005;4(9):1265-1272. doi:10.1074/mcp.M500061-MCP200

151. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces c erevisiae*. *J Proteome Res*. 2006;5(9):2339-2347. doi:10.1021/pr060161n

152. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

153. Vaudel M, Burkhart JM, Zahedi RP, et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotechnol*. 2015;33(1):22-24. doi:10.1038/nbt.3109

154. Solntsev SK, Shortreed MR, Frey BL, Smith LM. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J Proteome Res*. 2018;17(5):1844-1851. doi:10.1021/acs.jproteome.7b00873

CHAPTER THREE

PROTEOMICS ANALYSIS OF THE TRANSITION BETWEEN AEROBIC AND ANAEROBIC GROWTH CONDITIONS IN ESCHERICHIA COLI

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Mohammed Y. Refai

Contributions: Performed data analyses, interpreted results, and wrote manuscript.

Co-Author: Nina Paris

Contributions: Sample preparation, performed MS data collection and analysis.

Co-Author: Monika Tokmina-Lukaszewska

Contributions: Sample preparation, performed MS data collection and analysis, discussed results and interpretations.

Co-Author: Ganesh Balasubramanian

Contributions: Sample preparation, performed MS data collection and analysis, discussed results and interpretations.

Co-Author: Brian Bothner

Contributions: Conceived the study, discussed results and interpretations, and edited manuscript.

Manuscript Information Page

Authors: Mohammed Refai, Nina Paris, Hunter Fausset, Monika Tokmina-Lukaszewska

Journal: Biochimica et Biophysica Acta

Status of Manuscript:

_X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Publisher: Elsevier

ABSTRACT

Escherichia coli is a gram-negative bacteria and a facultative anaerobe. The ability to switch metabolic programs in response to environmental factors, such as oxygen availability, is key to survival and pathogenicity. Thus, understanding how *E. coli* orchestrates the physiological transition between states is of value. A critical factor in the process and a gap in knowledge is a global understanding of proteins regulated by cellular redox potential. We have used shotgun proteomics and thiolomics methods to investigate whole-cell differences between E. coli in aerobic and anaerobic environments. Over 1,000 proteins were tracked in the two experimental groups. Functional annotation enrichment analysis indicated higher rates of glucose utilization through the methylglyoxal pathway in aerobic conditions, whereas amino acid degradation and TCA cycle associated enzymes are upregulated in the anaerobic group. Chemical tagging of cysteines indicated that the proteome was more oxidized when growing in anaerobic conditions. Analysis at the level of cysteine-containing peptides revealed that thiol reactivity is protein and residue-specific. The unexpected result that the proteome is more oxidized during anaerobic growth and that a large number of proteins have site-specific thiol redox potentials indicates that protein-based cysteines play has a wider role in cellular signaling and metabolic tuning than previously appreciated.

INTRODUCTION

Escherichia coli is a gram-negative bacteria that can live in aerobic and anaerobic environments. The ability to grow in the presence or absence of oxygen is rooted in the genome, and the transition between states is accompanied by specific changes in gene expression patterns [1]. As an inhabitant of the lower intestine, *E. coli* often occupies oxygen-starved nitches, however, it must be capable of surviving in an oxygen-rich environment if from the host. The ability to switch metabolic programs in response to environmental factors, such as oxygen availability, is key to survival and pathogenicity [21,30,155]. The availability of electron donors and acceptors plays an important role in the energetic state of an organism and changes significantly when shifting between aerobic and anaerobic conditions [22]. In fact, the transition between aerobic and anaerobic is associated with *E. coli* pathogenicity in mouse intestine [156]. Recently, Murashko, *et al.*, reported changes in the *E. coli* morphology in response to changes in the growing environment [157]. In anaerobic conditions, the organism adopts a filamentous shape that is linked to bacterial pathogenesis. Understanding the role of protein oxidation in the transition of *E. coli* during metabolic transitions will enhance our fundamental knowledge about redox-regulated proteins and redox signaling.

The anaerobic/aerobic transition in *E. coli* is tightly controlled through signaling mechanisms that are sensitive to oxygen. Intracellular signals lead to changes in gene and protein expression patterns, altering metabolic potential. This includes the expression of proteins that mitigate oxidative damage that accompanies the presence of reactive oxygen

species [158–160]. Even though Reactive Oxygen Species (ROS) are often considered the "bad boys" of biology, they play an important role in cellular redox signaling [161]. In the last two decades, it has become clear that ROS serve as signaling molecules to regulate physiological and biological processes [162]. In yeast, the global transcriptional regulator Yap1p senses the presence of ROS and responds by upregulating antioxidant proteins [163]. Recently, a mountain of evidence supports ROS plays an important role in a wide range of signaling process include differentiation, development, and growth has come out [164,165]. In plants, an interesting example is the pathogen-responsive NPR1 (non-expresser of pathogenesis-related gene 1) protein. [166]. NPR1 forms disulfide bridge oligomers and resides in the cytosol, however, in response to a pathogen, NPR1 become reduced monomers and translocate to the nucleus where it activates plant immune responses [167] While, it has long been accepted that excess ROS are deleterious, low levels can lead to favorable physiologic consequences as well [44] supporting the importance of maintaining and balancing homeostasis [168].

ROS, including hydrogen peroxide (H_2O_2), peroxide (O_2^{-2}), superoxide anion ($O_2^{\bullet-}$) and hydroxyl radical ($OH^{\bullet}$) are formed spontaneously as byproducts of normal oxygen metabolism. High ROS levels can trigger the rapid oxidation of transcription factors, including oxyR in bacteria [49]. Moreover, redox-regulated chaperones, such as Hsp33 and peroxiredoxin, are also affected by the high ROS levels [50,51]. Importantly, high redox levels alter the function of many redox-sensitive glycolytic enzymes, like GAPDH, to re-route metabolic fluxes essential for the energy production and redox homeostasis [46]. These downstream effects are not necessarily a direct response to oxidative stress, yet it has a significant influence on the overall metabolic process.

The oxygen-sensitive transcription factor oxyR is a much-studied redox-sensitive protein in bacteria. In *E. coli,* it is a master regulator that responds to specific post-translational modifications [160]. The ROS sensing mechanism of oxyR involves the formation of disulfide bonds between Cys199 and Cys208 [160]. For example, in response to H_2O_2, the oxidized form of oxyR binds to a promoter that induces the expression of antioxidant genes [159]. However, the controlled genes, like glutaredoxin 1, in turn, results in the reduction of oxyR at Cys199 as an autoregulation mechanism to repress the expression of antioxidant proteins [160]. Another set of genes also could be expressed in response oxyR nitrosylation at Cys199 that are important for the anaerobic growth [158]. OxyR is an example of redox-sensitive signaling molecule since the function is altered by oxidative state of its reactive thiol groups.

E. coli is a much-used model system and has been the subject of numerous "Omics" based analyses including genomics [2–7], transcriptomics [8–11], proteomics [12–15], and metabolomics [16–20]. Our interest is in redox-based signaling that occurs through the modification of protein thiols. A bottom-up proteomics approach using cysteine directed protein chemistry was used to generate a global profile of cysteine modifications, in particular, thiolomics. Thiolomics is the study of all thiol-containing proteins as a subset of the entire diverse organism proteome.

To our surprise, a global analysis of cysteine redox states in *E. coli,* and the role of thiol signaling in aerobic and anaerobic conditions had yet to be reported. The goal of this study was to understand the oxidative states of the reactive cysteine residues that are associated with the changes in the central carbon metabolic pathways enzymes. To achieve this, *E. coli* was used in this study as a model organism growing in two

conditions, aerobic and anaerobic. To investigate the oxidized state of the cysteine-containing proteins, we used differential alkylation approach to capture oxidative states of proteins. Mass spectrometry-based proteomics was used to investigate the global proteome to assess the metabolic state of the bacteria growing aerobically and anaerobically. Gel-based proteomics was applied to investigate the overall oxidative state of the entire set of proteins in the two experimental groups. Moreover, this experimental plan explores the oxidative state that are associated with enzymes regulated either aerobically or anaerobically, and that could reveal signaling pathways that are with these enzymes and eventually, the regulated pathways.

MATERIALS & METHODS

All chemicals were purchase from Sigma unless noted.

IAM (sigma, cas: 144-48-9), NEM (sigma, cas: 128-53-0), Urea (biosciences, cas: 57-13-6), DTT (sigma, cas: 3483-12-3), ABC (sigma, CAS Number 1066-33-7) and TRIS (sigma, CAS Number: 77-86-1).

Bacterial Strain and Growing Conditions

E. coli strain MG1655 (K12) cultures were growing in batches of LB media at 37°C with continuous sparging of either atmospheric air or Nitrogen gas to achieve aerobic and anaerobic conditions, respectively. Briefly, cultures were started by inoculating 50 mL of sterile LB media with 5 uL of an overnight starter culture grown at 37°C. After 18 hours, 100 uL of the starting culture is transferred into four different growing flasks containing 100 mL of LB and stirred at 300 rpm. Four biological

replicates were collected for every treatment. Growth was monitored by UV-VIS at 600 nm for the first 12 hours to determine the exponential and stationary phases. After that, 10 mL of each flask and at mid-exponential phase, at 0.4 OD for the aerobic samples and 0.3 OD for the anaerobic samples. Immediately after the harvesting step, all samples were centrifuged at 4,000 rpm for 10 minutes. The supernatant was discarded, and the cell pellet was processed immediately to avoid losing the oxidative states information and another batch of the cell pellet were stored for future analysis.

Protein Extraction and Differential Alkylation

The process was performed in parallel for all biological replicates to minimize variation. Cell pellets were carefully washed with PBS buffer and then resuspended in 100uL of cold extraction buffer containing 0.1M Tris-HCl pH 7.5 (4 °C), and 8M urea. The resulted suspension was subjected to three cycles of freeze/thaw in liquid nitrogen. Next, the samples were sonicated for 5 min at 40% duty using ultrasonicate homogenizer (Biologix – Model 3000). The samples then were centrifuged at 15,000 rpm for 15 min. The resulting supernatant was transferred to new Eppendorf tubes. Cold acetone precipitation was performed to clear proteins out of small molecules. The samples were stored in -20□C for an hour and then centrifuged at 5,000 rpm for 5 min. The resulted precipitated protein from the previous step were resuspended in 0.1M Tris-HCl pH 6.8, 5uM EDTA, 50mM N-ethylmaleimide (NEM) in 6M Urea and incubated at room temperature in the dark for 30 min. Then, what is equal to 100 ug of soluble protein was transferred to a 3K MWCO Nanosep centrifugal device and centrifuged at 5,000 rpm for 5 min to discard the excess NEM. After that, 100 uL of 10 mM dithiotheretol (DTT) in

50 mM ammonium bicarbonate pH 7.8 and 6M urea was added to the top membrane and vortexed for 30 seconds, incubated at room temperature for 1 hour. After that, the centrifugal device was centrifuged to discard the excess DTT. After that, 100 uL of 50 mM iodoacetamide (IAM) in 50 mM ammonium bicarbonate pH 7.8 was added to the top membrane and incubated for 30 min at room temperature in the dark. After that, four wash steps were performed to exchange any remaining small molecules and urea in the sample using 50 mM ammonium bicarbonate pH 7.8. The samples then were subjected to sequencing grade trypsin at the ratio of 100:5 protein to protease and incubated at 37□C for 18 hours. On the next day, the resulted peptides were collected by the centrifugation the Nanosep devices for 40 min at 8,000 rpm.

Two Dimensional SDS-PAGE

Cell pellets were carefully washed with PBS buffer and then lysed with a combination of freeze/thaw and lysis buffer (30 Mm Tris-HCl pH 8.5 (4 °C), 7 M urea, 2M thiourea, 4% CHAPS, 50 Mm DTT, 0.5% IPG carrier ampholytes and a cocktail of protease inhibitors). The supernatant was clarified by centrifugation. Proteins were purified and concentrated by precipitation with acetone and resuspended for one h in lysis buffer without DTT. Protein concentration was measured using a Bradford Protein Assay (Bio-Rad). The labeling reaction was performed according to [169].

LC-MS Instrumentation, Data Acquisition, and Data Analysis

The collected peptides were dried and resolubilized in 30 uL of 95/5 (*v/v*) water/acetonitrile with 0.1% formic acid. Injection volumes of 2 uL were processed via Dionex Ultimate 3000 Nano UHPLC, with an Acclaim PepMap100 C18 column used for

trapping (100 μm × 2cm) and an Acclaim PepMap RSLC C18 (75 μm × 50 cm, C18 2 μm 100A) for final peptide separation. Chromatography was as follows: solvent consisted of H_2O with 0.1% (*v/v*) formic acid for channel 'A' and 80/20 acetonitrile/water for channel 'B'. Following sample trapping for 2 min at a flow rate of 10 μL/min, the HPLC valve was switched to elution position. From 0.0 to 5.0 min, the elution solvent pump composition was held at 7% B. From 5 to 125 min, the elution solvent gradient was linearly changed from 3% to 30% B. From 125 to 127 min, the gradient was ramped from 30% to 80% B. From 127 to 130 min, the solvent was held at 80% B, and from 130 to 133 min the solvent was linearly ramped from 80% B to 3% B. From 133 to 140 min, the elution solvent was held at 3% B. During the entire run, the loading pump solvent was held at 10 μL/min of 97% water, 3% acetonitrile, and 0.1% formic acid. The mass spectrometer used was a Bruker maXis Impact with CaptiveSpray ESI source: resolution is approximately 50,000 and accuracy is 1 ppm. Spectra were collected in positive mode from 150 to 1750 *m/z* at a maximum rate of 2 Hz for both precursor and fragment spectra, and with adaptive acquisition time for highly-abundant ions.

LCMS data was analyzed using MaxQuant (v1.6.4.0) [170]. The identification settings used in the analysis were following the MaxQuant authors [171] with few modifications; variable modifications: N-ethylmaleimide on C (+125.047679 Da), carbamidomethyl on C (+57.021464 Da) and deamidation on N or Q (-0.984016 Da). iBAQ values were used in the quantitative analysis and further statistical tests under Perseus platform, the comprehensive platform for comprehensive proteomics data analysis [172].

RESULTS

This study aims to understand the role of protein thiols in the cell related to the metabolic shift between aerobic and anaerobic states. We selected *E. coli* MG1655 (K12) for our experiments because it is a facultative anaerobe and is an often-used model organism needs [155–157,173–175]. We used the same growing conditions, which is the optimum temperature for the organism (37°C), the only factor that determines these two metabolic states is the presence or absence of oxygen. The experimental setup we used is similar to the previous studies that established aerobic and anaerobic conditions presence or absence of oxygen [21,155,175–180]. *E. coli* (K12) cells were harvested during exponential (logarithmic) phase (Figure 3.1S). Cells growing aerobically were harvested at 0.6 OD at 600 nm and the anaerobic cultures at 0.4 OD at 600 nm. The protein extraction method was optimized for soluble proteins.

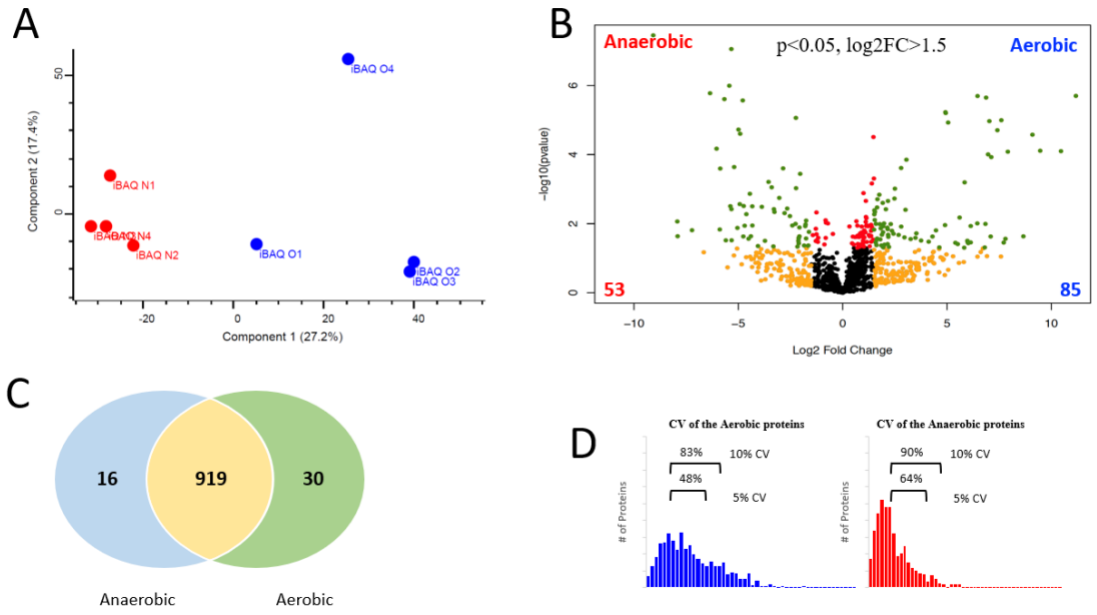


Figure 3.1 Global proteome analysis of the Aerobic and Anaerobic *E. coli* proteins. A, PCA analysis for the two experimental groups (Aerobic in Red and Anaerobic in Blue). B, Quantitative data analysis for the two experimental groups. Volcano plot for the –log10 p-value vs the log2 expression levels of the Aerobic and Anaerobic E. coli proteins. The two cutoffs are log2FC>1.5 and p-value<0.05. The green spots are for proteins that meets the two cutoffs, the yellow spots are for the proteins that meets only log2FC cutoff, the red spots are for the proteins that meets only the p-value cutoff and the black spots are for the proteins that do not meet any of the two cutoffs. C, Venn diagram for the overlapped identified proteins in the two experimental groups and the uniquely identified proteins in each group. D, Coefficient of variation Histogram for the identified proteins expression values in the two experimental groups the Aerobic in blue and the anaerobic in red.

Global Proteome Profiling

LC-MS/MS shotgun proteomics resulted in the identification of 4,107 peptides and 1,200 proteins at 1%FDR in the two experimental groups (Table S3.1). After removal of proteins not present in all biological replicates, 949 proteins remained. Intensity-Based Absolute Quantitation (iBAQ) values were used in the differential abundance analysis. Perseus software was used to perform Analysis of Variance (ANOVA) and student t-test according to the recommended settings, according to Tyanova *et al*., 2016 [172].

To assess the physiological difference between aerobic and anaerobic environments on *E. coli* MG1655 (K12), a global analysis of the proteome was conducted. A Principal Component Analysis (PCA) based on the log2 transformed iBAQ protein expression values was performed. The two experimental groups show clear separation, with the anaerobic group forming a tighter cluster (Figure 3.1A). To confirm a difference in the variability within the biological groups, coefficients of variation (CV) were compared. For the aerobic group, 48% of the proteins had a CV of < 5% while 64% for the anaerobic groups. Increasing the CV to 10%, the number of proteins increases to 83% and 90% for the aerobic and anaerobic experimental groups, respectively (Figure 3.1D). This suggests that most of the proteins in the anaerobic group behave similarly across the biological replicates (in terms of abundance), unlike the aerobic group, which indicate that the anaerobic group shows signs of growth in a stressful environment.

Based on general knowledge of cellular adaptation to oxygen, we expected to find a subset of proteins unique to each of the growth conditions. We were able to identify 919 protein that were shared by the two experimental groups with 30 and 16 proteins

found only in aerobic and anaerobic groups, respectively (Figure 3.1C). To investigate the differences among the global proteome at the level of proteins, a volcano plot was constructed for the entire list of proteins that are present in the two experimental groups. Differential analysis utilizing p-value less than 0.05 and Log2 fold change greater than 1.5 was done. The analysis showed 85 differentially regulated proteins in the aerobic group and 35 regulated proteins in the anaerobic group. The majority of regulated proteins are enzymes associated with central carbon metabolism (Table 3.1S). To understand the metabolic pathways enriched for the differentially regulated proteins in the two experimental groups, functional annotation analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology for Biological Processes (GOPB), Gene Ontology for Molecular Functions (GOMF) and Gene Ontology for Cellular Compartment (GOCC) was conducted. This was accomplished by loading the entire list of identified proteins into Perseus as described by Tyanova, *et al*., (2011) [181]. The enriched terms related to central carbon metabolism under aerobic conditions were glycolysis, fructose and mannose metabolism, hexose metabolic and catabolic metabolism ($p<0.01$, 1% FDR). Our analysis revealed that enzymes associated with glycolysis were up-regulated in the aerobic experimental group (Table 3.1). Fold changes ranged from Log2 1.3-4.4 for the glycolytic enzymes which included: ATP-dependent 6-phosphofructokinase isoenzyme 1, Fructose-1,6-bisphosphatase 1 class 2, Triosephosphate isomerase, Enolase, Glyceraldehyde-3-phosphate dehydrogenase, Acetate Kinase, Mannose-6-phosphate isomerase, Adenylate kinase, Fructose-bisphosphate aldolase class 2 and Phosphoglycerate kinase. The methylglyoxal pathway is an interesting alternative to the canonical glycolysis pathway found in bacteria that also

was up-regulated in the aerobic experimental group. When activated, this pathway balances the flow of sugar phosphate like glucose-6-phosphate, mannose-6-phosphate, and fructose-6-phosphate to the end product of the Glycolysis [182].

Table 3.1: Enrichment analysis for the associated KEGG, GOBP, GOMF and GOCC terms. The first part of the table shows the significantly enriched terms associated with the differentially regulated proteins in the Aerobic experimental group ($p<0.05$) and the second part shows the significantly enriched terms associated with the differentially regulated proteins in the Anaerobic experimental group ($p<0.05$)

Experimental Group	Annotation class	Annotation term	# of Proteins	Enrichment factor	p-value	Benj. Hoch. FDR
Aerobic	GOBP	**Glycolysis**	**10**	5.9752	7.494E-07	0.00035415
	KEGG	**Glycolysis / Gluconeogenesis**	**14**	5.4807	1.4393E-08	1.9287E-06
	KEGG	Methane metabolism	7	4.9669	0.00017322	0.0092846
	KEGG	**Fructose and mannose metabolism**	**8**	4.7802	7.9396E-05	0.0053195
	GOBP	**Glucose catabolic process**	**10**	4.3665	2.4123E-05	0.007705
	GOBP	**Glucose metabolic process**	**12**	3.5851	3.4289E-05	0.0094524
	GOBP	**Hexose catabolic process**	**12**	3.4059	6.0045E-05	0.012754
	GOBP	**Hexose metabolic process**	**14**	3.1165	4.1435E-05	0.0097906
	GOBP	**Monosaccharide catabolic process**	**14**	2.9989	6.5543E-05	0.012754
	GOBP	Alacohol catabolic process	16	2.9298	2.5621E-05	0.007705
	GOBP	Alcohol metabolic process	20	2.3408	7.5124E-05	0.013806
Anaerobic	GOBP	**Glyoxylate metabolic process**	**4**	14.566	3.8695E-05	0.0097906
	KEGG	**Citrate cycle (TCA cycle)**	**14**	13.416	3.5767E-15	9.5856E-13
	GOBP	**Tricarboxylic acid cycle**	**12**	12.852	1.0419E-12	1.1489E-09
	GOBP	**Acetyl-CoA catabolic process**	**12**	12.852	1.0419E-12	1.1489E-09
	GOBP	Cofactor catabolic process	12	12.138	2.9906E-12	1.9786E-09
	GOBP	Coenzyme catabolic process	12	12.138	2.9906E-12	1.9786E-09
	GOBP	**Acetyl-CoA metabolic process**	**13**	11.835	5.1782E-13	1.1489E-09
	KEGG	**Carbon fixation pathways in prokaryotes**	**8**	7.6663	2.1708E-06	0.00019392
	KEGG	**Glyoxylate and dicarboxylate metabolism**	**5**	7.5865	0.00022971	0.01026
	GOMF	4 iron, 4 sulfur cluster binding	9	4.9657	2.8234E-05	0.019453
	GOBP	Coenzyme metabolic process	14	4.0461	1.7256E-06	0.00071352
	GOMF	Coenzyme binding	13	3.3814	0.00003437	0.019453
	GOCC	Protein complex	14	3.268	2.4219E-05	0.0058609
	GOMF	Cofactor binding	17	3.1585	4.178E-06	0.009459
	GOBP	Cellular catabolic process	16	3.0346	1.4829E-05	0.0054506
	GOBP	Cofactor metabolic process	14	2.9989	6.5543E-05	0.012754
	GOBP	Catabolic process	23	2.7193	6.9713E-07	0.00035415
	GOMF	Oxidoreductase activity	21	2.451	1.5216E-05	0.017225

In contrast, glyoxylate metabolic processes, citrate cycle (TCA cycle) and acetyl-CoA were included in the enriched terms for the anaerobic group ($p<0.01$, 1% FDR). Moreover, the glutamate synthase [NADPH] large chain (gltB) and glutamate synthase [NADPH] small chain (gltD) were differentially upregulated by 2.2- and 1.3-fold change

($p<0.01$) in the anaerobic experimental group, respectively. These two enzymes increase amino acids utilization in the TCA cycle to maximize the energy yield. Most of the TCA cycle enzymes are differentially up-regulated in the anaerobic group including Fumarate hydratase class I (anaerobic), Aconitate hydratase B, 2-oxoglutarate dehydrogenase E1 component, Citrate synthase, Fumarate hydratase class I (aerobic), Succinate dehydrogenase flavoprotein subunit, and Succinate dehydrogenase iron0sulfer subunit. Most fold changes ranged from 2.4-4.5 in comparison to the aerobic experimental group, which supports the dependence on TCA cycle for energy generation. (Table 3.1). We also found that Glyoxylate cycle enzymes were up-regulated in the anaerobic experimental group like malate synthase G with 3-fold changes.

Table 3.2 The differential analysis comparison between the two experimental groups. Cluster 1 is the proteins up-regulated in the Aerobic group; Cluster 2 is for the proteins up-regulated in the Anaerobic group. Log2FC is the log2 fold change where the negative values are higher in Anaerobic group and the positive values.

Cluster	Peptides	Mol. weight [kDa]	MS/MS count	-Log T-test p-value	Log2FC	Protein IDs	Protein names	Gene names
1	2	6.3715	6	1.30	-6.33	P0A7N9	50S ribosomal protein L33hghf	rpmG
1	3	11.306	55	1.54	-5.92	P0AEH5	Protein ElaB	elaB
1	5	11.852	25	2.01	-4.96	P0AB46	Uncharacterized protein YmgD	ymgD
1	5	12.141	8	3.28	-4.89	P0A8U6	Met repressor	metJ
1	17	61.529	43	3.40	-4.49	P0A6T1	Glucose-6-phosphate isomerase	pgi
1	3	11.905	7	1.71	-4.44	P67603	UPF0267 protein YqfB	yqfB
1	7	24.073	21	2.00	-3.88	P0A761	Putative N-acetylmannosamine-6-phosphate 2-epimerase	nanE
1	5	32.593	22	1.38	-3.87	P0A6L4	N-acetylneuraminate lyase	nanA
1	3	28.642	8	1.82	-3.81	P0AFZ1	Protein SseB	sseB
1	12	90.398	27	2.35	-3.80	P18775	Dimethyl sulfoxide reductase DmsA	dmsA
1	2	22.869	5	2.47	-3.75	P18776	Anaerobic dimethyl sulfoxide reductase chain B	dmsB
1	5	13.304	20	2.17	-3.63	P05706	Glucitol/sorbitol-specific phosphotransferase enzyme IIA component	srlB
1	1	8.7527	3	1.42	-3.56	P64503	Uncharacterized protein YebV	yebV
1	13	18.695	63	2.52	-3.50	P0ABT2	DNA protection during starvation protein	dps

1	10	31.19	30	1.79	-3.40	P31658	Molecular chaperone Hsp31 and glyoxalase 3	hchA
1	35	96.126	219	3.02	-3.30	P0A9Q7	Aldehyde-alcohol dehydrogenase;Alcohol dehydrogenase;Acetaldehyde dehydrogenase [acetylating];Pyruvate-formate-lyase deactivase	adhE
1	7	27.864	28	2.13	-3.26	P0AEK4	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI	fabI
1	5	14.89	4	1.31	-3.20	P19931	Hydrogenase-1 operon protein HyaE	hyaE
1	5	34.842	17	2.15	-3.12	P0A796	ATP-dependent 6-phosphofructokinase isozyme 1	pfkA
1	2	13.696	4	2.06	-3.01	P0ADK8	Uncharacterized protein YibL	yibL
1	2	15.094	5	2.05	-2.95	P0A6Z6	Nickel-responsive regulator	nikR
1	3	18.858	4	2.51	-2.95	P45470	Protein YhbO	yhbO
1	7	36.082	16	1.69	-2.89	P76193	Probable L,D-transpeptidase YnhG	ynhG
1	1	7.1771	5	1.34	-2.88	P0ADN6	Uncharacterized lipoprotein YifL	yifL
1	7	22.632	19	1.37	-2.87	P76014	PTS-dependent dihydroxyacetone kinase, ADP-binding subunit DhaL	dhaL
1	7	31.109	39	1.76	-2.79	Q46857	2,5-diketo-D-gluconic acid reductase A	dkgA
1	5	23.353	15	1.73	-2.76	P0A7J0	3,4-dihydroxy-2-butanone 4-phosphate synthase	ribB
1	2	35.934	4	1.81	-2.74	P0A8K1	Phosphatidylserine decarboxylase proenzyme;Phosphatidylserine decarboxylase alpha chain;Phosphatidylserine decarboxylase beta chain	psd
1	3	33.675	14	1.52	-2.68	P76187	Oxidoreductase YdhF	ydhF
1	3	17.751	5	2.06	-2.65	P37182	Hydrogenase 2 maturation protease	hybD
1	8	35.232	21	1.44	-2.64	P0AGF6	L-threonine dehydratase catabolic TdcB	tdcB
1	5	35.852	6	2.02	-2.58	P0A9C9	Fructose-1,6-bisphosphatase 1 class 2	glpX
1	3	29.436	4	2.40	-2.56	P30863	2,5-diketo-D-gluconic acid reductase B	dkgB
1	3	46.11	6	1.68	-2.54	P06988	Histidinol dehydrogenase	hisD
1	2	19.997	9	1.66	-2.53	P0ABZ4	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase KdsC	kdsC
1	20	52.022	140	2.98	-2.49	P0ADG7	Inosine-5-monophosphate dehydrogenase	guaB
1	11	32.456	48	2.48	-2.43	P06999	ATP-dependent 6-phosphofructokinase isozyme 2	pfkB
1	4	36.361	8	1.94	-2.42	P0A6S7	Glycerol-3-phosphate dehydrogenase [NAD(P)+]	gpsA
1	4	21.082	3	1.60	-2.41	P0A7D1	Peptidyl-tRNA hydrolase	pth
1	14	62.49	24	1.68	-2.40	P0ACE0	Hydrogenase-2 large chain	hybC
1	4	31.488	5	2.91	-2.34	P0AG84	Uncharacterized oxidoreductase YghA	yghA
1	1	10.68	1	1.71	-2.15	P64602	Probable phospholipid ABC transporter-binding protein MlaB	mlaB
1	17	56.193	139	1.86	-2.03	P37689	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	gpmI
1	7	20.563	32	1.90	-2.01	P0ACR9	Transcriptional repressor MprA	mprA
1	28	50.729	336	3.60	-1.96	P0AD61	Pyruvate kinase I	pykF

1	1	11.426	4	2.83	-1.86	P69795	N,N-diacetylchitobiose-specific phosphotransferase enzyme IIB component	chbB
1	16	36.85	106	1.72	-1.85	P00805	L-asparaginase 2	ansB
1	8	15.935	67	1.58	-1.77	P39177	Universal stress protein G	uspG
1	2	26.8	5	3.03	-1.77	P0A9V1	Lipopolysaccharide export system ATP-binding protein LptB	lptB
1	4	15.603	39	3.07	-1.75	P0AF90	Regulator of ribonuclease activity B	rraB
1	2	21.691	2	2.44	-1.75	P0A729	Maf-like protein YceF	yceF
1	3	46.161	12	1.36	-1.74	P11557	Cell division protein DamX	damX
1	17	45.654	649	2.57	-1.65	P0A6P9	Enolase	eno
1	12	26.972	180	1.94	-1.61	P0A858	Triosephosphate isomerase	tpiA
1	7	15.096	45	2.23	-1.56	P62768	UPF0325 protein YaeH	yaeH
1	40	95.584	287	2.79	-1.55	P63284;P63284-2	Chaperone protein ClpB	clpB
1	20	35.532	553	1.64	-1.51	P0A9B2	Glyceraldehyde-3-phosphate dehydrogenase A	gapA
1	12	35.047	102	1.45	-1.51	P69797	PTS system mannose-specific EIIAB component;Mannose-specific phosphotransferase enzyme IIA component;Mannose-specific phosphotransferase enzyme IIB component	manX
1	7	11.316	98	1.33	-1.49	P60624	50S ribosomal protein L24	rplX
1	5	20.591	55	3.21	-1.47	P0A6N4	Elongation factor P	efp
1	6	15.088	30	2.49	-1.45	P0C0L2	Peroxiredoxin OsmC	osmC
1	6	15.347	98	1.32	-1.38	P0ACG1	DNA-binding protein StpA	stpA
1	3	16.918	21	1.68	-1.37	P0A731	Methylglyoxal synthase	mgsA
1	7	37.784	21	1.35	-1.36	P0AGJ5	Uncharacterized tRNA/rRNA methyltransferase YfiF	yfiF
1	25	41.118	771	3.72	-1.35	P0A799	Phosphoglycerate kinase	pgk
1	15	39.147	212	1.60	-1.28	P0AB71	Fructose-bisphosphate aldolase class 2	fbaA
1	12	23.586	167	1.52	-1.26	P69441;Q1PI83	Adenylate kinase	adk
1	11	42.849	43	1.86	-1.24	P00946	Mannose-6-phosphate isomerase	manA
1	16	43.29	122	2.89	-1.20	P0A6A3	Acetate kinase	ackA
1	5	12.784	106	2.27	-1.20	P0AD49	Ribosome-associated inhibitor A	raiA
1	6	26.892	32	1.46	-1.07	P0AFP6	Putative GTP cyclohydrolase 1 type 2	ybgI
1	5	32.829	27	1.32	-1.06	P0AAB6	UTP--glucose-1-phosphate uridylyltransferase	galF
1	13	20.761	4390	1.55	-1.02	P0AE08	Alkyl hydroperoxide reductase subunit C	ahpC
1	9	30.636	32	1.35	-1.01	P18843	NH(3)-dependent NAD(+) synthetase	nadE
1	5	27.249	36	2.20	-0.99	P39831	NADP-dependent 3-hydroxy acid dehydrogenase YdfG	ydfG
1	8	23.905	31	1.61	-0.93	P38489;F2X7N5	Oxygen-insensitive NAD(P)H nitroreductase	nfsB;NTR
1	11	18.251	168	1.31	-0.89	P69783	Glucose-specific phosphotransferase enzyme IIA component	crr
1	17	45.316	303	1.68	-0.88	P0A825	Serine hydroxymethyltransferase	glyA

1	9	24.35	67	1.36	-0.80	P0AC59	Glutaredoxin-2	grxB
1	5	33.325	40	2.14	-0.79	P0AAX8	Probable L,D-transpeptidase YbiS	ybiS
1	9	22.533	94	1.53	-0.67	P0A8F0	Uracil phosphoribosyltransferase	upp
1	46	69.114	1284	1.74	-0.63	P0A6Y8	Chaperone protein DnaK	dnaK
1	5	38.19	30	2.43	-0.60	P0A9C3	Aldose 1-epimerase	galM
1	3	8.9863	47	1.42	-0.58	P0A7T7	30S ribosomal protein S18	rpsR
1	29	96.031	171	1.96	-0.43	P00957	Alanine--tRNA ligase	alaS
2	28	66.095	403	1.60	0.88	P06959	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	aceF
2	6	14.98	106	2.88	1.02	P02413	50S ribosomal protein L15	rplO
2	7	36.831	32	1.33	1.05	P08312	Phenylalanine--tRNA ligase alpha subunit	pheS
2	6	48.369	15	1.99	1.10	P0AFK0	Metalloprotease PmbA	pmbA
2	38	99.667	376	1.42	1.19	P0AFG8	Pyruvate dehydrogenase E1 component	aceE
2	7	77.515	10	1.33	1.23	P24171	Peptidyl-dipeptidase dcp	dcp
2	2	23.713	24	1.52	1.28	P0ACA7	Glutathione S-transferase GstB	gstB
2	18	50.688	249	1.77	1.30	P0A9P0	Dihydrolipoyl dehydrogenase	lpdA
2	3	55.035	6	1.45	1.36	P0A6P5	GTPase Der	der
2	10	52.015	19	1.77	1.38	P09832;P37127	Glutamate synthase [NADPH] small chain	gltD
2	5	56.415	9	1.89	1.54	P0AAG8	Galactose/methyl galactoside import ATP-binding protein MglA	mglA
2	2	54.845	5	1.51	1.72	P0C0L7	Proline/betaine transporter	proP
2	1	36.038	5	1.39	1.74	P21645	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase	lpxD
2	14	45.756	111	2.90	1.79	P08200;Q1PI72	Isocitrate dehydrogenase [NADP]	icd
2	4	127.21	4	1.74	1.84	P77338	Mechanosensitive channel MscK	mscK
2	2	63.997	5	2.30	1.88	P17846	Sulfite reductase [NADPH] hemoprotein beta-component	cysI
2	6	45.774	24	1.43	1.93	P22256	4-aminobutyrate aminotransferase GabT	gabT
2	2	34.593	7	1.41	1.93	P36771	Probable HTH-type transcriptional regulator LrhA	lrhA
2	5	27.751	27	1.98	1.96	P28861	Ferredoxin--NADP reductase	fpr
2	14	68.235	43	1.49	2.09	P33599	NADH-quinone oxidoreductase subunit C/D	nuoC
2	5	41.137	9	1.41	2.12	P0AB24	Iron uptake system component EfeO	efeO
2	8	44.011	62	2.27	2.15	P0AFG6	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	sucB
2	19	35.712	91	1.74	2.18	P0AEE5	D-galactose-binding periplasmic protein	mglB
2	3	48.906	6	1.45	2.19	P16095	L-serine dehydratase 1	sdaA
2	21	163.3	53	1.40	2.22	P09831	Glutamate synthase [NADPH] large chain	gltB
2	15	56.75	52	2.01	2.29	P13035	Aerobic glycerol-3-phosphate dehydrogenase	glpD
2	28	56.23	284	2.00	2.32	P0A6F3	Glycerol kinase	glpK

2	3	28.145	6	1.65	2.33	P0A9L8	Pyrroline-5-carboxylate reductase	proC
2	12	60.105	32	2.01	2.41	P14407	Fumarate hydratase class I, anaerobic	fumB
2	13	82.416	44	1.44	2.56	P76558	NADP-dependent malic enzyme	maeB
2	3	20.538	10	1.75	2.56	P0AFD6	NADH-quinone oxidoreductase subunit I	nuoI
2	5	26.103	17	1.91	2.62	P0AE22	Class B acid phosphatase	aphA
2	6	50.657	12	1.97	2.73	P06720	Alpha-galactosidase	melA
2	10	27.19	65	1.51	2.87	P0AEQ3	Glutamine-binding periplasmic protein	glnH
2	10	85.774	28	1.48	2.87	P00452;P00452-2	Ribonucleoside-diphosphate reductase 1 subunit alpha	nrdA
2	9	29.777	44	1.80	2.90	P0AGE9	Succinyl-CoA ligase [ADP-forming] subunit alpha	sucD
2	12	80.488	29	1.48	3.02	P37330	Malate synthase G	glcB
2	7	55.041	26	3.09	3.13	P04983	Ribose import ATP-binding protein RbsA	rbsA
2	11	41.392	82	1.83	3.18	P0A836	Succinyl-CoA ligase [ADP-forming] subunit beta	sucC
2	8	58.066	21	2.18	3.29	P11349;P19318	Respiratory nitrate reductase 1 beta chain	narH
2	5	26.449	19	2.58	3.37	P0AF26	Nitrate reductase molybdenum cofactor assembly chaperone NarJ	narJ
2	3	26.77	14	2.13	3.43	P07014	Succinate dehydrogenase iron-sulfur subunit	sdhB
2	24	93.497	99	3.45	3.43	P36683	Aconitate hydratase B	acnB
2	9	64.421	38	2.14	3.92	P0AC41	Succinate dehydrogenase flavoprotein subunit	sdhA
2	18	60.298	52	4.42	4.02	P0AC33	Fumarate hydratase class I, aerobic	fumA
2	14	52.272	63	1.58	4.05	P25553	Lactaldehyde dehydrogenase	aldA
2	9	30.427	38	1.77	4.06	P0ABQ2	2-hydroxy-3-oxopropionate reductase	garR
2	8	48.014	25	2.23	4.46	P0ABH7	Citrate synthase	gltA
2	20	105.06	82	1.77	4.53	P0AFG3	2-oxoglutarate dehydrogenase E1 component	sucA
2	8	38.612	35	3.24	4.84	P76116	Uncharacterized protein YncE	yncE
2	8	23.097	20	4.37	5.12	P00448	Superoxide dismutase [Mn]	sodA
2	5	15.463	26	1.80	5.49	P0A763	Nucleoside diphosphate kinase	ndk
2	4	8.6389	29	6.02	6.57	P0AA31	UPF0033 protein YedF	yedF

Hierarchical clustering analysis for the regulated proteins in the two experimental groups showed two distinct clusters (Figure 3.2S). One cluster is for the upregulated protein in the aerobic experimental group while the other is for the upregulated proteins in the anaerobic experimental group. The hierarchical cluster also shows that the four

biological replicates in every group cluster one group which is expected. The log2 iBAQ values were Z- score normalized and mapped to the row-scaled expression values (+1.5 to -1.5): the red values are for upregulated proteins, and the blue values are for down-regulated proteins.

Differential Alkylation Analysis

To measure changes in the global redox state of *E. coli*, we used a novel reagent to specifically label cysteine side chains [183]-[169]. This protein tag consists of a highly soluble (Z)-CyDye coupled to a thio-reactive maleimide group. (Z)-CyDyes were specifically developed to be compatible with 2D gels. Cell pellet were collected at the same time points for both aerobic and anaerobic conditions. The soluble protein fractions were let to react with the (Z)-CyDyes as described by Maaty *et al.*, (2012) [169] After quenching the reaction, only 50 ug of each sample was loaded onto 2D gel apparatus as described in the method section. In order to visualize the reduced thiols, which had fluorescent probes attached, the 2D gels were scanned using the Typhoon (gel scanner model, etc) utilizing the fluorescent excitiation and the (blah Blah) filters. The variation of proteins and color saturation on the aerobic gel versus the anaerobic gel show the larger number of proteins and higher total fluorescence intensity in the aerobic samples. This points to a greater number of oxidized proteins in the anaerobic sample in comparison to the aerobic gel samples. These data suggest a high cellular reduced environment in the aerobic samples at the level of proteins which was never reported in the literature (Figure 3.3A). Moreover, to compare the overall number of the reduced to oxidized cysteine residues across the two experimental groups, the ratio of the total

number of reduced peptides to the ratio of the total number of oxidized peptides is calculated for every biological replicate as shown in Figure 3.3B. The ratio of the reduced to an oxidized number of peptides is 1.27±0.05 and 0.50±0.10 for the anaerobic and aerobic samples, respectively. The ratios show that more of the peptides are oxidized in the anaerobic condition in comparison to the aerobic condition

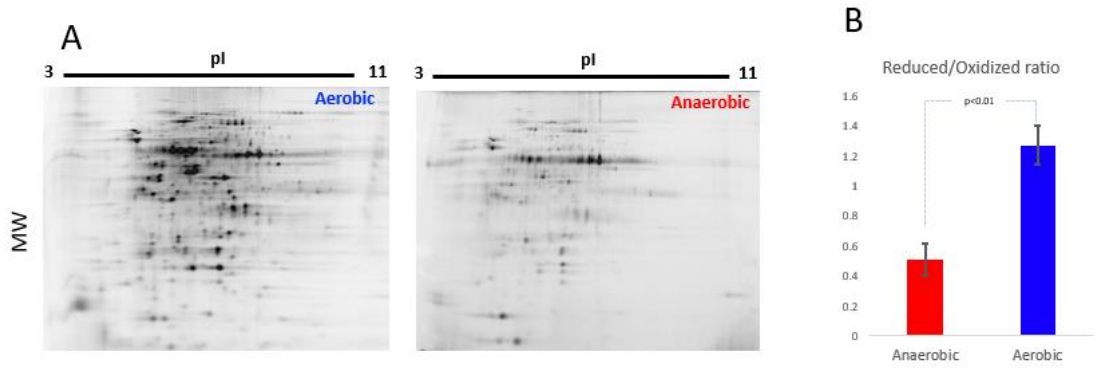


Figure 3.3 **Protein Differential alkylation analysis**. A, fluorescently labeled protein using maleimide Z-CyDye5 2D gels for the experimental groups, aerobic and anaerobic. B, the ratio of total number of the reduced peptides to the total number of the oxidized peptides in the two experimental groups.

It is essential to identify the cysteine-containing proteins involved in cell signaling in both conditions in order to link their potential role to the metabolic state. This could reveal an oxidation state that is linked to a signaling protein in the two experimental groups. Therefore, the next part of the study is focused on the ratio of oxidized to reduced cysteines by focusing on cysteine-containing peptides. We identified 214 Cys-containing peptides that were shared between experimental groups (Table 3.2S). Using the abundance of reduced and oxidized forms of the peptides, hierarchical

clustering showed that the two experimental groups sorted as expected into four classes. There were aerobic reduced (Ae-SH), aerobic oxidized (Ae-SS), anaerobic reduced (An-SH) and anaerobic oxidized (An-SS) classes, Figure 3.3A. The heatmap shows that the An-SS peptides were globally the most abundant class. This is consistent with our 2D gel analysis. Within this general description were subclasses that behaved differently. For example, the cluster with Glucose-1-phosphate (399) and (391), Cytosol aminopeptidase (360), 2-Dehydro-3-deoxygluconokinase (234) and Acid stress chaperone HdeB (39) show higher abundance of oxidized peptide classes in both aerobic and anaerobic conditions, meaning that the cysteines in these peptides are predominantly oxidized independent of metabolic state. In contrast, the cluster, which contains peptides Adenylosuccinate synthase (104), Tryptophanase (148) and 60 kDa chaperonin (519), predominantly had reduced peptide classes in both aerobic and anaerobic experimental groups. It is noteworthy mentioning that these data suggest that peptides reduction potential is specific, and does not always follow the global protein expression pattern.

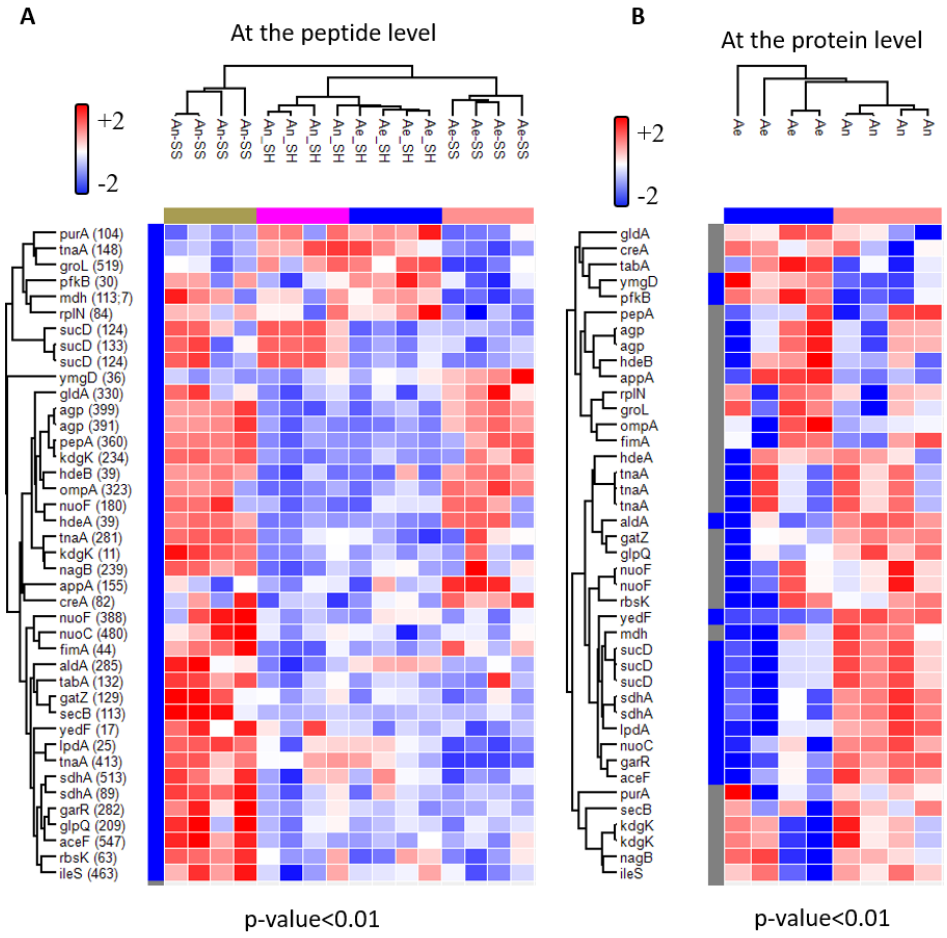


Figure 3.4: Hierarchical cluster analysis for the differentially regulated Cysteine-containing peptide intensities (A) and with their corresponding total protein expression values (B). The heat map colored scale is Z-score of the highest and lowest intensities and represented as (-2 to +2) where Blue is low and Red is High abundance. The blue vertical line next to the gene names represent a significant ANOVA comparison (p-value < 0.01). The proteins with gray lines are the ones that did not show any differential expression at the level of proteins. An-SS and An-SH are the log2 transformed values for the oxidized and reduced peptides in the Anaerobic experimental group, respectively. Ae-SS and Ae-SH are the log2 transformed values for the oxidized and reduced peptides in the Aerobic experimental group, respectively. All of these peptides are ANOVA significant at p<0.01. The numbers in the parentheses are for the position of the cysteine within the protein. Aerobic (An) and Anaerobic (Ae) are for the total protein expression values.

An example of redox tuning of cysteines is the protein tnaA. behavior at the protein level and at the cysteine-containing peptides level. At the protein level, no significant changes were associated with the global expression level. Meanwhile, at the level of peptides, we were able to identify the two oxidative states for the Cys281, Cys148 and Cys413. The three peptides show statistically significant (ANOVA). Cys281 and Cys413 are significantly higher as oxidized in the anaerobic experimental group. This behavior would be expected for a molecule involved in cell signaling.

Discussion and Conclusion

The goal of the study is to investigate the role of protein-containing reactive cysteine groups. We used *E. coli* as an experimental model because it can grow in the presence and absence of oxygen. LC-MS/MS shotgun proteomics and thiolomics were performed to cells that were collected at the logarithmic (exponential) phase. We also performed differential alkylation reactions to capture two oxidative cysteine states. The enrichment analysis we performed highlighted the regulated metabolic pathways in both treatments. Finally, we performed differential analysis for the level of Cys-containing peptides that were identified in both oxidative states.

Global Proteome Profiling

Numerous studies have reported global changes associated with the aerobic and anaerobic metabolism, in particular the transition from one state to another at the level of genomics and transcripts in *E. coli* [155,176,179,180,184]. To our knowledge, this is the first

report to discuss cysteine oxidative states of proteins and differential abundance in the aerobic and anaerobic conditions.

The PCA analysis for the four replicates of the two experimental groups shows clear separation across the two experimental groups as expected. However, the clustering per group is different as the anaerobic cluster in a small area in comparison to the aerobic which cluster in a bigger area (Figure 3.1A). To investigate this behavior, we constructed a coefficient of variation (CV) histogram that visualizes the identified proteins' CV values (Figure 3.1D). The CV histogram shows two different trends. These two trends indicate similarities across the anaerobic experimental group, whereas we see less similarities across the aerobic experimental groups. This anaerobic behavior indicates that growth in an anaerobic environment is stressful in comparison to the aerobic environment. Moreover, 64% of the overall identified proteins has less than 5% CV, which indicates that the organism has less space of freedom to be in. Biologically, a stressed system focuses on responding to the stress rather than performing other biological activities as suggested by Tokmina-Lukaszewska *et al*., (2014) [185] and Heinemann *et al*., (2014) [186] (Figure 1S).

Our differential analysis of the proteome showed that 14% (138 out of 949) of the identified proteins changed significantly in abundance (Table 3.1S & Figure 3.1B). Enzymes involved with central carbon metabolism were over-represented in this group and not the presence of newly synthesized proteins as reported by Kramer, *et al*., (2010) [177]. In addition, a small number of proteins present only in each group, 30 and 16 proteins in the aerobic and anaerobic, respectively. Among these proteins, four proteins associated with response to oxidative stress, response to DNA damage, and redox homeostasis,

which only present in the anaerobic experimental group (Figure1C). These proteins are glutathione S-transferase, Iron uptake system component EfeO, Superoxide dismutase [Mn] and UPF0033 Yedf (Table 3.1S). Why would superoxide dismutase [Mn] only present in an anaerobic experimental group? Kargaliogly & Imalay (1994)[187] found that superoxide dismutase [Mn] is essential for the anaerobic growth to facilitates the entry to an aerobic environment [187]. They also found that the SOD-deficient strain lagged three hours upon aeration. Their data suggest the importance of preemptively synthesize SOD during the anaerobic metabolism in preparation for the transition into the aerobic condition.

The 30 proteins that were present only in the aerobic group consists of nine proteins that are associated with protein maturation and protein synthesis quality control (Table 3.1S). These proteins are 50S ribosomal protein L33, hydrogenase isoenzymes nickel incorporation protein HypB, hydrogenase-1 maturation protease, hydrogenase-1operon proteins HyaE, hydrogenase 2 maturation protease, met repressor, histidinol dehydrogenase, peptidyl-t-RNA hydrolase, and isoasparyl dipeptidase. As expected, and shown by the growth curve, the aerobic experimental group bacteria grow better in the presence of oxygen as reported by Macdaniel, *et al.*, (1965) [188]. Biologically, the protein synthesis machinery is active, producing more proteins to facilitate the rapid growth rate. It is also noteworthy to mention that one protein is also identified in the aerobic experimental group that is associated to the with the Methylglyoxal breakdown which is 2.5-diketo-D-gluconic acid reductase B. The Methylglyoxal pathway is a diversion of the Glycolysis pathway that is found in few prokaryotes which convert glucose to lactate and eventually to pyruvate [182]. This diversion does not result in energy production, as the

pathway doesn't yield ATP, but rather it is to lower the concentrations of sugar phosphates and replenish stores of inorganic phosphate needed for ATP generation [182]. This evidence suggests a better energetic state for the aerobic state than in the anaerobic.

Enrichment analysis for biological terms using the 14% of the total number of proteins and the total number of identified proteins as background shows that Glycolysis / Gluconeogenesis (14 protein), Fructose and Mannose metabolism (8 proteins) and Monosaccharides catabolic process (14 protein) in the aerobic group (Table 1). It has been reported by the transcriptomics aerobic and anaerobic transition studies that the main altered metabolic pathways are the central carbon metabolism [155,176]. However, the enrichment analysis also shows different central carbon metabolism regulated pathways; like Citrate cycle (TCA) (14 proteins), Acetyl-CoA metabolic pathway (13 proteins) and Glyoxylate metabolic process (4 proteins) (Table 3.1S). The enrichment of the TCA cycle does not necessarily mean that these enzymes are present only in this experimental group but rather means that they are present in both experimental groups but higher abundance in the second group. The increased abundance in the TCA cycle proteins is also associated with two of the amino acid degradation enzymes: Glutamate synthase [NADPH] large chain and Glutamate synthase [NADPH] small chain, at a fold change of 2.2- and 1.3-, respectively (Table 2). The glutamate synthase [NADPH] reaction is reversible, and one of the products is 2-oxoglutarate (alpha-ketoglutarate) which is one of the important TCA cycle precursors. This suggests there is a flux of amino acids into the TCA cycle as a source of energy in the anaerobic metabolism. At the same time, Glyoxylate pathway, which is a TCA bypass, is also enriched in the anaerobic group. It has been found that the Glyoxylate pathway is activated when acetate accumulates and is

used as a carbon source [189–192]. In anaerobic metabolism, acetate is one of the products generated by the mixed acid fermentation reactions [193]. This potentially explains why this pathway is only activated in the anaerobic experimental group as shown in this data.

Oxidized Anaerobic Environment

The collected 2D fluorescence images (Figure 3.2A) show a high degree of oxidation at the level of proteins in the anaerobic proteins sample. The proteomics shotgun data analysis shows the same trend of oxidation to reduction at the level of peptides (Figure 3.2B). This could be because the aerobic metabolism is more energetic and active, whereas the anaerobic metabolism is less active. Also, this could be because anaerobic metabolism is experiencing some metabolic stress that is also suggested by the CV histogram (Figure 3.1C).

Redox analysis at the Level of Peptides and Potential Redox Signaling Molecule

The hierarchal clustering analysis shows more than one behavior at the level of Cys- containing peptides. The first cluster that contains Adenylosuccinate synthase (104), Tryptophanase (148) and Acid stress chaperone HdeB (39) shows a trend where these Cys containing peptides shows a high degree of reduction and low degree of oxidation in both samples (Figure 3.3A). This indicates that these proteins may not involve in any oxidation or reduction reactions in the two experimental conditions. In fact, 60 kDa chaperonin (519) peptide is an excellent example of this cluster. While this peptide is increased in abundance, the overall expression value does not show significant change (Figure 3B). This could be due to the function of this 60 kDa chaperonin not requiring the

involvement of Cys159 in the protein folding in response to heat. A second behavior of the cluster containing Glucose-1-phosphate (399), (391), Cytosol aminopeptidase (360), 2-Dehydro-3-deoxygluconokinase (234) and Acid stress chaperone HdeB (39) show high abundance in the two reduced states in the aerobic and anaerobic (Figure 3.3A) where none the overall protein expression values shows any significant difference. This could be because these proteins present in the two experimental groups but at the peptide level, favors the oxidized state. Moreover, these Cys residues might be involved in a structure-based disulfide bridge and not involved in any oxidation-reduction reactions.

One of the interesting behaviors we observed was the case of Tryptophanase Cys-oxidative peptides regarding the overall protein expression value (Figure 3.4A, 3.4B). The analysis at the peptide level shows different trends of the three identified peptides at the two oxidative states. It is worth mentioning that the abundance of Tryptophanase does not show a significant difference (Figure 3.3B). The investigation of the biological role of Tryptophanase Cys residues shows that all Cys residues are reduced and do not involve in any disulfide bonds [194]. Moreover, Tryptophanase is an enzyme involved in the degradation of Tryptophan in *E. coli*, to generate ammonium, pyruvate, and indole [195]. Nihra, *et al*., (1985) [196] suggested a functional role of the cysteine residues in Tryprophanase [196]. It has been shown that blocking the Cys residues irreversibly results in lowering the affinity of the enzyme to its cofactor, pyridoxal 5-phosphate [196]. In fact, proteomics studied indicated that nitrosylation is reported in tnaA by Liu, *et al*., (2010)[197] and Forrester, *et al*., (2009)[198]. Such behavior is expected from a protein involved in thiol signaling. Furthermore, Indole production is associated with many physiological

activities. These activities range from biofilm formation [199], motility [200], virulence [201]and antibiotic resistance [202].

In conclusion, we were able to help untangle some of the unknown threads connecting the shift from aerobic to anaerobic metabolism in *E. coli*. First, we were able to report uniquely present proteins in the two experimental groups that are essential for the metabolic states. We also reported the enrichment of central carbon metabolic pathways and the Methylglyoxal pathway in the aerobic group. In contrast, Glyoxylate pathway and TCA cycle pathway were enriched in the anaerobic group. Moreover, by utilizing a MS based proteomics approach, with a pipeline suited to analyzing PTM's, the differential alkylation of cysteines in anaerobic versus aerobic conditions revealed higher high degree of protein oxidation about this niche of *E. coli* metabolism. There are varying behaviors of reduction/oxidation states of individual Cys containing peptides. One of such is the group of proteins whose their Cys containing peptides showed differential abundance at the level peptides in the two environments. We were able to potentially identify Tryptophanse as a thiol switch. Trptophanse is an idol producing enzyme, which is a critical signaling molecule in many pathways in bacterial metabolism. Not only were we able to identify new potential switches from our refined pipeline, our data suggests that there is a large number of oxidized proteins in the anaerobic culture as compared to the aerobic counterpart, which has never been reported in the literature.

REFERENCE CITED

1. Blattner FR. The Complete Genome Sequence of Escherichia coli K-12. *Science (80-)*. 1997;277(5331):1453-1462. doi:10.1126/science.277.5331.1453

2. Feist AM, Palsson BØ. The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol*. 2008;26(6):659-667. doi:10.1038/nbt1401

3. Barrick JE, Yu DS, Yoon SH, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*. 2009;461(7268):1243-1247. doi:10.1038/nature08480

4. Richmond C, Glasner JD, Mau R, Jin H, Blattner FR. Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Res*. 1999;27(19):3821-3835. doi:10.1093/nar/27.19.3821

5. Chang D-E, Smalley DJ, Tucker DL, et al. Noninvasive measurement of anatomic structure and intraluminal oxygenation in the gastrointestinal tract of living mice with spatial and spectral EPR imaging. *PNAS*. 2004;96(8):4586-4591. doi:10.1073/pnas.96.8.4586

6. Kolisnychenko V, Plunkett G, Herring CD, et al. Engineering a reduced Escherichia coli genome. *Genome Res*. 2002;12(4):640-647. doi:10.1101/gr.217202

7. Chen S, Lesnik EA, Hall TA, et al. A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. *Biosystems*. 2002;65(2-3):157-177. doi:10.1016/S0303-2647(02)00013-8

8. Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol*. 2015;23:133-140. doi:10.1016/J.MIB.2014.11.011

9. Hain T, Hossain H, Chatterjee SS, et al. Temporal transcriptomic analysis of the Listeria monocytogenes EGD-e σB regulon. *BMC Microbiol*. 2008;8(1):20. doi:10.1186/1471-2180-8-20

10. Chueca B, Pagán R, García-Gonzalo D. Transcriptomic analysis of Escherichia coli MG1655 cells exposed to pulsed electric fields. *Innov Food Sci Emerg Technol*. 2015;29:78-86. doi:10.1016/J.IFSET.2014.09.003

11. Nie L, Wu G, Culley DE, Scholten JCM, Zhang W. Integrative Analysis of Transcriptomic and Proteomic Data: Challenges, Solutions and Applications. *Crit Rev Biotechnol*. 2007;27(2):63-75. doi:10.1080/07388550701334212

12. Weber A, Kögl SA, Jung K. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli. *J Bacteriol*.

2006;188(20):7165-7175. doi:10.1128/JB.00508-06

13. Vertommen D, Ruiz N, Leverrier P, Silhavy TJ, Collet J-F. Characterization of the role of the Escherichia coli periplasmic chaperone SurA using differential proteomics. *Proteomics*. 2009;9(9):2432-2443. doi:10.1002/pmic.200800794

14. Schmidt A, Kochanowski K, Vedelaar S, et al. The quantitative and condition-dependent Escherichia coli proteome. *Nat Biotechnol*. 2016;34(1):104-110. doi:10.1038/nbt.3418

15. Lai E-M, Nair U, Phadke ND, Maddock JR. Proteomic screening and identification of differentially distributed membrane proteins in Escherichia coli. *Mol Microbiol*. 2004;52(4):1029-1044. doi:10.1111/j.1365-2958.2004.04040.x

16. Rabinowitz JD, Kimball E. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Anal Chem*. 2007;79(16):6167-6173. doi:10.1021/ac070470c

17. Nobeli I, Ponstingl H, Krissinel EB, Thornton JM. A Structure-based Anatomy of the E. coli Metabolome. *J Mol Biol*. 2003;334(4):697-719. doi:10.1016/J.JMB.2003.10.008

18. Heinemann J, Noon B, Mohigmi MJ, Mazurie A, Dickensheets DL, Bothner B. Real-Time Digitization of Metabolomics Patterns from a Living System Using Mass Spectrometry. *J Am Soc Mass Spectrom*. 2014;25(10):1755-1762. doi:10.1007/s13361-014-0922-z

19. Yuan J, Doucette CD, Fowler WU, et al. Metabolomics-driven quantitative analysis of ammonia assimilation in E. coli. *Mol Syst Biol*. 2009;5(1):302. doi:10.1038/msb.2009.60

20. Tweeddale H, Notley-McRobb L, Ferenci T. Effect of Slow Growth on Metabolism of Escherichia coli, as Revealed by Global Metabolite Pool ("Metabolome") Analysis. *J Bacteriol*. 1998;180(19):5109-5116. https://jb.asm.org/content/180/19/5109.short. Accessed April 16, 2019.

21. Partridge JD, Scott C, Tang Y, Poole RK, Green J. Escherichia coli transcriptome dynamics during the transition from anaerobic to aerobic conditions. *J Biol Chem*. 2006;281(38):27806-27815. doi:10.1074/jbc.M603450200

22. Guest JR, Green, J. Irvine AS, Spiro S. *Regulation of Gene Expression in Escherichia Coli*. Boston, MA: Springer US; 1996. doi:10.1007/978-1-4684-8601-8

23. Gennis RB S V. *Respiration, Escherichia Coli and Salmonella Cellular and Molecular Biology*. 2nd ed. (Neidhardt FC, ed.). Washington, D.C: ASM Press; 1996.

24. Unden G, Bongaerts J. Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. *Biochim Biophys Acta*. 1997;1320(3):217-234. http://www.ncbi.nlm.nih.gov/pubmed/9230919. Accessed May 19, 2019.

25. Pecher A, Blaschkowski HP, Knappe K, Böck A. Expression of pyruvate formate-lyase of <i>Escherichia coli</i> from the cloned structural gene. *Arch Microbiol*. 1982;132(4):365-371. doi:10.1007/BF00413390

26. Dawes EA, Foster SM. The formation of ethanol in Escherichia coli. *Biochim Biophys Acta*. 1956;22(2):253-265. http://www.ncbi.nlm.nih.gov/pubmed/13382840. Accessed November 9, 2016.

27. Gupta S, Clark DP. Escherichia coli derivatives lacking both alcohol dehydrogenase and phosphotransacetylase grow anaerobically by lactate fermentation. *J Bacteriol*. 1989;171(7):3650-3655. http://www.ncbi.nlm.nih.gov/pubmed/2661531. Accessed November 9, 2016.

28. Stryer L. *Biochemistry*. 4th ed. (W. H. Freeman and Company., ed.). New York-Basingstoke; 1995.

29. Murashko ON, Lin-Chao S. Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. doi:10.1073/pnas.1703731114

30. Jones SA, Chowdhury FZ, Fabich AJ, et al. Respiration of Escherichia coli in the mouse intestine. *Infect Immun*. 2007;75(10):4891-4899. doi:10.1128/IAI.00484-07

31. Marteyn B, West NP, Browning DF, et al. Modulation of Shigella virulence in response to available oxygen in vivo. *Nature*. 2010;465(7296):355-358. doi:10.1038/nature08970

32. Davies K. Oxidative Stress, Antioxidant Defenses, and Damage Removal, Repair, and Replacement Systems. *IUBMB Life*. 2001;50(4):279-289. doi:10.1080/713803728

33. Davies KJ. Oxidative stress: the paradox of aerobic life. *Biochem Soc Symp*. 1995;61:1-31. doi:10.1042/BSS0610001

34. Riley PA. Free radicals in biology: oxidative stress and the effects of ionizing radiation. *Int J Radiat Biol*. 1994;65(1):27-33. http://www.ncbi.nlm.nih.gov/pubmed/7905906. Accessed May 8, 2019.

35. Sahnoun Z, Jamoussi K, Zeghal KM. [Free radicals and antioxidants: human physiology, pathology and therapeutic aspects]. *Therapie*. 1997;52(4):251-270. http://www.ncbi.nlm.nih.gov/pubmed/9437876. Accessed May 8, 2019.

36. Sakac V, Sakac M. [Free oxygen radiacals and kidney diseases--part I]. *Med Pregl*. 2000;53(9-10):463-74. http://www.ncbi.nlm.nih.gov/pubmed/11320727. Accessed May 8, 2019.

37. Devasagayam TPA, Tilak JC, Boloor KK, Sane KS, Ghaskadbi SS, Lele RD. Free radicals and antioxidants in human health: current status and future prospects. *J Assoc Physicians India*. 2004;52:794-804. http://www.ncbi.nlm.nih.gov/pubmed/15909857. Accessed May 8, 2019.

38. Finkel T. Signal transduction by reactive oxygen species. *J Cell Biol*. 2011;194(1):7-15. doi:10.1083/jcb.201102095

39. Turkan I. ROS and RNS: key signalling molecules in plants. *J Exp Bot*. 2018;69(14):3313-3315. doi:10.1093/jxb/ery198

40. Bigarella CL, Liang R, Ghaffari S. Stem cells and the impact of ROS signaling. *Development*. 2014;141(22):4206-4218. doi:10.1242/dev.107086

41. Zhang J, Wang X, Vikash V, et al. ROS and ROS-Mediated Cellular Signaling. *Oxid Med Cell Longev*. 2016;2016:1-18. doi:10.1155/2016/4350965

42. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol*. 2014;24(10):R453-62. doi:10.1016/j.cub.2014.03.034

43. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8(10):813-824. doi:10.1038/nrm2256

44. Knoefler D, Leichert LIO, Thamsen M, et al. About the dangers, costs and benefits of living an aerobic lifestyle. *Biochem Soc Trans*. 2014;42(4):917-921. doi:10.1042/BST20140108

45. Halliwell, B.; Gutteridge J, ed. Free radicals in biology and medicine. In: 4th ed. Oxford: Oxford University Press; 2007.

46. Brandes N, Schmitt S, Jakob U. Thiol-based redox switches in eukaryotic proteins. *Antioxid Redox Signal*. 2009;11(5):997-1014. doi:10.1089/ars.2008.2285

47. Klomsiri C, Karplus PA, Poole LB. Cysteine-based redox switches in enzymes. *Antioxid Redox Signal*. 2011;14(6):1065-1077. doi:10.1089/ars.2010.3376

48. Paulsen CE, Carroll KS. Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chem Biol*. 2010;5(1):47-62. doi:10.1021/cb900258z

49. Antelmann H, Helmann JD. Thiol-based redox switches and gene regulation. *Antioxid Redox Signal*. 2011;14(6):1049-1063. doi:10.1089/ars.2010.3400

50. Jang HH, Lee KO, Chi YH, et al. Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell*. 2004;117(5):625-635. doi:10.1016/j.cell.2004.05.002

51. Winter J, Ilbert M, Graf PCF, Ozcelik D, Jakob U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell*. 2008;135(4):691-701. doi:10.1016/j.cell.2008.09.024

52. Dalle-Donne I, Rossi R, Colombo G, Giustarini D, Milzani A. Protein S-glutathionylation: a regulatory device from bacteria to humans. *Trends Biochem Sci*. 2009;34(2):85-96. doi:10.1016/j.tibs.2008.11.002

53. Burgoyne JR, Eaton P. A rapid approach for the detection, quantification, and discovery of novel sulfenic acid or S-nitrosothiol modified proteins using a biotin-switch method. *Methods Enzymol*. 2010;473:281-303. doi:10.1016/S0076-6879(10)73015-9

54. Yeh JI, Claiborne A, Hol WG. Structure of the native cysteine-sulfenic acid redox center of enterococcal NADH peroxidase refined at 2.8 A resolution. *Biochemistry*. 1996;35(31):9951-9957. doi:10.1021/bi961037s

55. Foster MW, Hess DT, Stamler JS. Protein S-nitrosylation in health and disease: a current perspective. *Trends Mol Med*. 2009;15(9):391-404. doi:10.1016/j.molmed.2009.06.007

56. Kokkola T, Savinainen JR, Mönkkönen KS, Retamal MD, Laitinen JT. S-nitrosothiols modulate G protein-coupled receptor signaling in a reversible and highly receptor-specific manner. *BMC Cell Biol*. 2005;6(1):21. doi:10.1186/1471-2121-6-21

57. Nozik-Grayck E, Whalen EJ, Stamler JS, McMahon TJ, Chitano P, Piantadosi CA. S-nitrosoglutathione inhibits alpha1-adrenergic receptor-mediated vasoconstriction and ligand binding in pulmonary artery. *Am J Physiol Lung Cell Mol Physiol*. 2006;290(1):L136-43. doi:10.1152/ajplung.00230.2005

58. Whalen EJ, Foster MW, Matsumoto A, et al. Regulation of beta-adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2. *Cell*. 2007;129(3):511-522. doi:10.1016/j.cell.2007.02.046

59. Benhar M, Forrester MT, Hess DT, Stamler JS. Regulated protein denitrosylation by cytosolic and mitochondrial thioredoxins. *Science*. 2008;320(5879):1050-1054. doi:10.1126/science.1158265

60. Chanvorachote P, Nimmannit U, Wang L, et al. Nitric oxide negatively regulates Fas CD95-induced apoptosis through inhibition of ubiquitin-proteasome-mediated degradation of FLICE inhibitory protein. *J Biol Chem*. 2005;280(51):42044-42050. doi:10.1074/jbc.M510080200

61. Hoffmann J, Haendeler J, Zeiher AM, Dimmeler S. TNFalpha and oxLDL reduce protein S-nitrosylation in endothelial cells. *J Biol Chem*. 2001;276(44):41383-41387. doi:10.1074/jbc.M107566200

62. Kim J-E, Tannenbaum SR. S-Nitrosation regulates the activation of endogenous procaspase-9 in HT-29 human colon carcinoma cells. *J Biol Chem*. 2004;279(11):9758-9764. doi:10.1074/jbc.M312722200

63. Kim YM, Talanian R V, Billiar TR. Nitric oxide inhibits apoptosis by preventing increases in caspase-3-like activity via two distinct mechanisms. *J Biol Chem*. 1997;272(49):31138-31148. http://www.ncbi.nlm.nih.gov/pubmed/9388267. Accessed March 30, 2016.

64. Mannick JB, Hausladen A, Liu L, et al. Fas-induced caspase denitrosylation. *Science*. 1999;284(5414):651-654. http://www.ncbi.nlm.nih.gov/pubmed/10213689. Accessed March 30, 2016.

65. Huang Y, Man H-Y, Sekine-Aizawa Y, et al. S-nitrosylation of N-ethylmaleimide sensitive factor mediates surface expression of AMPA receptors. *Neuron*. 2005;46(4):533-540. doi:10.1016/j.neuron.2005.03.028

66. Lipton SA, Choi YB, Pan ZH, et al. A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. *Nature*. 1993;364(6438):626-632. doi:10.1038/364626a0

67. Mustafa AK, Kumar M, Selvakumar B, et al. Nitric oxide S-nitrosylates serine racemase, mediating feedback inhibition of D-serine formation. *Proc Natl Acad Sci U S A*. 2007;104(8):2950-2955. doi:10.1073/pnas.0611620104

68. Takahashi H, Shin Y, Cho S-J, et al. Hypoxia enhances S-nitrosylation-mediated NMDA receptor inhibition via a thiol oxygen sensor motif. *Neuron*. 2007;53(1):53-64. doi:10.1016/j.neuron.2006.11.023

69. Lowenstein CJ. Nitric oxide regulation of protein trafficking in the cardiovascular system. *Cardiovasc Res*. 2007;75(2):240-246. doi:10.1016/j.cardiores.2007.03.024

70. Matsushita K, Morrell CN, Cambien B, et al. Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. *Cell*. 2003;115(2):139-150. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2846406&tool=pmcentrez&rendertype=abstract. Accessed March 30, 2016.

71. Morrell CN, Matsushita K, Chiles K, et al. Regulation of platelet granule exocytosis by S-nitrosylation. *Proc Natl Acad Sci U S A*. 2005;102(10):3782-3787. doi:10.1073/pnas.0408310102

72. Wang G, Moniri NH, Ozawa K, Stamler JS, Daaka Y. Nitric oxide regulates endocytosis by S-nitrosylation of dynamin. *Proc Natl Acad Sci U S A*.

2006;103(5):1295-1300. doi:10.1073/pnas.0508354103

73. Kim SF, Huri DA, Snyder SH. Inducible nitric oxide synthase binds, S-nitrosylates, and activates cyclooxygenase-2. *Science*. 2005;310(5756):1966-1970. doi:10.1126/science.1119407

74. Xu L, Han C, Lim K, Wu T. Activation of cytosolic phospholipase A2alpha through nitric oxide-induced S-nitrosylation. Involvement of inducible nitric-oxide synthase and cyclooxygenase-2. *J Biol Chem*. 2008;283(6):3077-3087. doi:10.1074/jbc.M705709200

75. Tian J, Kim SF, Hester L, Snyder SH. S-nitrosylation/activation of COX-2 mediates NMDA neurotoxicity. *Proc Natl Acad Sci U S A*. 2008;105(30):10537-10540. doi:10.1073/pnas.0804852105

76. Uehara T, Nakamura T, Yao D, et al. S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature*. 2006;441(7092):513-517. doi:10.1038/nature04782

77. Dalle-Donne I, Scaloni A, Giustarini D, et al. Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. *Mass Spectrom Rev*. 2005;24(1):55-99. doi:10.1002/mas.20006

78. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol*. 1990;58(3):157-161. doi:10.1007/BF00548927

79. Hillion M, Antelmann H. Thiol-based redox switches in prokaryotes. *Biol Chem*. 2015;396(5):415-444. doi:10.1515/hsz-2015-0102

80. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie Van Leeuwenhoek*. 1990;58(3):157-161. http://www.ncbi.nlm.nih.gov/pubmed/2256675. Accessed May 20, 2019.

81. Kim SO, Merchant K, Nudelman R, et al. OxyR: a molecular code for redox-related signaling. *Cell*. 2002;109(3):383-396. http://www.ncbi.nlm.nih.gov/pubmed/12015987. Accessed May 20, 2019.

82. Li Y, He Z-G. The Mycobacterial LysR-Type Regulator OxyS Responds to Oxidative Stress and Negatively Regulates Expression of the Catalase-Peroxidase Gene. Cornelis P, ed. *PLoS One*. 2012;7(1):e30186. doi:10.1371/journal.pone.0030186

83. Domenech P, Honoré N, Heym B, Cole ST. Role of OxyS of Mycobacterium tuberculosis in oxidative stress: overexpression confers increased sensitivity to organic hydroperoxides. *Microbes Infect*. 2001;3(9):713-721. http://www.ncbi.nlm.nih.gov/pubmed/11489419. Accessed April 28, 2019.

84. Zheng C, Xu J, Li J, et al. Two Spx Regulators Modulate Stress Tolerance and Virulence in Streptococcus suis Serotype 2. Lemos JA, ed. *PLoS One*. 2014;9(9):e108197. doi:10.1371/journal.pone.0108197

85. Choi H, Kim S, Mukhopadhyay P, et al. Structural basis of the redox switch in the OxyR transcription factor. *Cell*. 2001;105(1):103-113. http://www.ncbi.nlm.nih.gov/pubmed/11301006. Accessed May 20, 2019.

86. Hildebrandt T, Knuesting J, Berndt C, Morgan B, Scheibe R. Cytosolic thiol switches regulating basic cellular functions: GAPDH as an information hub? *Biol Chem*. 2015;396(5):523-537. doi:10.1515/hsz-2014-0295

87. Grant CM, Quinn KA, Dawes IW. Differential protein S-thiolation of glyceraldehyde-3-phosphate dehydrogenase isoenzymes influences sensitivity to oxidative stress. *Mol Cell Biol*. 1999;19(4):2650-2656. http://www.ncbi.nlm.nih.gov/pubmed/10082531. Accessed April 28, 2019.

88. Ralser M, Wamelink MM, Kowald A, et al. Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. *J Biol*. 2007;6(4):10. doi:10.1186/jbiol61

89. Sampathkumar Krishnan ‡, Eva Y. Chi §, Stephen J. Wood ‖, et al. Oxidative Dimer Formation Is the Critical Rate-Limiting Step for Parkinson's Disease α-Synuclein Fibrillogenesis†. 2002. doi:10.1021/BI026528T

90. Wojdyla K, Rogowska-Wrzesinska A. Differential alkylation-based redox proteomics - Lessons learnt. *Redox Biol*. 2015;6:240-252. doi:10.1016/j.redox.2015.08.005

91. Jaffrey SR, Snyder SH. The biotin switch method for the detection of S-nitrosylated proteins. *Sci STKE*. 2001;2001(86):pl1. doi:10.1126/stke.2001.86.pl1

92. Shelton MD, Chock PB, Mieyal JJ. Glutaredoxin: Role in Reversible Protein *S* -Glutathionylation and Regulation of Redox Signal Transduction and Protein Translocation. *Antioxid Redox Signal*. 2005;7(3-4):348-366. doi:10.1089/ars.2005.7.348

93. García-Santamarina S, Boronat S, Espadas G, Ayté J, Molina H, Hidalgo E. The oxidized thiol proteome in fission yeast—Optimization of an ICAT-based method to identify H2O2-oxidized proteins. *J Proteomics*. 2011;74(11):2476-2486. doi:10.1016/J.JPROT.2011.05.030

94. Wojdyla K, Williamson J, Roepstorff P, Rogowska-Wrzesinska A. The SNO/SOH TMT strategy for combinatorial analysis of reversible cysteine oxidations. *J Proteomics*. 2015;113:415-434. doi:10.1016/J.JPROT.2014.10.015

95. Qu Z, Meng F, Bomgarden RD, et al. Proteomic Quantification and Site-Mapping

of *S* -Nitrosylated Proteins Using Isobaric iodoTMT Reagents. *J Proteome Res*. 2014;13(7):3200-3211. doi:10.1021/pr401179v

96. Fares A, Nespoulous C, Rossignol M, Peltier J-B. Simultaneous Identification and Quantification of Nitrosylation Sites by Combination of Biotin Switch and ICAT Labeling. In: Humana Press, Totowa, NJ; 2014:609-620. doi:10.1007/978-1-62703-631-3_41

97. James P. Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys*. 1997;30(4):279-331. http://www.ncbi.nlm.nih.gov/pubmed/9634650. Accessed January 11, 2019.

98. Feng J, Zhang Q, Zhou Y, et al. Integration of Proteomics and Metabolomics Revealed Metabolite–Protein Networks in ACTH-Secreting Pituitary Adenoma. *Front Endocrinol (Lausanne)*. 2018;9:678. doi:10.3389/fendo.2018.00678

99. Simithy J, Sidoli S, Garcia BA. Integrating Proteomics and Targeted Metabolomics to Understand Global Changes in Histone Modifications. *Proteomics*. 2018;18(18):1700309. doi:10.1002/pmic.201700309

100. Smoyer W, Klein J. Integrating Proteomics and Metabolomics to Understand Pediatric Glomerular Disease Pathophysiology and Prognosis. http://grantome.com/grant/NIH/R01-DK110077-02. Accessed January 11, 2019.

101. Li X, Mu L, Hu X. Integrating proteomics, metabolomics and typical analysis to investigate the uptake and oxidative stress of graphene oxide and polycyclic aromatic hydrocarbons. *Environ Sci Nano*. 2018;5(1):115-129. doi:10.1039/C7EN00803A

102. Perna F, Berman SH, Soni RK, et al. Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. *Cancer Cell*. 2017;32(4):506-519.e5. doi:10.1016/j.ccell.2017.09.004

103. Gene Regulation, Transcriptomics and Proteomics. https://www.caister.com/hsp/abstracts/borrelia/04.html. Accessed January 11, 2019.

104. Haider S, Pal R. Integrated analysis of transcriptomic and proteomic data. *Curr Genomics*. 2013;14(2):91-110. doi:10.2174/1389202911314020003

105. Taylor G. Disintegration of water drops in an electric field. *Proc R Soc London Ser A Math Phys Sci*. 1964;280(1382):383-397. doi:10.1098/rspa.1964.0151

106. Edmond de H, Vincent S. *Mass Spectrometry: Principles and Applications*. (John Wiley & Sons L, ed.). Toronto; 2003.

107. Siuzdak G. *The Expading Role of Mass Spectrometry in Biotechnology*. 2nd ed.

San Diego, California: MCC Press; 2006.

108. B.A. Mamyrin, V.I. Karataev, D.V. Shmikk VAZ. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *JETP*. 1972;37(1):45. http://www.jetp.ac.ru/cgi-bin/e/index/e/37/1/p45?a=list. Accessed May 11, 2019.

109. Deutsch EW, Mendoza L, Shteynberg D, et al. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*. 2010;10(6):1150-1159. doi:10.1002/pmic.200900375

110. Zhang Y, Fonslow BR, Shan B, Baek M, Yates JR. Protein Analysis by Shotgun / Bottom-up Proteomics. 2013:1-101. doi:10.1021/cr3003533.Protein

111. Loo JA, Udseth HR, Smith RD. Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Anal Biochem*. 1989;179(2):404-412. doi:10.1016/0003-2697(89)90153-X

112. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR, III. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev*. 2013;113(4):2343-2394. doi:10.1021/cr3003533

113. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004;101(26):9528-9533. doi:10.1073/pnas.0402700101

114. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2

115. Matthiesen R, Trelle MB, Højrup P, Bunkenborg J, Jensen ON. VEMS 3.0: Algorithms and Computational Tools for Tandem Mass Spectrometry Based Identification of Post-translational Modifications in Proteins. *J Proteome Res*. 2005;4(6):2338-2347. doi:10.1021/pr050264q

116. Matthiesen R, Bunkenborg J, Stensballe A, Jensen ON, Welinder KG, Bauw G. Database-independent, database-dependent, and extended interpretation of peptide mass spectra in VEMS V2.0. *Proteomics*. 2004;4(9):2583-2593. doi:10.1002/pmic.200300792

117. Matthiesen R, Lundsgaard M, Welinder KG, Bauw G. Interpreting peptide mass spectra by VEMS. *Bioinformatics*. 2003;19(6):792-793. doi:10.1093/bioinformatics/btg059

118. Listgarten J, Emili A. Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry. *Mol Cell Proteomics*. 2005;4(4):419-434. doi:10.1074/mcp.R500005-MCP200

119. Jimmy KE, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994;5(11):976-989. doi:10.1016/1044-0305(94)80016-2

120. Nesvizhskii AI, Aebersold R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discov Today*. 2004;9(4):173-181. doi:10.1016/S1359-6446(03)02978-7

121. Hunt DF, Yates JR, Shabanowitz J, Winston S, Hauer CR. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 1986;83(17):6233-6237. http://www.ncbi.nlm.nih.gov/pubmed/3462691. Accessed January 13, 2019.

122. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*. 1993;3(6):327-332. http://www.ncbi.nlm.nih.gov/pubmed/15335725. Accessed January 13, 2019.

123. Akhtar MN, Southey BR, Andrén PE, Sweedler J V, Rodriguez-Zas SL. Evaluation of database search programs for accurate detection of neuropeptides in tandem mass spectrometry experiments. *J Proteome Res*. 2012;11(12):6044-6055. doi:10.1021/pr3007123

124. Craig R, Beavis RC. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun Mass Spectrom*. 2003;17(20):2310-2316. doi:10.1002/rcm.1198

125. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004;20(9):1466-1467. doi:10.1093/bioinformatics/bth092

126. Eng JK, Fischer B, Grossmann J, MacCoss MJ. A Fast SEQUEST Cross Correlation Algorithm. *J Proteome Res*. 2008;7(10):4598-4602. doi:10.1021/pr800420s

127. Eng JK, Jahan TA, Hoopmann MR. Comet: An open-source MS/MS sequence database search tool. *Proteomics*. 2013;13(1):22-24. doi:10.1002/pmic.201200439

128. Phanstiel DH, Brumbaugh J, Wenger CD, et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat Methods*. 2011;8(10):821-827. doi:10.1038/nmeth.1699

129. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res*. 2011;10(4):1794-1805. doi:10.1021/pr101065j

130. Zhang N, Li X, Ye M, Pan S, Schwikowski B, Aebersold R. ProbIDtree: An automated software program capable of identifying multiple peptides from a single collision-induced dissociation spectrum collected by a tandem mass spectrometer. *Proteomics*. 2005;5(16):4096-4106. doi:10.1002/pmic.200401260

131. Bern M, Finney G, Hoopmann MR, Merrihew G, Toth MJ, MacCoss MJ. Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem*. 2010;82(3):833-841. doi:10.1021/ac901801b

132. Wang J, Pérez-Santiago J, Katz JE, Mallick P, Bandeira N. Peptide Identification from Mixture Tandem Mass Spectra. *Mol Cell Proteomics*. 2010;9(7):1476-1485. doi:10.1074/mcp.M000136-MCP201

133. Houel S, Abernathy R, Renganathan K, Meyer-Arendt K, Ahn NG, Old WM. Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies. *J Proteome Res*. 2010;9(8):4152-4160. doi:10.1021/pr1003856

134. Gupta N, Bandeira N, Keich U, Pevzner PA. Target-Decoy Approach and False Discovery Rate: When Things May Go Wrong. *J Am Soc Mass Spectrom*. 2011;22(7):1111-1120. doi:10.1007/s13361-011-0139-3

135. Shortreed MR, Wenger CD, Frey BL, et al. Global Identification of Protein Post-translational Modifications in a Single-Pass Database Search. *J Proteome Res*. 2015;14(11):4714-4720. doi:10.1021/acs.jproteome.5b00599

136. Qiyao L, Shortreed MR, Wenger CD, et al. Global Post-Translational Modification Discovery. *J Proteome Res*. 2017;16(4):1383-1390. doi:10.1021/acs.jproteome.6b00034

137. Zhang G, Ueberheide BM, Waldemarson S, et al. Protein quantitation using mass spectrometry. *Methods Mol Biol*. 2010;673:211-222. doi:10.1007/978-1-60761-842-3_13

138. Wang M, You J, Bemis KG, Tegeler TJ, Brown DPG. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Briefings Funct Genomics Proteomics*. 2008;7(5):329-339. doi:10.1093/bfgp/eln031

139. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13(9):2513-2526. doi:10.1074/mcp.M113.031591

140. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098

141. Pisanu S, Ghisaura S, Pagnozzi D, et al. The sheep milk fat globule membrane proteome. *J Proteomics*. 2011;74(3):350-358. doi:10.1016/J.JPROT.2010.11.011

142. Zybailov BL, Florens L, Washburn MP. Quantitative shotgun proteomics using a

protease with broad specificity and normalized spectral abundance factors. *Mol Biosyst*. 2007;3(5):354. doi:10.1039/b701483j

143. Arike L, Peil L. Spectral Counting Label-Free Proteomics. In: Humana Press, New York, NY; 2014:213-222. doi:10.1007/978-1-4939-0685-7_14

144. Paoletti AC, Parmely TJ, Tomomori-Sato C, et al. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci U S A*. 2006;103(50):18928-18933. doi:10.1073/pnas.0606379103

145. Dheilly NM, Raftos DA, Haynes PA, Smith LC, Nair S V. Shotgun proteomics of coelomic fluid from the purple sea urchin, Strongylocentrotus purpuratus. *Dev Comp Immunol*. 2013;40(1):35-50. doi:10.1016/J.DCI.2013.01.007

146. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

147. Timmins-Schiffman E, Coffey WD, Hua W, Nunn BL, Dickinson GH, Roberts SB. Shotgun proteomics reveals physiological response to ocean acidification in Crassostrea gigas. *BMC Genomics*. 2014;15(1):951. doi:10.1186/1471-2164-15-951

148. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem*. 2007;389(4):1017-1031. doi:10.1007/s00216-007-1486-6

149. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

150. Ishihama Y, Oda Y, Tabata T, et al. Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein. *Mol Cell Proteomics*. 2005;4(9):1265-1272. doi:10.1074/mcp.M500061-MCP200

151. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces c erevisiae*. *J Proteome Res*. 2006;5(9):2339-2347. doi:10.1021/pr060161n

152. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

153. Vaudel M, Burkhart JM, Zahedi RP, et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotechnol*. 2015;33(1):22-24.

doi:10.1038/nbt.3109

154. Solntsev SK, Shortreed MR, Frey BL, Smith LM. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J Proteome Res*. 2018;17(5):1844-1851. doi:10.1021/acs.jproteome.7b00873

155. Partridge JD, Sanguinetti G, Dibden DP, Roberts RE, Poole RK, Green J. Transition of Escherichia coli from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components. *J Biol Chem*. 2007;282(15):11230-11237. doi:10.1074/jbc.M700728200

156. Jones SA, Gibson T, Maltby RC, et al. Anaerobic respiration of Escherichia coli in the mouse intestine. *Infect Immun*. 2011;79(10):4218-4226. doi:10.1128/IAI.05395-11

157. Murashko ON, Lin-Chao S. Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. *Proc Natl Acad Sci U S A*. 2017;114(38):E8025-E8034. doi:10.1073/pnas.1703731114

158. Kim SO, Merchant K, Nudelman R, et al. OxyR: A Molecular Code for Redox-Related Signaling. *Cell*. 2002;109(3):383-396. doi:10.1016/S0092-8674(02)00723-7

159. Zheng M, Aslund F, Storz G. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science*. 1998;279(5357):1718-1721. http://www.ncbi.nlm.nih.gov/pubmed/9497290. Accessed April 10, 2019.

160. Storz G, Tartaglia LA, Ames BN. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science*. 1990;248(4952):189-194. http://www.ncbi.nlm.nih.gov/pubmed/2183352. Accessed April 10, 2019.

161. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol*. 2014;24(10):R453-62. doi:10.1016/j.cub.2014.03.034

162. Lu J, Holmgren A. The thioredoxin antioxidant system. *Free Radic Biol Med*. 2014;66:75-87. doi:10.1016/j.freeradbiomed.2013.07.036

163. Boronat S, Domènech A, Paulo E, et al. Thiol-based H2O2 signalling in microbial systems. *Redox Biol*. 2014;2:395-399. doi:10.1016/j.redox.2014.01.015

164. Ray PD, Huang B-W, Tsuji Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cell Signal*. 2012;24(5):981-990. doi:10.1016/j.cellsig.2012.01.008

165. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8(10):813-

824. doi:10.1038/nrm2256

166. Couturier J, Chibani K, Jacquot J-P, Rouhier N. Cysteine-based redox regulation and signaling in plants. *Front Plant Sci*. 2013;4:105. doi:10.3389/fpls.2013.00105

167. Mou Z, Fan W, Dong X. Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell*. 2003;113(7):935-944. http://www.ncbi.nlm.nih.gov/pubmed/12837250. Accessed April 18, 2019.

168. Finkel T, Holbrook NJ. Oxidants, oxidative stress and the biology of ageing. *Nature*. 2000;408(6809):239-247. doi:10.1038/35041687

169. Maaty WS, Selvig K, Ryder S, et al. Proteomic analysis of sulfolobus solfataricus during sulfolobus turreted icosahedral virus infection. *J Proteome Res*. 2012;11(2):1420-1432. doi:10.1021/pr201087v

170. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

171. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-1372. doi:10.1038/nbt.1511

172. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13(9):731-740. doi:10.1038/nmeth.3901

173. Finn TJ, Shewaramani S, Leahy SC, Janssen PH, Moon CD. Dynamics and genetic diversification of Escherichia coli during experimental adaptation to an anaerobic environment. *PeerJ*. 2017;5:e3244. doi:10.7717/peerj.3244

174. von Wulffen J, RecogNice-Team, Sawodny O, Feuer R. Transition of an Anaerobic Escherichia coli Culture to Aerobiosis: Balancing mRNA and Protein Levels in a Demand-Directed Dynamic Flux Balance Analysis. *PLoS One*. 2016;11(7):e0158711. doi:10.1371/journal.pone.0158711

175. Trotter EW, Rolfe MD, Hounslow AM, et al. Reprogramming of Escherichia coli K-12 Metabolism during the Initial Phase of Transition from an Anaerobic to a Micro-Aerobic Environment. Roop RM, ed. *PLoS One*. 2011;6(9):e25501. doi:10.1371/journal.pone.0025501

176. Constantinidou C, Hobman JL, Griffiths L, et al. A Reassessment of the FNR Regulon and Transcriptomic Analysis of the Effects of Nitrate, Nitrite, NarXL, and NarQP as Escherichia coli K12 Adapts from Aerobic to Anaerobic Growth. *J Biol Chem*. 2006;281(8):4802-4815. doi:10.1074/jbc.M512312200

177. Kramer G, Sprenger RR, Speijer D, et al. Proteome-wide Alterations in Escherichia coli Translation Rates upon Anaerobiosis . *Mol Cell Proteomics*. 2010;9(11):2508-2516. doi:10.1074/mcp.m110.001826

178. Weber A, Kögl SA, Jung K. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli. *J Bacteriol*. 2006;188(20):7165-7175. doi:10.1128/JB.00508-06

179. Kang Y, Weber KD, Qiu Y, Kiley PJ, Blattner FR. Genome-wide expression analysis indicates that FNR of Escherichia coli K-12 regulates a large number of genes of unknown function. *J Bacteriol*. 2005;187(3):1135-1160. doi:10.1128/JB.187.3.1135-1160.2005

180. Salmon KA, Hung S -p., Steffen NR, et al. Global Gene Expression Profiling in Escherichia coli K12: EFFECTS OF OXYGEN AVAILABILITY AND ArcA. *J Biol Chem*. 2005;280(15):15084-15096. doi:10.1074/jbc.M414030200

181. Tyanova S, Cox J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. In: Humana Press, New York, NY; 2018:133-148. doi:10.1007/978-1-4939-7493-1_7

182. Weber J, Kayser A, Rinas U. Metabolic flux analysis of Escherichia coli in glucose-limited continuous culture. II. Dynamic response to famine and feast, activation of the methylglyoxal pathway and oscillatory behaviour. *Microbiology*. 2005;151(3):707-716. doi:10.1099/mic.0.27482-0

183. Epstein MG, Reeves BD, Maaty WS, et al. Enhanced Sensitivity Employing Zwitterionic and pI Balancing Dyes (Z-CyDyes) Optimized for 2D-Gel Electrophoresis Based on Side Chain Modifications of CyDye Fluorophores. New Tools For Use in Proteomics and Diagnostics. *Bioconjug Chem*. 2013;24(9):1552-1561. doi:10.1021/bc4002213

184. Constantinidou C, Hobman JL, Griffiths L, et al. A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as Escherichia coli K12 adapts from aerobic to anaerobic growth. *J Biol Chem*. 2006;281(8):4802-4815. doi:10.1074/jbc.M512312200

185. Tokmina-Lukaszewska M, Movahed N, Lusczek E, Mulier K, Beilman G, Bothner B. Transformation of UPLC-MS Data Overcomes Extreme Variability in Urine Concentration and Metabolite Fold Change. *Curr Metabolomics*. 2014;2(2):78-87. doi:10.2174/2213235X0202141127102223

186. Heinemann J, Mazurie A, Tokmina-Lukaszewska M, Beilman GJ, Bothner B. Application of support vector machines to metabolomics experiments with limited replicates. *Metabolomics*. 2014;10(6):1121-1128. doi:10.1007/s11306-014-0651-0

187. Kargalioglu Y, Imlay JA. Importance of anaerobic superoxide dismutase synthesis

in facilitating outgrowth of Escherichia coli upon entry into an aerobic habitat. *J Bacteriol*. 1994;176(24):7653-7658. doi:10.1128/JB.176.24.7653-7658.1994

188. MCDANIEL LE, BAILEY EG, ZIMMERLI A. EFFECT OF OXYGEN SUPPLY RATES ON GROWTH OF ESCHERICHIA COLI. *Appl Microbiol*. 1965;13(1):109-114. http://www.ncbi.nlm.nih.gov/pubmed/14264837. Accessed April 21, 2019.

189. Nimmo GA, Nimmo HG. The regulatory properties of isocitrate dehydrogenase kinase and isocitrate dehydrogenase phosphatase from Escherichia coli ML308 and the roles of these activities in the control of isocitrate dehydrogenase. *Eur J Biochem*. 1984;141(2):409-414. http://www.ncbi.nlm.nih.gov/pubmed/6329757. Accessed April 22, 2019.

190. LaPorte DC, Walsh K, Koshland DE. The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. *J Biol Chem*. 1984;259(22):14068-14075. http://www.ncbi.nlm.nih.gov/pubmed/6389540. Accessed April 22, 2019.

191. Walsh K, Koshland DE. Determination of flux through the branch point of two metabolic cycles. The tricarboxylic acid cycle and the glyoxylate shunt. *J Biol Chem*. 1984;259(15):9646-9654. http://www.ncbi.nlm.nih.gov/pubmed/6378912. Accessed April 22, 2019.

192. Cortay JC, Bleicher F, Duclos B, et al. Utilization of acetate in Escherichia coli: structural organization and differential expression of the ace operon. *Biochimie*. 1989;71(9-10):1043-1049. http://www.ncbi.nlm.nih.gov/pubmed/2512996. Accessed April 22, 2019.

193. Madigan M, Martinko J. *Brock's Biology of Microorganisms,*. 11th editi. NJ: Pearson Prentice Hall; 2006. http://enacademic.com/dic.nsf/enwiki/2639682/Mixed. Accessed April 22, 2019.

194. Ku SY, Yip P, Howell PL. Structure of Escherichia coli tryptophanase. *Acta Crystallogr D Biol Crystallogr*. 2006;62(Pt 7):814-823. doi:10.1107/S0907444906019895

195. Phillips RS, Demidkina T V., Faleev NG. Structure and mechanism of tryptophan indole-lyase and tyrosine phenol-lyase. *Biochim Biophys Acta - Proteins Proteomics*. 2003;1647(1-2):167-172. doi:10.1016/S1570-9639(03)00089-X

196. Nihira T, Yasuda T, Kakizono T, et al. Functional role of cysteinyl residues in tryptophanase. *Eur J Biochem*. 1985;149(1):129-133. http://www.ncbi.nlm.nih.gov/pubmed/3888623. Accessed April 13, 2019.

197. Liu M, Hou J, Huang L, et al. Site-specific proteomics approach for study protein S-nitrosylation. *Anal Chem*. 2010;82(17):7160-7168. doi:10.1021/ac100569d

198. Forrester MT, Thompson JW, Foster MW, Nogueira L, Moseley MA, Stamler JS. Proteomic analysis of S-nitrosylation and denitrosylation by resin-assisted capture. *Nat Biotechnol*. 2009;27(6):557-559. doi:10.1038/nbt.1545

199. Lee J, Maeda T, Hong SH, Wood TK. Reconfiguring the Quorum-Sensing Regulator SdiA of Escherichia coli To Control Biofilm Formation via Indole and N-Acylhomoserine Lactones. *Appl Environ Microbiol*. 2009;75(6):1703-1716. doi:10.1128/AEM.02081-08

200. Lee Y, Yeom J, Kim J, Jung J, Jeon CO, Park W. Phenotypic and physiological alterations by heterologous acylhomoserine lactone synthase expression in Pseudomonas putida. *Microbiology*. 2010;156(12):3762-3772. doi:10.1099/mic.0.041095-0

201. Chu W, Zere TR, Weber MM, et al. Indole Production Promotes Escherichia coli Mixed-Culture Growth with Pseudomonas aeruginosa by Inhibiting Quorum Signaling. *Appl Environ Microbiol*. 2012;78(2):411-419. doi:10.1128/AEM.06396-11

202. Molina-Santiago C, Daddaoua A, Fillet S, Duque E, Ramos J-L. Interspecies signalling: *Pseudomonas putida* efflux pump TtgGHI is activated by indole to increase antibiotic resistance. *Environ Microbiol*. 2014;16(5):1267-1281. doi:10.1111/1462-2920.12368

CHAPTER FOUR

CONCLUDING REMARKS

E. coli is one of the best-characterized prokaryotes and has long served as a valuable model system to investigate microbial responses to changes in their environments. Despite being well studied and examined under growth conditions in the presence or absence of oxygen, a global understanding of its proteome changes, i.e. of the molecular changes occurring at the level of protein expression and protein post-translational modifications, during reversible transitions from aerobic to anaerobic growth conditions remains incomplete. Particularly, the impact of transitioning between aerobic and anaerobic conditions had not been investigated at the level of thiol signaling in *E. coli.*

The work presented in this thesis was motivated by an interest to better understand the impact and biological significance of biomolecular oxidation and reduction changes in biological processes. This led to carefully examining how changes in thiol oxidation states mediate a facultative anaerobe's (such as *E. coli*) ability to adapt its metabolism during the transitions from aerobic to anaerobic growth and vice versa. Thiol signaling, referred to herein as "thiol switches", mediates numerous biological activities and biological responses including changes in gene expression, protein synthesis and protein post-translational modifications, and microbial responses to environmental stresses including oxidative stress. The activity of the OxyR transcription factor and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) proteins represent excellent examples of proteins whose function is modulated by thiol switches and whose functions respond to thiol signaling.

To better explore the relatively uncharted territory of “thiol switches”, a mass-spectrometry (MS)-based proteomics workflow, i.e. a global assessment of protein changes accompanying transitions from aerobic to anaerobic microbial growth conditions, was developed and refined. High throughput MS-based proteomics was ideally suited for assessing changes in *E. coli* protein expression profiles, regulation of protein function and post-translational modifications (PTM) following transitions from aerobic to anaerobic modes of growth. Development of a high-throughput MS pipeline was motivated by the fact that MS datasets are often large complex, and require sophisticated software pipelines for data conversion, processing, normalization, and statistical analysis of data significance, and ultimately to identify the best peptide/protein match from a database of user-selected protein sequences.

After extensive optimization of a general protocol for HPLC-MS/MS data processing, data normalization, peptide pattern matching, and statistical analysis, the analytical pipeline was fine-tuned to investigate the specific proteome-wide changes and chemical modifications of cysteines in *E. coli* proteins. This analytical approach has been possible because cysteine residues can be reversibly covalently modified, leading to different thiol oxidation/reduction states , i.e. different detectable signals that are dependent on the type of oxidizing or reducing agent used. The mass spectrometry workflow developed here utilized all open-source software and publicly accessible databases, creating a transparent and reproducible proteomics method, which can be easily shared among research laboratories, and can be easily adapted to a wide range of diverse research applications.

Following the development of a comprehensive MS-based proteomics approach, the research presented here focused on characterizing changes in oxidation/reduction states and the chemical forms of protein thiol switches, as *E. coli* transitions from aerobic to anaerobic growth conditions (or vice versa) over time. Unexpectedly, differential alkylation analysis of a large number of cysteine-containing proteins of *E. coli* revealed the presence of a higher proportion of oxidized cysteines when the organism grows under oxygen-limited anaerobic conditions, compared to the number of oxidized cysteines found in cysteine-rich proteins when *E. coli* is grown under oxygenic conditions (i.e with oxygen). This result is not only unexpected, but also highly novel as such a phenomenon has never been reported in the literature prior to this study. Interestingly, different cysteine-containing proteins exhibited different patterns of thiol oxidation/reduction states as *E. coli* transitioned from aerobic to anaerobic growth conditions. For example, for some of the cysteine-containing proteins, their thiol oxidation/reduction states appeared to be influenced by the metabolic state of *E. coli* during the transition from aerobic to anaerobic (i.e. O_2 absent or limited) growth conditions. In contrast, the oxidation/reduction states of other cysteine-containing proteins did not seem to be impacted by changes in *E. coli* metabolism, as the organism transitions from aerobic to anaerobic growth, or vice versa,

Most interestingly, a small subset of cysteine-containing proteins, which we identified as potential "thiol switches," displayed specific patterns of oxidization/reduction states following *E. coli*'s transitions from aerobic to anaerobic growth conditions. An interesting example of a thiol-switch protein identified included the enzyme, tryptophanase, which is responsible for generating indole, a small molecule

which has been identified as an important modulator of bacterial quorum sensing in addition being a metabolic byproduct of tryptophan degradation. The HPLC-MS/MS based differential alkylation method used to identify this thiol switch can be easily applied to other thiol signaling/"thiol-switches" in other facultative anaerobic microbes.

In order to investigate the physiological role of cysteine (Cys) oxidative modification, it is important to monitor the presence of multiple Cys oxidation states as the organism transitions between different metabolic states. This was achieved in this by using specific reducing agents which enabled the capture of different Cys oxidation/reduction and chemically modified forms (see Figure 1.4 for example). Specific alkylation reagents, such as ascorbate could be used in the presence of Cu^{2+} to target nitrosylated forms of cysteines as described in chapter 1. This approach revealed a deeper layer of molecular information which is specific to different Cys oxidative modification forms. However, this type of reversible and labile modification is potentially challenging to reliably monitor over time as a facultative anaerobe transitions from aerobic to anaerobic environments. There exists therefore a need to develop more stable Cys modification agents such as alkylation probes that can specifically report on Cys PTM modifications. Such probes can reliably report on changes in Cys oxidation/reduction states over time, inform on an organism's physiological changes, and may be used to better understand the molecular changes taking place at the protein levels in health and disease. Applications of this approach have already been reported in the literature, providing support to the notion that thiol signaling via rapid changes in Cys oxidation/reduction states play an important role in heart diseases and cancers [69].

Mass spectrometry-based proteomics has proven to be an extremely valuable high throughput method to monitor changes in protein expression, protein function, and protein post-translational modifications (PTMs). With new advancements in the field, mass spectrometry-based proteomics has gained the power to monitor changes in PTMs confidently, allowing researchers to better understand the role of PTMs over time during physiological or environmental changes. Herein, we implemented Cys differential alkylation together with shotgun proteomics to investigate the biological function of thiol switches/signaling and to enhance our understanding of its role in *E. coli*'s ability to adapt to changing aerobic/anaerobic environments and to report on the versatility of reversible modifications of cysteines such as S-nitrosylation.

Lastly, the mass spectrometry approach that we have established allows us to answer questions at the system level, reporting on global changes in the proteome of an organism, and focusing on specific molecular processes, such as better understanding how changes in the oxidation/reduction states of Cys-containing proteins impact an organism's response to changing environmental conditions. Ultimately, this MS-based analysis has enabled us to generate meaningful biological information and to uncover previously unknown phenomena in an organism such as *E. coli* that has been so well studied over the past 100 years.

REFERENCES CITED

1. Blattner FR. The Complete Genome Sequence of Escherichia coli K-12. *Science (80-)*. 1997;277(5331):1453-1462. doi:10.1126/science.277.5331.1453

2. Feist AM, Palsson BØ. The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol*. 2008;26(6):659-667. doi:10.1038/nbt1401

3. Barrick JE, Yu DS, Yoon SH, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*. 2009;461(7268):1243-1247. doi:10.1038/nature08480

4. Richmond C, Glasner JD, Mau R, Jin H, Blattner FR. Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Res*. 1999;27(19):3821-3835. doi:10.1093/nar/27.19.3821

5. Chang D-E, Smalley DJ, Tucker DL, et al. Noninvasive measurement of anatomic structure and intraluminal oxygenation in the gastrointestinal tract of living mice with spatial and spectral EPR imaging. *PNAS*. 2004;96(8):4586-4591. doi:10.1073/pnas.96.8.4586

6. Kolisnychenko V, Plunkett G, Herring CD, et al. Engineering a reduced Escherichia coli genome. *Genome Res*. 2002;12(4):640-647. doi:10.1101/gr.217202

7. Chen S, Lesnik EA, Hall TA, et al. A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. *Biosystems*. 2002;65(2-3):157-177. doi:10.1016/S0303-2647(02)00013-8

8. Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Curr Opin Microbiol*. 2015;23:133-140. doi:10.1016/J.MIB.2014.11.011

9. Hain T, Hossain H, Chatterjee SS, et al. Temporal transcriptomic analysis of the Listeria monocytogenes EGD-e σB regulon. *BMC Microbiol*. 2008;8(1):20. doi:10.1186/1471-2180-8-20

10. Chueca B, Pagán R, García-Gonzalo D. Transcriptomic analysis of Escherichia coli MG1655 cells exposed to pulsed electric fields. *Innov Food Sci Emerg Technol*. 2015;29:78-86. doi:10.1016/J.IFSET.2014.09.003

11. Nie L, Wu G, Culley DE, Scholten JCM, Zhang W. Integrative Analysis of Transcriptomic and Proteomic Data: Challenges, Solutions and Applications. *Crit Rev Biotechnol*. 2007;27(2):63-75. doi:10.1080/07388550701334212

12. Weber A, Kögl SA, Jung K. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli. *J Bacteriol*. 2006;188(20):7165-7175. doi:10.1128/JB.00508-06

13. Vertommen D, Ruiz N, Leverrier P, Silhavy TJ, Collet J-F. Characterization of the role of the Escherichia coli periplasmic chaperone SurA using differential proteomics. *Proteomics*. 2009;9(9):2432-2443. doi:10.1002/pmic.200800794

14. Schmidt A, Kochanowski K, Vedelaar S, et al. The quantitative and condition-dependent Escherichia coli proteome. *Nat Biotechnol*. 2016;34(1):104-110. doi:10.1038/nbt.3418

15. Lai E-M, Nair U, Phadke ND, Maddock JR. Proteomic screening and identification of differentially distributed membrane proteins in Escherichia coli. *Mol Microbiol*. 2004;52(4):1029-1044. doi:10.1111/j.1365-2958.2004.04040.x

16. Rabinowitz JD, Kimball E. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Anal Chem*. 2007;79(16):6167-6173. doi:10.1021/ac070470c

17. Nobeli I, Ponstingl H, Krissinel EB, Thornton JM. A Structure-based Anatomy of the E. coli Metabolome. *J Mol Biol*. 2003;334(4):697-719. doi:10.1016/J.JMB.2003.10.008

18. Heinemann J, Noon B, Mohigmi MJ, Mazurie A, Dickensheets DL, Bothner B. Real-Time Digitization of Metabolomics Patterns from a Living System Using Mass Spectrometry. *J Am Soc Mass Spectrom*. 2014;25(10):1755-1762. doi:10.1007/s13361-014-0922-z

19. Yuan J, Doucette CD, Fowler WU, et al. Metabolomics-driven quantitative analysis of ammonia assimilation in E. coli. *Mol Syst Biol*. 2009;5(1):302. doi:10.1038/msb.2009.60

20. Tweeddale H, Notley-McRobb L, Ferenci T. Effect of Slow Growth on Metabolism of Escherichia coli, as Revealed by Global Metabolite Pool ("Metabolome") Analysis. *J Bacteriol*. 1998;180(19):5109-5116. https://jb.asm.org/content/180/19/5109.short. Accessed April 16, 2019.

21. Partridge JD, Scott C, Tang Y, Poole RK, Green J. Escherichia coli transcriptome dynamics during the transition from anaerobic to aerobic conditions. *J Biol Chem*. 2006;281(38):27806-27815. doi:10.1074/jbc.M603450200

22. Guest JR, Green, J. Irvine AS, Spiro S. *Regulation of Gene Expression in Escherichia Coli*. Boston, MA: Springer US; 1996. doi:10.1007/978-1-4684-8601-8

23. Gennis RB S V. *Respiration, Escherichia Coli and Salmonella Cellular and*

Molecular Biology. 2nd ed. (Neidhardt FC, ed.). Washington, D.C: ASM Press; 1996.

24. Unden G, Bongaerts J. Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. *Biochim Biophys Acta*. 1997;1320(3):217-234. http://www.ncbi.nlm.nih.gov/pubmed/9230919. Accessed May 19, 2019.

25. Pecher A, Blaschkowski HP, Knappe K, Böck A. Expression of pyruvate formate-lyase of <i>Escherichia coli</i> from the cloned structural gene. *Arch Microbiol*. 1982;132(4):365-371. doi:10.1007/BF00413390

26. Dawes EA, Foster SM. The formation of ethanol in Escherichia coli. *Biochim Biophys Acta*. 1956;22(2):253-265. http://www.ncbi.nlm.nih.gov/pubmed/13382840. Accessed November 9, 2016.

27. Gupta S, Clark DP. Escherichia coli derivatives lacking both alcohol dehydrogenase and phosphotransacetylase grow anaerobically by lactate fermentation. *J Bacteriol*. 1989;171(7):3650-3655. http://www.ncbi.nlm.nih.gov/pubmed/2661531. Accessed November 9, 2016.

28. Stryer L. *Biochemistry*. 4th ed. (W. H. Freeman and Company., ed.). New York-Basingstoke; 1995.

29. Murashko ON, Lin-Chao S. Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. doi:10.1073/pnas.1703731114

30. Jones SA, Chowdhury FZ, Fabich AJ, et al. Respiration of Escherichia coli in the mouse intestine. *Infect Immun*. 2007;75(10):4891-4899. doi:10.1128/IAI.00484-07

31. Marteyn B, West NP, Browning DF, et al. Modulation of Shigella virulence in response to available oxygen in vivo. *Nature*. 2010;465(7296):355-358. doi:10.1038/nature08970

32. Davies K. Oxidative Stress, Antioxidant Defenses, and Damage Removal, Repair, and Replacement Systems. *IUBMB Life*. 2001;50(4):279-289. doi:10.1080/713803728

33. Davies KJ. Oxidative stress: the paradox of aerobic life. *Biochem Soc Symp*. 1995;61:1-31. doi:10.1042/BSS0610001

34. Riley PA. Free radicals in biology: oxidative stress and the effects of ionizing radiation. *Int J Radiat Biol*. 1994;65(1):27-33. http://www.ncbi.nlm.nih.gov/pubmed/7905906. Accessed May 8, 2019.

35. Sahnoun Z, Jamoussi K, Zeghal KM. [Free radicals and antioxidants: human

physiology, pathology and therapeutic aspects]. *Therapie*. 1997;52(4):251-270. http://www.ncbi.nlm.nih.gov/pubmed/9437876. Accessed May 8, 2019.

36. Sakac V, Sakac M. [Free oxygen radiacals and kidney diseases--part I]. *Med Pregl*. 2000;53(9-10):463-74. http://www.ncbi.nlm.nih.gov/pubmed/11320727. Accessed May 8, 2019.

37. Devasagayam TPA, Tilak JC, Boloor KK, Sane KS, Ghaskadbi SS, Lele RD. Free radicals and antioxidants in human health: current status and future prospects. *J Assoc Physicians India*. 2004;52:794-804. http://www.ncbi.nlm.nih.gov/pubmed/15909857. Accessed May 8, 2019.

38. Finkel T. Signal transduction by reactive oxygen species. *J Cell Biol*. 2011;194(1):7-15. doi:10.1083/jcb.201102095

39. Turkan I. ROS and RNS: key signalling molecules in plants. *J Exp Bot*. 2018;69(14):3313-3315. doi:10.1093/jxb/ery198

40. Bigarella CL, Liang R, Ghaffari S. Stem cells and the impact of ROS signaling. *Development*. 2014;141(22):4206-4218. doi:10.1242/dev.107086

41. Zhang J, Wang X, Vikash V, et al. ROS and ROS-Mediated Cellular Signaling. *Oxid Med Cell Longev*. 2016;2016:1-18. doi:10.1155/2016/4350965

42. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol*. 2014;24(10):R453-62. doi:10.1016/j.cub.2014.03.034

43. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8(10):813-824. doi:10.1038/nrm2256

44. Knoefler D, Leichert LIO, Thamsen M, et al. About the dangers, costs and benefits of living an aerobic lifestyle. *Biochem Soc Trans*. 2014;42(4):917-921. doi:10.1042/BST20140108

45. Halliwell, B.; Gutteridge J, ed. Free radicals in biology and medicine. In: 4th ed. Oxford: Oxford University Press; 2007.

46. Brandes N, Schmitt S, Jakob U. Thiol-based redox switches in eukaryotic proteins. *Antioxid Redox Signal*. 2009;11(5):997-1014. doi:10.1089/ars.2008.2285

47. Klomsiri C, Karplus PA, Poole LB. Cysteine-based redox switches in enzymes. *Antioxid Redox Signal*. 2011;14(6):1065-1077. doi:10.1089/ars.2010.3376

48. Paulsen CE, Carroll KS. Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chem Biol*. 2010;5(1):47-62. doi:10.1021/cb900258z

49. Antelmann H, Helmann JD. Thiol-based redox switches and gene regulation. *Antioxid Redox Signal*. 2011;14(6):1049-1063. doi:10.1089/ars.2010.3400

50. Jang HH, Lee KO, Chi YH, et al. Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell*. 2004;117(5):625-635. doi:10.1016/j.cell.2004.05.002

51. Winter J, Ilbert M, Graf PCF, Ozcelik D, Jakob U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell*. 2008;135(4):691-701. doi:10.1016/j.cell.2008.09.024

52. Dalle-Donne I, Rossi R, Colombo G, Giustarini D, Milzani A. Protein S-glutathionylation: a regulatory device from bacteria to humans. *Trends Biochem Sci*. 2009;34(2):85-96. doi:10.1016/j.tibs.2008.11.002

53. Burgoyne JR, Eaton P. A rapid approach for the detection, quantification, and discovery of novel sulfenic acid or S-nitrosothiol modified proteins using a biotin-switch method. *Methods Enzymol*. 2010;473:281-303. doi:10.1016/S0076-6879(10)73015-9

54. Yeh JI, Claiborne A, Hol WG. Structure of the native cysteine-sulfenic acid redox center of enterococcal NADH peroxidase refined at 2.8 A resolution. *Biochemistry*. 1996;35(31):9951-9957. doi:10.1021/bi961037s

55. Foster MW, Hess DT, Stamler JS. Protein S-nitrosylation in health and disease: a current perspective. *Trends Mol Med*. 2009;15(9):391-404. doi:10.1016/j.molmed.2009.06.007

56. Kokkola T, Savinainen JR, Mönkkönen KS, Retamal MD, Laitinen JT. S-nitrosothiols modulate G protein-coupled receptor signaling in a reversible and highly receptor-specific manner. *BMC Cell Biol*. 2005;6(1):21. doi:10.1186/1471-2121-6-21

57. Nozik-Grayck E, Whalen EJ, Stamler JS, McMahon TJ, Chitano P, Piantadosi CA. S-nitrosoglutathione inhibits alpha1-adrenergic receptor-mediated vasoconstriction and ligand binding in pulmonary artery. *Am J Physiol Lung Cell Mol Physiol*. 2006;290(1):L136-43. doi:10.1152/ajplung.00230.2005

58. Whalen EJ, Foster MW, Matsumoto A, et al. Regulation of beta-adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2. *Cell*. 2007;129(3):511-522. doi:10.1016/j.cell.2007.02.046

59. Benhar M, Forrester MT, Hess DT, Stamler JS. Regulated protein denitrosylation by cytosolic and mitochondrial thioredoxins. *Science*. 2008;320(5879):1050-1054. doi:10.1126/science.1158265

60. Chanvorachote P, Nimmannit U, Wang L, et al. Nitric oxide negatively regulates

Fas CD95-induced apoptosis through inhibition of ubiquitin-proteasome-mediated degradation of FLICE inhibitory protein. *J Biol Chem*. 2005;280(51):42044-42050. doi:10.1074/jbc.M510080200

61. Hoffmann J, Haendeler J, Zeiher AM, Dimmeler S. TNFalpha and oxLDL reduce protein S-nitrosylation in endothelial cells. *J Biol Chem*. 2001;276(44):41383-41387. doi:10.1074/jbc.M107566200

62. Kim J-E, Tannenbaum SR. S-Nitrosation regulates the activation of endogenous procaspase-9 in HT-29 human colon carcinoma cells. *J Biol Chem*. 2004;279(11):9758-9764. doi:10.1074/jbc.M312722200

63. Kim YM, Talanian R V, Billiar TR. Nitric oxide inhibits apoptosis by preventing increases in caspase-3-like activity via two distinct mechanisms. *J Biol Chem*. 1997;272(49):31138-31148. http://www.ncbi.nlm.nih.gov/pubmed/9388267. Accessed March 30, 2016.

64. Mannick JB, Hausladen A, Liu L, et al. Fas-induced caspase denitrosylation. *Science*. 1999;284(5414):651-654. http://www.ncbi.nlm.nih.gov/pubmed/10213689. Accessed March 30, 2016.

65. Huang Y, Man H-Y, Sekine-Aizawa Y, et al. S-nitrosylation of N-ethylmaleimide sensitive factor mediates surface expression of AMPA receptors. *Neuron*. 2005;46(4):533-540. doi:10.1016/j.neuron.2005.03.028

66. Lipton SA, Choi YB, Pan ZH, et al. A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. *Nature*. 1993;364(6438):626-632. doi:10.1038/364626a0

67. Mustafa AK, Kumar M, Selvakumar B, et al. Nitric oxide S-nitrosylates serine racemase, mediating feedback inhibition of D-serine formation. *Proc Natl Acad Sci U S A*. 2007;104(8):2950-2955. doi:10.1073/pnas.0611620104

68. Takahashi H, Shin Y, Cho S-J, et al. Hypoxia enhances S-nitrosylation-mediated NMDA receptor inhibition via a thiol oxygen sensor motif. *Neuron*. 2007;53(1):53-64. doi:10.1016/j.neuron.2006.11.023

69. Lowenstein CJ. Nitric oxide regulation of protein trafficking in the cardiovascular system. *Cardiovasc Res*. 2007;75(2):240-246. doi:10.1016/j.cardiores.2007.03.024

70. Matsushita K, Morrell CN, Cambien B, et al. Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. *Cell*. 2003;115(2):139-150. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2846406&tool=pmcentrez&rendertype=abstract. Accessed March 30, 2016.

71. Morrell CN, Matsushita K, Chiles K, et al. Regulation of platelet granule exocytosis by S-nitrosylation. *Proc Natl Acad Sci U S A*. 2005;102(10):3782-3787.

doi:10.1073/pnas.0408310102

72. Wang G, Moniri NH, Ozawa K, Stamler JS, Daaka Y. Nitric oxide regulates endocytosis by S-nitrosylation of dynamin. *Proc Natl Acad Sci U S A*. 2006;103(5):1295-1300. doi:10.1073/pnas.0508354103

73. Kim SF, Huri DA, Snyder SH. Inducible nitric oxide synthase binds, S-nitrosylates, and activates cyclooxygenase-2. *Science*. 2005;310(5756):1966-1970. doi:10.1126/science.1119407

74. Xu L, Han C, Lim K, Wu T. Activation of cytosolic phospholipase A2alpha through nitric oxide-induced S-nitrosylation. Involvement of inducible nitric-oxide synthase and cyclooxygenase-2. *J Biol Chem*. 2008;283(6):3077-3087. doi:10.1074/jbc.M705709200

75. Tian J, Kim SF, Hester L, Snyder SH. S-nitrosylation/activation of COX-2 mediates NMDA neurotoxicity. *Proc Natl Acad Sci U S A*. 2008;105(30):10537-10540. doi:10.1073/pnas.0804852105

76. Uehara T, Nakamura T, Yao D, et al. S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature*. 2006;441(7092):513-517. doi:10.1038/nature04782

77. Dalle-Donne I, Scaloni A, Giustarini D, et al. Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. *Mass Spectrom Rev*. 2005;24(1):55-99. doi:10.1002/mas.20006

78. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol*. 1990;58(3):157-161. doi:10.1007/BF00548927

79. Hillion M, Antelmann H. Thiol-based redox switches in prokaryotes. *Biol Chem*. 2015;396(5):415-444. doi:10.1515/hsz-2015-0102

80. Storz G, Tartaglia LA, Ames BN. The OxyR regulon. *Antonie Van Leeuwenhoek*. 1990;58(3):157-161. http://www.ncbi.nlm.nih.gov/pubmed/2256675. Accessed May 20, 2019.

81. Kim SO, Merchant K, Nudelman R, et al. OxyR: a molecular code for redox-related signaling. *Cell*. 2002;109(3):383-396. http://www.ncbi.nlm.nih.gov/pubmed/12015987. Accessed May 20, 2019.

82. Li Y, He Z-G. The Mycobacterial LysR-Type Regulator OxyS Responds to Oxidative Stress and Negatively Regulates Expression of the Catalase-Peroxidase Gene. Cornelis P, ed. *PLoS One*. 2012;7(1):e30186. doi:10.1371/journal.pone.0030186

83. Domenech P, Honoré N, Heym B, Cole ST. Role of OxyS of Mycobacterium

tuberculosis in oxidative stress: overexpression confers increased sensitivity to organic hydroperoxides. *Microbes Infect*. 2001;3(9):713-721. http://www.ncbi.nlm.nih.gov/pubmed/11489419. Accessed April 28, 2019.

84. Zheng C, Xu J, Li J, et al. Two Spx Regulators Modulate Stress Tolerance and Virulence in Streptococcus suis Serotype 2. Lemos JA, ed. *PLoS One*. 2014;9(9):e108197. doi:10.1371/journal.pone.0108197

85. Choi H, Kim S, Mukhopadhyay P, et al. Structural basis of the redox switch in the OxyR transcription factor. *Cell*. 2001;105(1):103-113. http://www.ncbi.nlm.nih.gov/pubmed/11301006. Accessed May 20, 2019.

86. Hildebrandt T, Knuesting J, Berndt C, Morgan B, Scheibe R. Cytosolic thiol switches regulating basic cellular functions: GAPDH as an information hub? *Biol Chem*. 2015;396(5):523-537. doi:10.1515/hsz-2014-0295

87. Grant CM, Quinn KA, Dawes IW. Differential protein S-thiolation of glyceraldehyde-3-phosphate dehydrogenase isoenzymes influences sensitivity to oxidative stress. *Mol Cell Biol*. 1999;19(4):2650-2656. http://www.ncbi.nlm.nih.gov/pubmed/10082531. Accessed April 28, 2019.

88. Ralser M, Wamelink MM, Kowald A, et al. Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. *J Biol*. 2007;6(4):10. doi:10.1186/jbiol61

89. Sampathkumar Krishnan ‡, Eva Y. Chi §, Stephen J. Wood ⦀, et al. Oxidative Dimer Formation Is the Critical Rate-Limiting Step for Parkinson's Disease α-Synuclein Fibrillogenesis†. 2002. doi:10.1021/BI026528T

90. Wojdyla K, Rogowska-Wrzesinska A. Differential alkylation-based redox proteomics - Lessons learnt. *Redox Biol*. 2015;6:240-252. doi:10.1016/j.redox.2015.08.005

91. Jaffrey SR, Snyder SH. The biotin switch method for the detection of S-nitrosylated proteins. *Sci STKE*. 2001;2001(86):pl1. doi:10.1126/stke.2001.86.pl1

92. Shelton MD, Chock PB, Mieyal JJ. Glutaredoxin: Role in Reversible Protein *S* -Glutathionylation and Regulation of Redox Signal Transduction and Protein Translocation. *Antioxid Redox Signal*. 2005;7(3-4):348-366. doi:10.1089/ars.2005.7.348

93. García-Santamarina S, Boronat S, Espadas G, Ayté J, Molina H, Hidalgo E. The oxidized thiol proteome in fission yeast—Optimization of an ICAT-based method to identify H2O2-oxidized proteins. *J Proteomics*. 2011;74(11):2476-2486. doi:10.1016/J.JPROT.2011.05.030

94. Wojdyla K, Williamson J, Roepstorff P, Rogowska-Wrzesinska A. The SNO/SOH

TMT strategy for combinatorial analysis of reversible cysteine oxidations. *J Proteomics*. 2015;113:415-434. doi:10.1016/J.JPROT.2014.10.015

95. Qu Z, Meng F, Bomgarden RD, et al. Proteomic Quantification and Site-Mapping of *S* -Nitrosylated Proteins Using Isobaric iodoTMT Reagents. *J Proteome Res*. 2014;13(7):3200-3211. doi:10.1021/pr401179v

96. Fares A, Nespoulous C, Rossignol M, Peltier J-B. Simultaneous Identification and Quantification of Nitrosylation Sites by Combination of Biotin Switch and ICAT Labeling. In: Humana Press, Totowa, NJ; 2014:609-620. doi:10.1007/978-1-62703-631-3_41

97. James P. Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys*. 1997;30(4):279-331. http://www.ncbi.nlm.nih.gov/pubmed/9634650. Accessed January 11, 2019.

98. Feng J, Zhang Q, Zhou Y, et al. Integration of Proteomics and Metabolomics Revealed Metabolite–Protein Networks in ACTH-Secreting Pituitary Adenoma. *Front Endocrinol (Lausanne)*. 2018;9:678. doi:10.3389/fendo.2018.00678

99. Simithy J, Sidoli S, Garcia BA. Integrating Proteomics and Targeted Metabolomics to Understand Global Changes in Histone Modifications. *Proteomics*. 2018;18(18):1700309. doi:10.1002/pmic.201700309

100. Smoyer W, Klein J. Integrating Proteomics and Metabolomics to Understand Pediatric Glomerular Disease Pathophysiology and Prognosis. http://grantome.com/grant/NIH/R01-DK110077-02. Accessed January 11, 2019.

101. Li X, Mu L, Hu X. Integrating proteomics, metabolomics and typical analysis to investigate the uptake and oxidative stress of graphene oxide and polycyclic aromatic hydrocarbons. *Environ Sci Nano*. 2018;5(1):115-129. doi:10.1039/C7EN00803A

102. Perna F, Berman SH, Soni RK, et al. Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. *Cancer Cell*. 2017;32(4):506-519.e5. doi:10.1016/j.ccell.2017.09.004

103. Gene Regulation, Transcriptomics and Proteomics. https://www.caister.com/hsp/abstracts/borrelia/04.html. Accessed January 11, 2019.

104. Haider S, Pal R. Integrated analysis of transcriptomic and proteomic data. *Curr Genomics*. 2013;14(2):91-110. doi:10.2174/1389202911314020003

105. Taylor G. Disintegration of water drops in an electric field. *Proc R Soc London Ser A Math Phys Sci*. 1964;280(1382):383-397. doi:10.1098/rspa.1964.0151

106. Edmond de H, Vincent S. *Mass Spectrometry: Principles and Applications*. (John Wiley & Sons L, ed.). Toronto; 2003.

107. Siuzdak G. *The Expading Role of Mass Spectrometry in Biotechnology*. 2nd ed. San Diego, California: MCC Press; 2006.

108. B.A. Mamyrin, V.I. Karataev, D.V. Shmikk VAZ. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *JETP*. 1972;37(1):45. http://www.jetp.ac.ru/cgi-bin/e/index/e/37/1/p45?a=list. Accessed May 11, 2019.

109. Deutsch EW, Mendoza L, Shteynberg D, et al. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*. 2010;10(6):1150-1159. doi:10.1002/pmic.200900375

110. Zhang Y, Fonslow BR, Shan B, Baek M, Yates JR. Protein Analysis by Shotgun / Bottom-up Proteomics. 2013:1-101. doi:10.1021/cr3003533.Protein

111. Loo JA, Udseth HR, Smith RD. Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Anal Biochem*. 1989;179(2):404-412. doi:10.1016/0003-2697(89)90153-X

112. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR, III. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev*. 2013;113(4):2343-2394. doi:10.1021/cr3003533

113. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004;101(26):9528-9533. doi:10.1073/pnas.0402700101

114. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2

115. Matthiesen R, Trelle MB, Højrup P, Bunkenborg J, Jensen ON. VEMS 3.0: Algorithms and Computational Tools for Tandem Mass Spectrometry Based Identification of Post-translational Modifications in Proteins. *J Proteome Res*. 2005;4(6):2338-2347. doi:10.1021/pr050264q

116. Matthiesen R, Bunkenborg J, Stensballe A, Jensen ON, Welinder KG, Bauw G. Database-independent, database-dependent, and extended interpretation of peptide mass spectra in VEMS V2.0. *Proteomics*. 2004;4(9):2583-2593. doi:10.1002/pmic.200300792

117. Matthiesen R, Lundsgaard M, Welinder KG, Bauw G. Interpreting peptide mass spectra by VEMS. *Bioinformatics*. 2003;19(6):792-793. doi:10.1093/bioinformatics/btg059

118. Listgarten J, Emili A. Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry. *Mol Cell Proteomics*. 2005;4(4):419-434. doi:10.1074/mcp.R500005-MCP200

119. Jimmy KE, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994;5(11):976-989. doi:10.1016/1044-0305(94)80016-2

120. Nesvizhskii AI, Aebersold R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discov Today*. 2004;9(4):173-181. doi:10.1016/S1359-6446(03)02978-7

121. Hunt DF, Yates JR, Shabanowitz J, Winston S, Hauer CR. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 1986;83(17):6233-6237. http://www.ncbi.nlm.nih.gov/pubmed/3462691. Accessed January 13, 2019.

122. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*. 1993;3(6):327-332. http://www.ncbi.nlm.nih.gov/pubmed/15335725. Accessed January 13, 2019.

123. Akhtar MN, Southey BR, Andrén PE, Sweedler J V, Rodriguez-Zas SL. Evaluation of database search programs for accurate detection of neuropeptides in tandem mass spectrometry experiments. *J Proteome Res*. 2012;11(12):6044-6055. doi:10.1021/pr3007123

124. Craig R, Beavis RC. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun Mass Spectrom*. 2003;17(20):2310-2316. doi:10.1002/rcm.1198

125. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004;20(9):1466-1467. doi:10.1093/bioinformatics/bth092

126. Eng JK, Fischer B, Grossmann J, MacCoss MJ. A Fast SEQUEST Cross Correlation Algorithm. *J Proteome Res*. 2008;7(10):4598-4602. doi:10.1021/pr800420s

127. Eng JK, Jahan TA, Hoopmann MR. Comet: An open-source MS/MS sequence database search tool. *Proteomics*. 2013;13(1):22-24. doi:10.1002/pmic.201200439

128. Phanstiel DH, Brumbaugh J, Wenger CD, et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat Methods*. 2011;8(10):821-827. doi:10.1038/nmeth.1699

129. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res*. 2011;10(4):1794-1805. doi:10.1021/pr101065j

130. Zhang N, Li X, Ye M, Pan S, Schwikowski B, Aebersold R. ProbIDtree: An automated software program capable of identifying multiple peptides from a single collision-induced dissociation spectrum collected by a tandem mass spectrometer. *Proteomics*. 2005;5(16):4096-4106. doi:10.1002/pmic.200401260

131. Bern M, Finney G, Hoopmann MR, Merrihew G, Toth MJ, MacCoss MJ. Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem*. 2010;82(3):833-841. doi:10.1021/ac901801b

132. Wang J, Pérez-Santiago J, Katz JE, Mallick P, Bandeira N. Peptide Identification from Mixture Tandem Mass Spectra. *Mol Cell Proteomics*. 2010;9(7):1476-1485. doi:10.1074/mcp.M000136-MCP201

133. Houel S, Abernathy R, Renganathan K, Meyer-Arendt K, Ahn NG, Old WM. Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies. *J Proteome Res*. 2010;9(8):4152-4160. doi:10.1021/pr1003856

134. Gupta N, Bandeira N, Keich U, Pevzner PA. Target-Decoy Approach and False Discovery Rate: When Things May Go Wrong. *J Am Soc Mass Spectrom*. 2011;22(7):1111-1120. doi:10.1007/s13361-011-0139-3

135. Shortreed MR, Wenger CD, Frey BL, et al. Global Identification of Protein Post-translational Modifications in a Single-Pass Database Search. *J Proteome Res*. 2015;14(11):4714-4720. doi:10.1021/acs.jproteome.5b00599

136. Qiyao L, Shortreed MR, Wenger CD, et al. Global Post-Translational Modification Discovery. *J Proteome Res*. 2017;16(4):1383-1390. doi:10.1021/acs.jproteome.6b00034

137. Zhang G, Ueberheide BM, Waldemarson S, et al. Protein quantitation using mass spectrometry. *Methods Mol Biol*. 2010;673:211-222. doi:10.1007/978-1-60761-842-3_13

138. Wang M, You J, Bemis KG, Tegeler TJ, Brown DPG. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Briefings Funct Genomics Proteomics*. 2008;7(5):329-339. doi:10.1093/bfgp/eln031

139. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13(9):2513-2526. doi:10.1074/mcp.M113.031591

140. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098

141. Pisanu S, Ghisaura S, Pagnozzi D, et al. The sheep milk fat globule membrane proteome. *J Proteomics*. 2011;74(3):350-358. doi:10.1016/J.JPROT.2010.11.011

142. Zybailov BL, Florens L, Washburn MP. Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Mol Biosyst*. 2007;3(5):354. doi:10.1039/b701483j

143. Arike L, Peil L. Spectral Counting Label-Free Proteomics. In: Humana Press, New York, NY; 2014:213-222. doi:10.1007/978-1-4939-0685-7_14

144. Paoletti AC, Parmely TJ, Tomomori-Sato C, et al. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci U S A*. 2006;103(50):18928-18933. doi:10.1073/pnas.0606379103

145. Dheilly NM, Raftos DA, Haynes PA, Smith LC, Nair S V. Shotgun proteomics of coelomic fluid from the purple sea urchin, Strongylocentrotus purpuratus. *Dev Comp Immunol*. 2013;40(1):35-50. doi:10.1016/J.DCI.2013.01.007

146. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

147. Timmins-Schiffman E, Coffey WD, Hua W, Nunn BL, Dickinson GH, Roberts SB. Shotgun proteomics reveals physiological response to ocean acidification in Crassostrea gigas. *BMC Genomics*. 2014;15(1):951. doi:10.1186/1471-2164-15-951

148. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem*. 2007;389(4):1017-1031. doi:10.1007/s00216-007-1486-6

149. Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat Biotechnol*. 2010;28(1):83-89. doi:10.1038/nbt.1592

150. Ishihama Y, Oda Y, Tabata T, et al. Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein. *Mol Cell Proteomics*. 2005;4(9):1265-1272. doi:10.1074/mcp.M500061-MCP200

151. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces c erevisiae*. *J Proteome Res*. 2006;5(9):2339-2347. doi:10.1021/pr060161n

152. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319.

doi:10.1038/nprot.2016.136

153. Vaudel M, Burkhart JM, Zahedi RP, et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotechnol*. 2015;33(1):22-24. doi:10.1038/nbt.3109

154. Solntsev SK, Shortreed MR, Frey BL, Smith LM. Enhanced Global Post-translational Modification Discovery with MetaMorpheus. *J Proteome Res*. 2018;17(5):1844-1851. doi:10.1021/acs.jproteome.7b00873

155. Partridge JD, Sanguinetti G, Dibden DP, Roberts RE, Poole RK, Green J. Transition of Escherichia coli from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components. *J Biol Chem*. 2007;282(15):11230-11237. doi:10.1074/jbc.M700728200

156. Jones SA, Gibson T, Maltby RC, et al. Anaerobic respiration of Escherichia coli in the mouse intestine. *Infect Immun*. 2011;79(10):4218-4226. doi:10.1128/IAI.05395-11

157. Murashko ON, Lin-Chao S. Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. *Proc Natl Acad Sci U S A*. 2017;114(38):E8025-E8034. doi:10.1073/pnas.1703731114

158. Kim SO, Merchant K, Nudelman R, et al. OxyR: A Molecular Code for Redox-Related Signaling. *Cell*. 2002;109(3):383-396. doi:10.1016/S0092-8674(02)00723-7

159. Zheng M, Aslund F, Storz G. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science*. 1998;279(5357):1718-1721. http://www.ncbi.nlm.nih.gov/pubmed/9497290. Accessed April 10, 2019.

160. Storz G, Tartaglia LA, Ames BN. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science*. 1990;248(4952):189-194. http://www.ncbi.nlm.nih.gov/pubmed/2183352. Accessed April 10, 2019.

161. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol*. 2014;24(10):R453-62. doi:10.1016/j.cub.2014.03.034

162. Lu J, Holmgren A. The thioredoxin antioxidant system. *Free Radic Biol Med*. 2014;66:75-87. doi:10.1016/j.freeradbiomed.2013.07.036

163. Boronat S, Domènech A, Paulo E, et al. Thiol-based H2O2 signalling in microbial systems. *Redox Biol*. 2014;2:395-399. doi:10.1016/j.redox.2014.01.015

164. Ray PD, Huang B-W, Tsuji Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cell Signal*. 2012;24(5):981-990.

doi:10.1016/j.cellsig.2012.01.008

165. D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat Rev Mol Cell Biol*. 2007;8(10):813-824. doi:10.1038/nrm2256

166. Couturier J, Chibani K, Jacquot J-P, Rouhier N. Cysteine-based redox regulation and signaling in plants. *Front Plant Sci*. 2013;4:105. doi:10.3389/fpls.2013.00105

167. Mou Z, Fan W, Dong X. Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell*. 2003;113(7):935-944. http://www.ncbi.nlm.nih.gov/pubmed/12837250. Accessed April 18, 2019.

168. Finkel T, Holbrook NJ. Oxidants, oxidative stress and the biology of ageing. *Nature*. 2000;408(6809):239-247. doi:10.1038/35041687

169. Maaty WS, Selvig K, Ryder S, et al. Proteomic analysis of sulfolobus solfataricus during sulfolobus turreted icosahedral virus infection. *J Proteome Res*. 2012;11(2):1420-1432. doi:10.1021/pr201087v

170. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

171. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-1372. doi:10.1038/nbt.1511

172. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13(9):731-740. doi:10.1038/nmeth.3901

173. Finn TJ, Shewaramani S, Leahy SC, Janssen PH, Moon CD. Dynamics and genetic diversification of Escherichia coli during experimental adaptation to an anaerobic environment. *PeerJ*. 2017;5:e3244. doi:10.7717/peerj.3244

174. von Wulffen J, RecogNice-Team, Sawodny O, Feuer R. Transition of an Anaerobic Escherichia coli Culture to Aerobiosis: Balancing mRNA and Protein Levels in a Demand-Directed Dynamic Flux Balance Analysis. *PLoS One*. 2016;11(7):e0158711. doi:10.1371/journal.pone.0158711

175. Trotter EW, Rolfe MD, Hounslow AM, et al. Reprogramming of Escherichia coli K-12 Metabolism during the Initial Phase of Transition from an Anaerobic to a Micro-Aerobic Environment. Roop RM, ed. *PLoS One*. 2011;6(9):e25501. doi:10.1371/journal.pone.0025501

176. Constantinidou C, Hobman JL, Griffiths L, et al. A Reassessment of the FNR

Regulon and Transcriptomic Analysis of the Effects of Nitrate, Nitrite, NarXL, and NarQP as Escherichia coli K12 Adapts from Aerobic to Anaerobic Growth. *J Biol Chem*. 2006;281(8):4802-4815. doi:10.1074/jbc.M512312200

177. Kramer G, Sprenger RR, Speijer D, et al. Proteome-wide Alterations in Escherichia coli Translation Rates upon Anaerobiosis . *Mol Cell Proteomics*. 2010;9(11):2508-2516. doi:10.1074/mcp.m110.001826

178. Weber A, Kögl SA, Jung K. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli. *J Bacteriol*. 2006;188(20):7165-7175. doi:10.1128/JB.00508-06

179. Kang Y, Weber KD, Qiu Y, Kiley PJ, Blattner FR. Genome-wide expression analysis indicates that FNR of Escherichia coli K-12 regulates a large number of genes of unknown function. *J Bacteriol*. 2005;187(3):1135-1160. doi:10.1128/JB.187.3.1135-1160.2005

180. Salmon KA, Hung S -p., Steffen NR, et al. Global Gene Expression Profiling in Escherichia coli K12: EFFECTS OF OXYGEN AVAILABILITY AND ArcA. *J Biol Chem*. 2005;280(15):15084-15096. doi:10.1074/jbc.M414030200

181. Tyanova S, Cox J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. In: Humana Press, New York, NY; 2018:133-148. doi:10.1007/978-1-4939-7493-1_7

182. Weber J, Kayser A, Rinas U. Metabolic flux analysis of Escherichia coli in glucose-limited continuous culture. II. Dynamic response to famine and feast, activation of the methylglyoxal pathway and oscillatory behaviour. *Microbiology*. 2005;151(3):707-716. doi:10.1099/mic.0.27482-0

183. Epstein MG, Reeves BD, Maaty WS, et al. Enhanced Sensitivity Employing Zwitterionic and pI Balancing Dyes (Z-CyDyes) Optimized for 2D-Gel Electrophoresis Based on Side Chain Modifications of CyDye Fluorophores. New Tools For Use in Proteomics and Diagnostics. *Bioconjug Chem*. 2013;24(9):1552-1561. doi:10.1021/bc4002213

184. Constantinidou C, Hobman JL, Griffiths L, et al. A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as Escherichia coli K12 adapts from aerobic to anaerobic growth. *J Biol Chem*. 2006;281(8):4802-4815. doi:10.1074/jbc.M512312200

185. Tokmina-Lukaszewska M, Movahed N, Lusczek E, Mulier K, Beilman G, Bothner B. Transformation of UPLC-MS Data Overcomes Extreme Variability in Urine Concentration and Metabolite Fold Change. *Curr Metabolomics*. 2014;2(2):78-87. doi:10.2174/2213235X0202141127102223

186. Heinemann J, Mazurie A, Tokmina-Lukaszewska M, Beilman GJ, Bothner B.

Application of support vector machines to metabolomics experiments with limited replicates. *Metabolomics*. 2014;10(6):1121-1128. doi:10.1007/s11306-014-0651-0

187. Kargalioglu Y, Imlay JA. Importance of anaerobic superoxide dismutase synthesis in facilitating outgrowth of Escherichia coli upon entry into an aerobic habitat. *J Bacteriol*. 1994;176(24):7653-7658. doi:10.1128/JB.176.24.7653-7658.1994

188. MCDANIEL LE, BAILEY EG, ZIMMERLI A. EFFECT OF OXYGEN SUPPLY RATES ON GROWTH OF ESCHERICHIA COLI. *Appl Microbiol*. 1965;13(1):109-114. http://www.ncbi.nlm.nih.gov/pubmed/14264837. Accessed April 21, 2019.

189. Nimmo GA, Nimmo HG. The regulatory properties of isocitrate dehydrogenase kinase and isocitrate dehydrogenase phosphatase from Escherichia coli ML308 and the roles of these activities in the control of isocitrate dehydrogenase. *Eur J Biochem*. 1984;141(2):409-414. http://www.ncbi.nlm.nih.gov/pubmed/6329757. Accessed April 22, 2019.

190. LaPorte DC, Walsh K, Koshland DE. The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. *J Biol Chem*. 1984;259(22):14068-14075. http://www.ncbi.nlm.nih.gov/pubmed/6389540. Accessed April 22, 2019.

191. Walsh K, Koshland DE. Determination of flux through the branch point of two metabolic cycles. The tricarboxylic acid cycle and the glyoxylate shunt. *J Biol Chem*. 1984;259(15):9646-9654. http://www.ncbi.nlm.nih.gov/pubmed/6378912. Accessed April 22, 2019.

192. Cortay JC, Bleicher F, Duclos B, et al. Utilization of acetate in Escherichia coli: structural organization and differential expression of the ace operon. *Biochimie*. 1989;71(9-10):1043-1049. http://www.ncbi.nlm.nih.gov/pubmed/2512996. Accessed April 22, 2019.

193. Madigan M, Martinko J. *Brock's Biology of Microorganisms,*. 11th editi. NJ: Pearson Prentice Hall; 2006. http://enacademic.com/dic.nsf/enwiki/2639682/Mixed. Accessed April 22, 2019.

194. Ku SY, Yip P, Howell PL. Structure of Escherichia coli tryptophanase. *Acta Crystallogr D Biol Crystallogr*. 2006;62(Pt 7):814-823. doi:10.1107/S0907444906019895

195. Phillips RS, Demidkina T V., Faleev NG. Structure and mechanism of tryptophan indole-lyase and tyrosine phenol-lyase. *Biochim Biophys Acta - Proteins Proteomics*. 2003;1647(1-2):167-172. doi:10.1016/S1570-9639(03)00089-X

196. Nihira T, Yasuda T, Kakizono T, et al. Functional role of cysteinyl residues in tryptophanase. *Eur J Biochem*. 1985;149(1):129-133. http://www.ncbi.nlm.nih.gov/pubmed/3888623. Accessed April 13, 2019.

197. Liu M, Hou J, Huang L, et al. Site-specific proteomics approach for study protein S-nitrosylation. *Anal Chem*. 2010;82(17):7160-7168. doi:10.1021/ac100569d

198. Forrester MT, Thompson JW, Foster MW, Nogueira L, Moseley MA, Stamler JS. Proteomic analysis of S-nitrosylation and denitrosylation by resin-assisted capture. *Nat Biotechnol*. 2009;27(6):557-559. doi:10.1038/nbt.1545

199. Lee J, Maeda T, Hong SH, Wood TK. Reconfiguring the Quorum-Sensing Regulator SdiA of Escherichia coli To Control Biofilm Formation via Indole and N-Acylhomoserine Lactones. *Appl Environ Microbiol*. 2009;75(6):1703-1716. doi:10.1128/AEM.02081-08

200. Lee Y, Yeom J, Kim J, Jung J, Jeon CO, Park W. Phenotypic and physiological alterations by heterologous acylhomoserine lactone synthase expression in Pseudomonas putida. *Microbiology*. 2010;156(12):3762-3772. doi:10.1099/mic.0.041095-0

201. Chu W, Zere TR, Weber MM, et al. Indole Production Promotes Escherichia coli Mixed-Culture Growth with Pseudomonas aeruginosa by Inhibiting Quorum Signaling. *Appl Environ Microbiol*. 2012;78(2):411-419. doi:10.1128/AEM.06396-11

202. Molina-Santiago C, Daddaoua A, Fillet S, Duque E, Ramos J-L. Interspecies signalling: *Pseudomonas putida* efflux pump TtgGHI is activated by indole to increase antibiotic resistance. *Environ Microbiol*. 2014;16(5):1267-1281. doi:10.1111/1462-2920.12368

APPENDICES

APPENDIX I

SUPPLEMENTARY FIGURES AND TABLES

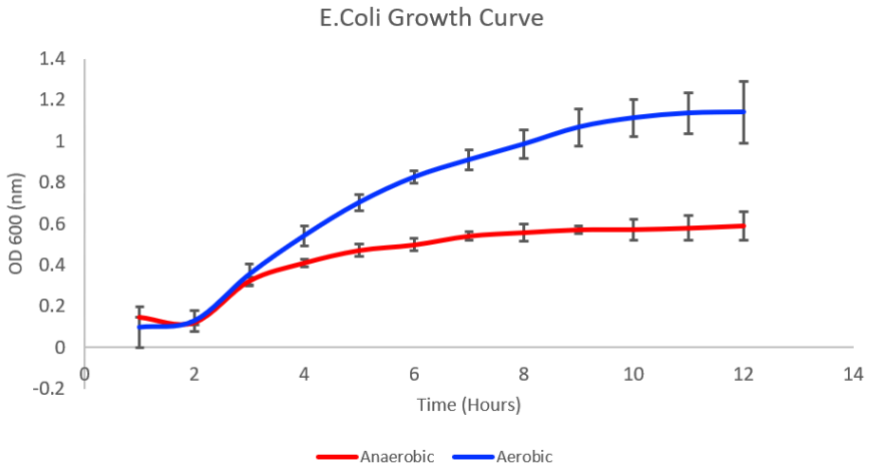


Figure 3.1S: Growth curve of the E. coli strain MG1655 (K12). The growth curve was constructed by reading the optical density at 600 nm hourly for 12 hours. The blue trace is for the aerobic and the red trace is for the anaerobic experimental groups. All readings were taken in triplicates.

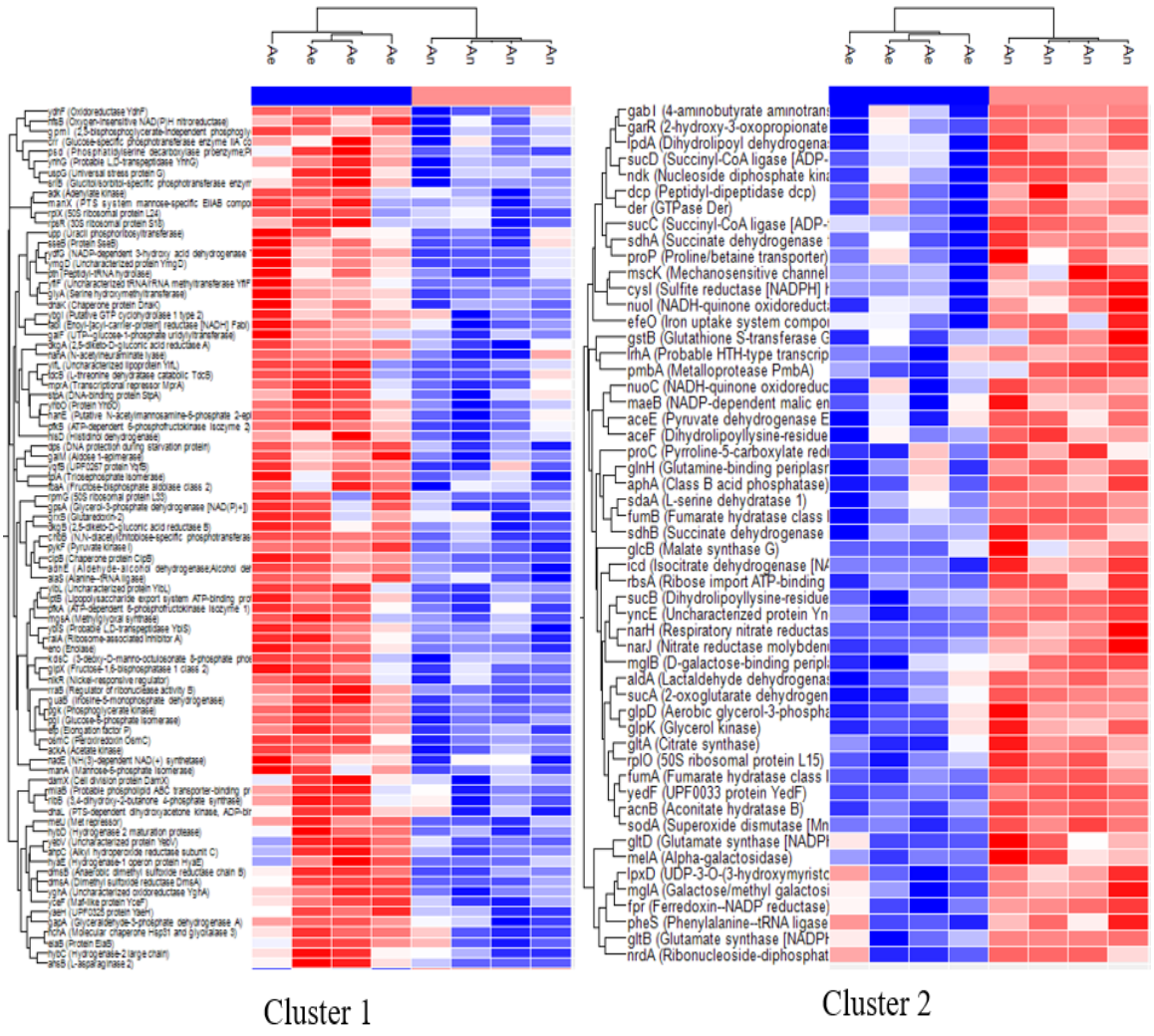


Figure 3.2S Hierarchical Clustering Analysis for the Differentially Regulated Proteins in the two Experimental groups, Aerobic (Ae) and Anaerobic (An). The log2 iBAQ values were Z-scored and mapped to the row-scaled expression values (+1.5 to -1.5). Cluster 1 is for the proteins up-regulated in the Ae (85 proteins). Cluster 2 is for the proteins up-regulated in the Ae (53 proteins).

Table 3.1S The Identified Proteins in the Two Experimental Groups. The Identification Process Was Carried out Using MaxQuant Platform. Gene names are Derived from UniProt Database and used in the Functional Annotation Analysis.

# of Peptides	Mol. weight [kDa]	Score	MS/MS count	Protein names	Gene names
6	19.419	73.832	51	Aerobic respiration control protein ArcA	ArcA
8	116.48	35.833	11	Beta-galactosidase	lacZ
1	17.219	4.8209	4	SsrA-binding protein	smpB
22	51.481	233.49	257	6-phosphogluconate dehydrogenase, decarboxylating	gnd
12	65.971	82.881	33	Fumarate reductase flavoprotein subunit	frdA
6	47.358	21.964	7	NADH dehydrogenase	ndh
8	23.097	34.903	20	Superoxide dismutase [Mn]	sodA
10	85.774	61.137	28	Ribonucleoside-diphosphate reductase 1 subunit alpha	nrdA
13	90.521	50.651	33	Maltodextrin phosphorylase	malP
9	43.573	33.663	28	Aspartate aminotransferase	aspC
6	67.971	11.012	9	PTS system mannitol-specific EIICBA component;Mannitol permease IIC component;Mannitol-specific phosphotransferase enzyme IIB component;Mannitol-specific phosphotransferase enzyme IIA component	mtlA
10	70.263	19.497	21	RNA polymerase sigma factor RpoD	rpoD
5	103.12	4.1167	6	DNA polymerase I	polA
16	36.85	172.7	106	L-asparaginase 2	ansB
6	99.061	16.134	12	Phosphoenolpyruvate carboxylase	ppc
5	47.9	12.016	10	D-serine dehydratase	dsdA
11	42.849	45.86	43	Mannose-6-phosphate isomerase	manA
7	37.437	51.733	39	Tryptophan--tRNA ligase	trpS
29	104.3	238.23	133	Isoleucine--tRNA ligase	ileS
29	96.031	235.23	171	Alanine--tRNA ligase	alaS
21	76.254	296.44	119	Methionine--tRNA ligase	metG
19	76.812	188.43	72	Glycine--tRNA ligase beta subunit	glyS
15	63.477	50.23	66	Glutamine--tRNA ligase	glnS

12	117.84	35.777	27	Carbamoyl-phosphate synthase large chain	carB
7	15.703	177.7	321	30S ribosomal protein S6;30S ribosomal protein S6, fully modified isoform;30S ribosomal protein S6, non-modified isoform	rpsF
10	20.019	216.22	175	30S ribosomal protein S7	rpsG
6	14.98	88.607	106	50S ribosomal protein L15	rplO
4	35.54	7.5816	12	L-arabinose-binding periplasmic protein	araF
37	30.95	323.31	429	D-ribose-binding periplasmic protein	rbsB
9	53.74	20.284	12	Outer membrane protein TolC	tolC
1	15.807	3.2127	7	Protein MioC	mioC
9	31.943	48.266	51	Maltose operon periplasmic protein	malM
6	28.756	18.837	37	4-hydroxy-tetrahydrodipicolinate reductase	dapB
12	58.679	84.955	44	GMP synthase [glutamine-hydrolyzing]	guaA
4	18.111	79.723	22	Type-1 fimbrial protein, A chain	fimA
2	47.008	3.0648	3	Esterase FrsA	frsA
5	35.56	10.622	10	Glutathione synthetase	gshB
5	43.537	6.8586	6	Aromatic-amino-acid aminotransferase	tyrB
18	53.815	137.4	97	Glutamate--tRNA ligase	gltX
2	27.614	1.9668	4	3-deoxy-manno-octulosonate cytidylyltransferase	kdsB
7	15.292	137.1	73	D-ribose pyranase	rbsD
7	55.041	66.399	26	Ribose import ATP-binding protein RbsA	rbsA
11	38.827	78.422	66	Dihydroorotase	pyrC
21	77.1	146.91	123	Polyribonucleotide nucleotidyltransferase	pnp
5	27.466	6.0453	9	3-dehydroquinate dehydratase	aroD
7	41.367	20.638	14	Erythronate-4-phosphate dehydrogenase	pdxB
5	13.304	44.991	20	Glucitol/sorbitol-specific phosphotransferase enzyme IIA component	srlB
4	65.531	5.0186	5	Dihydroxy-acid dehydratase	ilvD
5	68.446	4.4672	4	Beta-glucuronidase	uidA
3	64.612	5.4761	8	D-lactate dehydrogenase	dld
4	20.469	13.673	20	Vitamin B12 transport periplasmic protein BtuE	btuE
9	97.349	18.853	26	DNA topoisomerase 1	topA
8	48.772	122.8	24	Glutathione reductase	gor
6	50.657	59.931	12	Alpha-galactosidase	melA

28	66.095	323.31	403	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	aceF
4	16.155	8.9329	12	Deoxyuridine 5-triphosphate nucleotidohydrolase	dut
6	33.851	66.951	34	Porphobilinogen deaminase	hemC
4	39.36	4.8799	7	Histidinol-phosphate aminotransferase	hisC
5	40.278	13.228	14	Histidine biosynthesis bifunctional protein HisB;Histidinol-phosphatase;Imidazoleglycerol-phosphate dehydratase	hisB
3	46.11	9.3959	6	Histidinol dehydrogenase	hisD
2	30.42	5.6489	6	Ribosomal RNA small subunit methyltransferase A	rsmA
11	40.368	140.1	46	Outer membrane protein C	ompC
11	32.456	203.92	48	ATP-dependent 6-phosphofructokinase isozyme 2	pfkB
4	44.63	33.493	14	Gamma-glutamyl phosphate reductase	proA
12	41.25	194.83	55	Peptide chain release factor 2	prfB
3	26.77	20.257	14	Succinate dehydrogenase iron-sulfur subunit	sdhB
9	60.823	38.749	31	Protein UshA;UDP-sugar hydrolase;5-nucleotidase	ushA
3	47.056	48.851	10	Periplasmic AppA protein;Phosphoanhydride phosphohydrolase;4-phytase	appA
21	108.19	63.218	52	Valine--tRNA ligase	valS
16	87.377	65.779	41	Phenylalanine--tRNA ligase beta subunit	pheT
13	47.207	143.93	51	Thymidine phosphorylase	deoA
2	84.336	1.1836	4	1,4-alpha-glucan branching enzyme GlgB	glgB
21	97.233	145.97	108	Leucine--tRNA ligase	leuS
1	32.839	3.2999	6	D-alanine--D-alanine ligase B	ddlB
6	37.239	30.847	38	L-threonine 3-dehydrogenase	tdh
5	36.854	39.826	15	Phosphoribosylformylglycinamidine cyclo-ligase	purM
2	23.238	23.138	13	Phosphoribosylglycinamide formyltransferase	purN
4	50.31	20.664	21	Glycerol-3-phosphate transporter	glpT
14	45.756	106.21	111	Isocitrate dehydrogenase [NADP]	icd
6	26.35	8.7386	6	Orotidine 5-phosphate decarboxylase	pyrF
7	36.831	41.375	32	Phenylalanine--tRNA ligase alpha subunit	pheS
9	70.831	22.871	17	2,3-cyclic-nucleotide 2-phosphodiesterase/3-nucleotidase	cpdB
3	36.364	8.8103	6	USG-1 protein	usg
8	41.1	43.027	27	Chaperone protein DnaJ	dnaJ

24	63.561	219.85	162	Phosphoenolpyruvate-protein phosphotransferase	ptsI
3	30.969	7.8326	7	Exodeoxyribonuclease III	xthA
7	42.962	20.34	19	Putative uroporphyrinogen-III C-methyltransferase	hemX
7	37.265	36.554	25	UDP-glucose 4-epimerase	galE
2	39.645	3.7903	4	Galactose-1-phosphate uridylyltransferase	galT
12	140.49	31.592	22	Respiratory nitrate reductase 1 alpha chain	narG
5	68.346	7.7042	14	PTS system N-acetylglucosamine-specific EIICBA component;N-acetylglucosamine permease IIC component;N-acetylglucosamine-specific phosphotransferase enzyme IIB component;N-acetylglucosamine-specific phosphotransferase enzyme IIA component	nagE
7	21.798	174.93	155	Protein GrpE	grpE
31	85.356	323.31	353	Formate acetyltransferase 1	pflB
18	40.843	323.31	329	Glycerophosphoryl diester phosphodiesterase	glpQ
7	41.139	59.638	27	Mannitol-1-phosphate 5-dehydrogenase	mtlD
7	143.81	14.73	10	Bifunctional protein PutA;Proline dehydrogenase;Delta-1-pyrroline-5-carboxylate dehydrogenase	putA
21	163.3	92.901	53	Glutamate synthase [NADPH] large chain	gltB
10	52.015	29.835	19	Glutamate synthase [NADPH] small chain	gltD
2	13.786	15.634	16	Uncharacterized protein YidB	yidB
16	43.29	157.69	122	Acetate kinase	ackA
4	8.6394	38.026	107	Acyl carrier protein	acpP
7	45.089	43.75	20	Cysteine desulfurase IscS	iscS
4	31.479	10.824	11	Endonuclease 4	nfo
3	16.994	33.725	18	Arginine repressor	argR
9	19.538	89.568	37	Shikimate kinase 1	aroK
5	15.068	130.11	38	ATP synthase epsilon chain	atpC
8	41.431	19.357	17	Carbamoyl-phosphate synthase small chain	carA
28	56.23	323.31	284	Glycerol kinase	glpK
31	57.328	323.31	1247	60 kDa chaperonin	groL
5	10.387	64.701	474	10 kDa chaperonin	groS
3	23.186	120.83	60	ATP-dependent Clp protease proteolytic subunit	clpP
12	46.355	72.08	51	ATP-dependent Clp protease ATP-binding subunit ClpX	clpX

17	49.593	247.48	109	ATP-dependent protease ATPase subunit HslU	hslU
8	24.746	35.931	13	Cytidylate kinase	cmk
4	47.607	7.0081	6	D-amino acid dehydrogenase	dadA
5	39.315	25.533	18	D-alanine--D-alanine ligase A	ddlA
6	19.328	16.338	11	Peptide deformylase	def
19	44.369	111.45	106	Phosphopentomutase	deoB
10	27.733	108	42	Deoxyribose-phosphate aldolase	deoC
7	31.27	71.774	50	4-hydroxy-tetrahydrodipicolinate synthase	dapA
5	32.593	25.421	22	N-acetylneuraminate lyase	nanA
7	54.808	7.5461	9	Altronate oxidoreductase	uxaB
3	20.137	21.015	8	Co-chaperone protein HscB	hscB
2	15.95	1.584	3	D-aminoacyl-tRNA deacylase	dtd
27	77.58	323.31	451	Elongation factor G	fusA
5	20.591	23.26	55	Elongation factor P	efp
6	21.532	20.462	18	Elongation factor P-like protein	yeiP
24	30.423	323.31	236	Elongation factor Ts	tsf
3	55.035	2.9679	6	GTPase Der	der
8	23.561	22.597	26	Probable GTP-binding protein EngB	engB
17	45.654	323.31	649	Enolase	eno
6	18.969	41.151	51	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase	fabA
7	33.515	37.209	29	3-oxoacyl-[acyl-carrier-protein] synthase 3	fabH
2	11.24	66.637	17	DNA-binding protein Fis	fis
4	36.361	15.66	8	Glycerol-3-phosphate dehydrogenase [NAD(P)+]	gpsA
17	61.529	53.932	43	Glucose-6-phosphate isomerase	pgi
6	41.442	31.76	28	Galactokinase	galK
5	24.83	23.748	22	GTP cyclohydrolase 1	folE
3	23.431	2.4938	3	Ribosomal RNA small subunit methyltransferase G	rsmG
2	48.697	6.9283	7	Glucose-1-phosphate adenylyltransferase	glgC
3	34.723	6.6596	12	Glucokinase	glk
11	17.641	74.452	27	Transcription elongation factor GreA	greA
6	11.354	92.035	47	Integration host factor subunit alpha	ihfA
2	10.651	28.35	17	Integration host factor subunit beta	ihfB
9	32.534	110.29	58	33 kDa chaperonin	hslO

46	69.114	323.31	1284	Chaperone protein DnaK	dnaK
8	65.652	31.663	28	Chaperone protein HscA	hscA
29	71.422	323.31	186	Chaperone protein HtpG	htpG
2	15.094	5.1256	5	Nickel-responsive regulator	nikR
36	97.349	220.09	169	Translation initiation factor IF-2	infB
8	20.564	170.89	74	Translation initiation factor IF-3;Translation initiation factor IF-3, N-terminally processed;Translation initiation factor IF-3S	infC
3	33.669	4.4367	3	2-keto-3-deoxygluconate permease	kdgT
6	30.832	19.37	20	2-dehydro-3-deoxyphosphooctonate aldolase	kdsA
9	34.218	76.741	56	Ribose-phosphate pyrophosphokinase	prs
3	23.783	11.937	9	Thymidylate kinase	tmk
1	28.08	3.9948	4	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	lpxA
2	21.691	1.2764	2	Maf-like protein YceF	yceF
3	16.918	91.845	21	Methylglyoxal synthase	mgsA
6	10.235	27.587	29	Cell division topological specificity factor	minE
1	17.467	5.1917	3	Cyclic pyranopterin monophosphate synthase accessory protein	moaC
3	23.315	11.278	5	Peptide methionine sulfoxide reductase MsrA	msrA
4	15.451	81.463	16	Peptide methionine sulfoxide reductase MsrB	msrB
5	29.774	19.268	15	Glucosamine-6-phosphate deaminase	nagB
7	24.073	22.713	21	Putative N-acetylmannosamine-6-phosphate 2-epimerase	nanE
5	15.463	116.5	26	Nucleoside diphosphate kinase	ndk
3	15.689	8.6192	7	N utilization substance protein B	nusB
5	20.815	14.812	12	Oligoribonuclease	orn
9	34.427	31.397	20	Aspartate carbamoyltransferase catalytic chain	pyrB
6	26.384	66.315	26	Pyridoxine 5-phosphate synthase	pdxJ
5	34.842	76.653	17	ATP-dependent 6-phosphofructokinase isozyme 1	pfkA
25	41.118	323.31	771	Phosphoglycerate kinase	pgk
8	19.703	124.78	85	Inorganic pyrophosphatase	ppa
3	39.056	1.1508	3	Glutamate 5-kinase	proB
5	19.093	70.116	13	ATP-dependent protease subunit HslV	hslV
6	22.357	16.595	25	LexA repressor	lexA

7	24.57	47.351	27	Peptidase E	pepE
4	21.082	15.291	3	Peptidyl-tRNA hydrolase	pth
19	47.344	175.69	112	Adenylosuccinate synthetase	purA
9	26.995	84.446	57	Phosphoribosylaminoimidazole-succinocarboxamide synthase	purC
7	36.774	24.828	19	Dihydroorotate dehydrogenase (quinone)	pyrD
2	23.567	1.9756	6	Orotate phosphoribosyltransferase	pyrE
9	60.373	27.803	17	CTP synthase	pyrG
3	25.97	20.557	16	Uridylate kinase	pyrH
4	17.12	40.001	30	Aspartate carbamoyltransferase regulatory chain	pyrI
6	30.384	39.227	15	S-adenosylmethionine decarboxylase proenzyme;S-adenosylmethionine decarboxylase beta chain;S-adenosylmethionine decarboxylase alpha chain	speD
2	39.43	3.1488	3	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	queA
5	15.154	43.039	42	Ribosome-binding factor A	rbfA
7	37.973	26.698	28	Protein RecA	recA
4	40.517	35.986	9	Peptide chain release factor 1	prfA
3	59.573	22.428	6	Peptide chain release factor 3	prfC
5	21.836	14.842	29	GTP cyclohydrolase-2	ribA
5	23.353	58.98	15	3,4-dihydroxy-2-butanone 4-phosphate synthase	ribB
9	17.711	136.81	198	50S ribosomal protein L10	rplJ
5	14.875	174.3	72	50S ribosomal protein L11	rplK
5	12.295	65.264	861	50S ribosomal protein L7/L12	rplL
7	13.133	50.5	86	50S ribosomal protein L19	rplS
18	24.729	321.43	403	50S ribosomal protein L1	rplA
3	13.497	6.5169	28	50S ribosomal protein L20	rplT
3	9.1244	6.2082	4	50S ribosomal protein L27	rpmA
3	9.0064	27.834	33	50S ribosomal protein L28	rpmB
2	7.2734	63.394	72	50S ribosomal protein L29	rpmC
3	7.871	57.461	59	50S ribosomal protein L31	rpmE
2	6.4463	51.047	8	50S ribosomal protein L32	rpmF
2	6.3715	1.9318	6	50S ribosomal protein L33	rpmG
1	7.2888	5.4775	4	50S ribosomal protein L35	rpmI
9	15.769	139.37	204	50S ribosomal protein L9	rplI

6	11.735	56.468	97	30S ribosomal protein S10	rpsJ
4	13.845	112.06	37	30S ribosomal protein S11	rpsK
2	13.737	12.707	21	30S ribosomal protein S12	rpsL
8	13.099	111.28	101	30S ribosomal protein S13	rpsM
6	9.1904	62.833	105	30S ribosomal protein S16	rpsP
3	8.9863	43.144	47	30S ribosomal protein S18	rpsR
5	10.43	23.202	39	30S ribosomal protein S19	rpsS
2	9.6843	43.489	15	30S ribosomal protein S20	rpsT
10	26.743	147.23	126	30S ribosomal protein S2	rpsB
8	25.983	112.08	129	30S ribosomal protein S3	rpsC
10	23.469	186.96	178	30S ribosomal protein S4	rpsD
7	17.603	235.11	199	30S ribosomal protein S5	rpsE
9	14.126	62.333	84	30S ribosomal protein S8	rpsH
5	14.856	40.323	50	30S ribosomal protein S9	rpsI
5	20.605	6.9998	18	Ribosome maturation factor RimM	rimM
5	22.86	46.642	39	Ribose-5-phosphate isomerase A	rpiA
19	36.511	198.07	224	DNA-directed RNA polymerase subunit alpha	rpoA
7	10.236	262.24	57	DNA-directed RNA polymerase subunit omega	rpoZ
10	20.638	204.97	143	Ribosome-recycling factor	frr
9	41.951	53.51	50	S-adenosylmethionine synthase	metK
5	50.607	28.63	8	L-seryl-tRNA(Sec) selenium transferase	selA
17	45.316	323.31	303	Serine hydroxymethyltransferase	glyA
11	41.392	82.643	82	Succinyl-CoA ligase [ADP-forming] subunit beta	sucC
28	48.192	323.31	357	Trigger factor	tig
28	52.773	323.31	2117	Tryptophanase	tnaA
10	45.955	72.53	56	Protein TolB	tolB
12	26.972	170.99	180	Triosephosphate isomerase	tpiA
7	17.835	219.79	101	Thiol peroxidase	tpx
8	35.658	42.31	48	Transaldolase A	talA
16	35.219	172.58	134	Transaldolase B	talB
1	28.422	2.6803	5	tRNA (guanine-N(1)-)-methyltransferase	trmD
4	28.073	12.959	12	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	ubiE

2	9.0945	9.2343	5	Sulfurtransferase TusA	tusA
1	17.525	13.801	4	Endoribonuclease YbeY	ybeY
7	26.422	46.312	35	Probable transcriptional regulatory protein YebC	yebC
4	25.867	44.594	27	Probable transcriptional regulatory protein YeeN	yeeN
3	16.651	2.7791	12	Ribosome maturation factor RimP	rimP
6	12.015	68.063	43	Nucleoid-associated protein YbaB	ybaB
1	11.437	7.9261	3	UPF0145 protein YbjQ	ybjQ
1	21.229	3.5145	4	UPF0149 protein YgfB	ygfB
3	17.229	14.549	9	Transcriptional repressor NrdR	nrdR
4	21.226	16.94	9	UPF0227 protein YcfP	ycfP
1	13.941	21.734	9	UPF0231 protein YacL	yacL
10	18.344	144.15	75	UPF0234 protein YajQ	yajQ
9	22.533	144.29	94	Uracil phosphoribosyltransferase	upp
9	53.987	50.067	17	Uronate isomerase	uxaC
8	20.845	177.34	101	NAD(P)H dehydrogenase (quinone)	wrbA
3	8.9518	12.003	14	Exodeoxyribonuclease 7 small subunit	xseB
6	19.059	24.947	17	Der GTPase-activating protein YihI	yihI
2	29.585	5.8999	4	UPF0246 protein YaaA	yaaA
3	19.455	3.8828	6	Primosomal protein 1	dnaT
4	9.8272	25.658	40	UPF0250 protein YbeD	ybeD
2	35.934	1.5444	4	Phosphatidylserine decarboxylase proenzyme;Phosphatidylserine decarboxylase alpha chain;Phosphatidylserine decarboxylase beta chain	psd
2	7.2141	13.393	8	UPF0253 protein YaeP	yaeP
16	48.413	96.307	63	Serine--tRNA ligase	serS
8	52.57	41.456	40	Asparagine--tRNA ligase	asnS
18	74.013	86.547	78	Threonine--tRNA ligase	thrS
6	12.778	99.231	23	UPF0265 protein YeeX	yeeX
2	17.693	9.1362	4	Macrodomain Ter protein	matP
22	57.603	193.53	119	Lysine--tRNA ligase	lysS
22	57.826	116.6	122	Lysine--tRNA ligase, heat inducible	lysU
7	10.952	34.09	25	Probable Fe(2+)-trafficking protein	yggX
1	8.2142	3.6781	2	Protein SlyX	slyX
3	31.877	6.6374	6	Ribosomal protein L11 methyltransferase	prmA

24	155.16	225.77	99	DNA-directed RNA polymerase subunit beta	rpoC
5	12.141	12.463	8	Met repressor	metJ
33	150.63	183.46	107	DNA-directed RNA polymerase subunit beta	rpoB
5	16.353	19.571	14	Transcriptional regulator SlyA	slyA
4	20.686	12.115	13	UPF0301 protein YqgE	yqgE
6	19.536	41.587	26	UPF0304 protein YfbU	yfbU
4	21.359	12.896	18	UPF0307 protein YjgA	yjgA
2	20.912	15.5	5	Protein YceI	yceI
4	22.732	19.164	12	Alpha-D-glucose-1-phosphate phosphatase YihX	yihX
5	29.721	24.429	14	Sugar phosphatase YidA	yidA
5	36.842	19.994	6	Outer membrane protein assembly factor BamC	bamC
4	15.601	9.8814	7	Outer membrane lipoprotein SlyB	slyB
3	27.831	33.061	11	MltA-interacting protein	mipA
14	37.2	154.9	150	Outer membrane protein A	ompA
3	18.824	5.7491	10	Peptidoglycan-associated lipoprotein	pal
2	22.928	2.635	6	Outer membrane protein W	ompW
3	33.589	12.078	10	Nucleoside-specific channel-forming protein tsx	tsx
10	90.552	29.581	31	Outer membrane protein assembly factor BamA	bamA
7	42.613	122.25	71	3-oxoacyl-[acyl-carrier-protein] synthase 1	fabB
11	22.284	144.11	173	KHG/KDPG aldolase;4-hydroxy-2-oxoglutarate aldolase;2-dehydro-3-deoxy-phosphogluconate aldolase	eda
2	7.9689	13.967	6	Cold shock-like protein CspD	cspD
3	7.4634	62.611	48	Cold shock-like protein CspE	cspE
14	38.109	84.101	59	Fructose-bisphosphate aldolase class 1	fbaB
3	36.833	9.9981	9	Fructose-1,6-bisphosphatase class 1	fbp
13	44.108	56.063	38	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC
5	19.424	15.654	17	Bacterial non-heme ferritin	ftnA
25	40.323	235.83	172	Cell division protein FtsZ	ftsZ
4	16.795	44.417	41	Ferric uptake regulation protein	fur
20	35.532	191.68	553	Glyceraldehyde-3-phosphate dehydrogenase A	gapA
9	58.958	60.49	25	Anaerobic glycerol-3-phosphate dehydrogenase subunit A	glpA
5	38.19	58.158	30	Aldose 1-epimerase	galM

18	51.903	166.59	135	Glutamine synthetase	glnA
5	35.852	7.4476	6	Fructose-1,6-bisphosphatase 1 class 2	glpX
4	22.868	5.9084	4	Glutathione S-transferase GstA	gstA
11	29.892	267.86	96	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	dapD
2	28.281	3.8971	6	Transcriptional regulator ModE	modE
5	18.673	4.866	5	G/U mismatch-specific DNA glycosylase	mug
2	43.909	3.5052	3	Cyclopropane-fatty-acyl-phospholipid synthase	cfa
2	12.284	14.749	3	Nitrite reductase (NADH) small subunit	nirD
5	55.364	10.373	7	Ribonuclease G	rng
7	32.29	58.912	49	Ribokinase	rbsK
4	43.111	13.012	4	P-protein;Chorismate mutase;Prephenate dehydratase	pheA
9	39.038	68.263	46	PhoH-like protein	ybeZ
6	20.853	154.3	70	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD	slyD
8	22.216	152.56	101	FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase	fklB
3	10.232	6.1829	12	Peptidyl-prolyl cis-trans isomerase C	ppiC
3	28.145	6.8572	6	Pyrroline-5-carboxylate reductase	proC
12	87.437	54.48	39	Lon protease	lon
4	20.115	82.316	24	Hypoxanthine phosphoribosyltransferase	hpt
14	77.171	92.961	61	Phosphate acetyltransferase	pta
4	28.204	27.718	12	Pyruvate formate-lyase 1-activating enzyme	pflA
18	50.688	213.21	249	Dihydrolipoyl dehydrogenase	lpdA
10	34.623	101.98	40	Thioredoxin reductase	trxB
7	70.545	23.797	15	ATP-dependent RNA helicase DeaD	deaD
4	33.322	18.814	11	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	accD
35	96.126	323.31	219	Aldehyde-alcohol dehydrogenase;Alcohol dehydrogenase;Acetaldehyde dehydrogenase [acetylating];Pyruvate-formate-lyase deactivase	adhE
7	40.017	41.348	17	Aspartate-semialdehyde dehydrogenase	asd
5	12.331	47.091	46	2Fe-2S ferredoxin	fdx
6	40.513	22.844	23	Lactaldehyde reductase	fucO
7	38.712	87.692	26	Glycerol dehydrogenase	gldA
9	59.857	30.867	17	Uncharacterized ABC transporter ATP-binding protein YbiT	ybiT

2	26.8	5.2637	5	Lipopolysaccharide export system ATP-binding protein LptB	lptB
12	62.442	30.403	31	Energy-dependent translational throttle protein EttA	ettA
6	20.245	23.918	22	Protein YrdA	yrdA
7	36.952	63.813	47	Rod shape-determining protein MreB	mreB
4	7.4032	17.317	16	Cold shock protein CspA	cspA
4	7.4023	149.51	188	Cold shock-like protein CspC	cspC
5	12.425	15.804	9	Nitrogen regulatory protein P-II 1	glnB
4	9.1193	172.63	83	Phosphocarrier protein HPr	ptsH
8	16.018	122.38	124	50S ribosomal protein L13	rplM
7	27.353	38.061	50	Transcriptional regulatory protein OmpR	ompR
4	11.806	86.375	153	Thioredoxin-1	trxA
4	8.6389	74.472	29	UPF0033 protein YedF	yedF
6	25.865	44.85	25	Ribosomal small subunit pseudouridine synthase A	rsuA
2	55.603	16.256	6	3-octaprenyl-4-hydroxybenzoate carboxy-lyase	ubiD
5	32.829	16.537	27	UTP--glucose-1-phosphate uridylyltransferase	galF
2	16.293	7.4281	11	Universal stress protein D	uspD
7	35.706	88.693	37	Universal stress protein E	uspE
1	11.556	3.2898	7	Iron-binding protein IscA	iscA
1	48.878	3.2695	5	Threonine/serine transporter TdcC	tdcC
5	56.415	13.774	9	Galactose/methyl galactoside import ATP-binding protein MglA	mglA
9	70.707	33.733	29	ATP-dependent zinc metalloprotease FtsH	ftsH
12	43.045	83.366	70	3-oxoacyl-[acyl-carrier-protein] synthase 2	fabF
7	32.417	110.6	83	Malonyl CoA-acyl carrier protein transacylase	fabD
2	22.752	4.7257	5	Probable anaerobic dimethyl sulfoxide reductase chain YnfG	ynfG
4	36.003	6.054	18	Hydrogenase-2 operon protein HybA	hybA
3	31.565	10.311	8	Hydrogenase isoenzymes nickel incorporation protein HypB	hypB
4	10.318	16.359	7	Transcriptional repressor FrmR	frmR
1	13.364	5.9258	8	Uncharacterized protein YajD	yajD
4	11.582	31.181	10	Ribosomal silencing factor RsfS	rsfS
6	18.797	75.583	26	Uncharacterized protein YbeL	ybeL
3	12.872	10.195	14	Uncharacterized protein YbgS	ybgS

5	33.325	45.692	40	Probable L,D-transpeptidase YbiS	ybiS
2	17.666	1.9548	2	Uncharacterized protein YbjN	ybjN
3	6.8549	12.468	10	UPF0434 protein YcaR	ycaR
5	8.5244	45.838	41	Uncharacterized protein YccJ	yccJ
5	41.137	16.462	9	Iron uptake system component EfeO	efeO
3	19.315	7.1166	5	Uncharacterized protein YceD	yceD
3	12.414	5.1887	4	Protein YcgL	ycgL
5	11.852	59.081	25	Uncharacterized protein YmgD	ymgD
5	12.693	59.991	45	Protein YchN	ychN
2	9.3855	123.38	26	Protein YciN	yciN
15	39.147	249.31	212	Fructose-bisphosphate aldolase class 2	fbaA
5	43.117	52.097	12	2-amino-3-ketobutyrate coenzyme A ligase	kbl
11	51.542	78.092	47	Adenylosuccinate lyase	purB
3	38.009	6.0129	9	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive	aroG
4	17.264	21.836	27	ATP synthase subunit b	atpF
6	19.332	46.031	45	ATP synthase subunit delta	atpH
10	31.577	111.7	56	ATP synthase gamma chain	atpG
21	55.221	323.31	350	ATP synthase subunit alpha	atpA
25	50.325	323.31	210	ATP synthase subunit beta	atpD
4	37.649	17.356	12	Modulator of FtsH protease HflC	hflC
9	45.544	70.63	34	Modulator of FtsH protease HflK	hflK
3	18.495	8.1844	7	Bacterioferritin	bfr
7	35.241	34.38	38	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	accA
4	16.687	13.856	19	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	accB
3	11.994	11.463	18	Protein BolA	bolA
4	31.54	14.806	7	Cytidine deaminase	cdd
8	48.014	67.216	25	Citrate synthase	gltA
9	84.206	25.272	25	ATP-dependent Clp protease ATP-binding subunit ClpA	clpA
4	58.204	13.249	20	Cytochrome bd-I ubiquinol oxidase subunit 1	cydA
14	34.489	73.606	89	Cysteine synthase A	cysK
13	25.95	122.66	53	Purine nucleoside phosphorylase DeoD-type	deoD

1	43.438	8.1842	5	Coenzyme A biosynthesis bifunctional protein CoaBC;Phosphopantothenoylcysteine decarboxylase;Phosphopantothenate--cysteine ligase	coaBC
9	30.427	87.267	38	2-hydroxy-3-oxopropionate reductase	garR
2	17.999	8.5012	5	Dihydrofolate reductase	folA
5	17.528	81.144	43	RNA polymerase-binding transcription factor DksA	dksA
2	8.8462	3.9086	3	DNA polymerase III subunit theta	holE
13	18.695	163.53	63	DNA protection during starvation protein	dps
6	31.633	18.154	10	1,4-dihydroxy-2-naphthoyl-CoA synthase	menB
12	39.667	104.89	89	Ribosome-binding ATPase YchF	ychF
5	22.981	66.694	54	Enhancing lycopene biosynthesis protein 2	elbB
1	25.597	5.5844	8	Protein TolQ	tolQ
2	19.997	3.4303	9	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase KdsC	kdsC
11	47.283	89.675	73	Chaperone SurA	surA
3	27.829	18.156	9	Outer membrane protein assembly factor BamD	bamD
5	30.615	4.3839	5	Dihydropteroate synthase	folP
2	14.082	16.917	10	D-erythro-7,8-dihydroneopterin triphosphate epimerase	folX
18	60.298	114.11	52	Fumarate hydratase class I, aerobic	fumA
15	52.356	116.32	111	Aspartate ammonia-lyase	aspA
9	64.421	75.429	38	Succinate dehydrogenase flavoprotein subunit	sdhA
6	27.123	59.456	37	Fumarate reductase iron-sulfur subunit	frdB
8	55.704	13.082	10	Glucose-6-phosphate 1-dehydrogenase;Extracellular death factor	zwf
9	24.35	85.686	67	Glutaredoxin-2	grxB
4	9.1374	53.959	32	Glutaredoxin-3	grxC
4	12.879	31.902	21	Glutaredoxin-4	grxD
3	14.92	8.9925	17	Lactoylglutathione lyase	gloA
6	28.434	80.664	22	Hydroxyacylglutathione hydrolase	gloB
2	6.5755	2.3951	14	Protein GnsA	gnsA
11	24.305	147.35	52	Stringent starvation protein A	sspA
2	23.713	16.221	24	Glutathione S-transferase GstB	gstB
7	35.624	50.708	28	Delta-aminolevulinic acid dehydratase	hemB
3	45.244	1.9893	6	Protein HemY	hemY
4	12.1	33.159	19	Iron-sulfur cluster insertion protein ErpA	erpA

7	49.19	21.251	21	Bifunctional protein GlmU;UDP-N-acetylglucosamine pyrophosphorylase;Glucosamine-1-phosphate N-acetyltransferase	glmU
3	13.848	41.709	8	Iron-sulfur cluster assembly scaffold protein IscU	iscU
9	66.253	19.426	12	Hydrogenase-1 large chain	hyaB
14	62.49	36.1	24	Hydrogenase-2 large chain	hybC
4	13.241	94.092	38	HIT-like protein HinT	hinT
5	9.5349	196.6	306	DNA-binding protein HU-alpha	hupA
4	9.2255	99.126	117	DNA-binding protein HU-beta	hupB
8	15.539	195.48	213	DNA-binding protein H-NS	hns
6	15.347	74.788	98	DNA-binding protein StpA	stpA
7	16.887	69.866	52	Regulatory protein AsnC	asnC
2	18.887	1.8769	4	Leucine-responsive regulatory protein	lrp
5	23.64	32.739	24	cAMP-activated global transcriptional regulator CRP	crp
2	29.425	5.7241	8	Pyruvate dehydrogenase complex repressor	pdhR
1	36.422	1.427	7	HTH-type transcriptional regulator GntR	gntR
1	34.276	4.2898	1	Hydrogen peroxide-inducible genes activator	oxyR
7	20.563	36.255	32	Transcriptional repressor MprA	mprA
2	11.288	21.684	13	Putative monooxygenase YdhR	ydhR
3	20.059	17.993	18	Putative NAD(P)H nitroreductase YdjA	ydjA
1	22.69	6.2878	3	Positive transcription regulator EvgA	evgA
2	25.039	6.5886	12	Uncharacterized protein YecA	yecA
2	9.1225	5.4761	3	Uncharacterized protein YecJ	yecJ
2	29.68	4.5768	5	Protein YeeZ	yeeZ
4	8.2884	5.4888	8	UPF0352 protein YejL	yejL
2	10.318	74.361	45	UPF0381 protein YfcZ	yfcZ
5	12.784	112.4	106	Ribosome-associated inhibitor A	raiA
6	16.872	31.196	45	Inhibitor of vertebrate lysozyme	ivy
28	50.729	323.31	336	Pyruvate kinase I	pykF
5	23.622	21.019	10	Acyl-CoA thioesterase I	tesA
1	40.149	5.1502	6	Murein hydrolase activator NlpD	nlpD
2	20.95	6.9317	9	Uncharacterized lipoprotein YajG	yajG
3	12.021	23.149	9	Osmotically-inducible lipoprotein E	osmE
1	4.8095	9.4014	11	Entericidin B	ecnB

1	21.356	4.0123	3	LPS-assembly lipoprotein LptE	lptE
9	36.094	49.234	52	tRNA-modifying protein YgfZ	ygfZ
4	11.106	6.4319	7	Acetolactate synthase isozyme 1 small subunit	ilvN
5	29.172	39.799	18	Inositol-1-monophosphatase	suhB
20	52.022	172.84	140	Inosine-5-monophosphate dehydrogenase	guaB
2	20.452	2.4657	6	Isochorismatase family protein YecD	yecD
3	25.663	15.536	7	Uncharacterized protein YiaF	yiaF
2	13.696	17.872	4	Uncharacterized protein YibL	yibL
9	13.133	94.939	183	UPF0438 protein YifE	yifE
1	7.1771	3.3404	5	Uncharacterized lipoprotein YifL	yifL
5	10.273	28.126	30	Protein YihD	yihD
3	12.594	45.596	22	Cell division protein ZapA	zapA
4	26.429	47.482	17	Uncharacterized protein YggN	yggN
3	45.025	1.9853	4	Putative acid--amine ligase YgiC	ygiC
2	11.532	5.8653	7	Probable quinol monooxygenase YgiN	ygiN
7	14.011	53.68	42	Protein YgiW	ygiW
3	20.127	31.297	18	Lipopolysaccharide export system protein LptA	lptA
1	37.129	3.6682	3	Uncharacterized protein YhbW	yhbW
6	23.962	42.34	30	Probable phospholipid-binding protein MlaC	mlaC
5	14.961	52.025	26	Inner membrane protein YhcB	yhcB
3	14.517	5.191	6	Protein YhfA	yhfA
5	16.624	18.929	24	Uncharacterized protein YhhA	yhhA
12	68.149	54.398	33	Peptidyl-prolyl cis-trans isomerase D	ppiD
7	13.541	247.41	90	50S ribosomal protein L14	rplN
4	15.281	88.334	53	50S ribosomal protein L16	rplP
3	11.199	12.699	43	50S ribosomal protein L23	rplW
3	10.269	79.387	35	30S ribosomal protein S15	rpsO
1	11.887	12.909	18	UPF0092 membrane protein YajC	yajC
5	27.048	9.8495	6	tRNA (cytidine/uridine-2-O-)-methyltransferase TrmJ	trmJ
11	42.196	44.633	35	Multidrug efflux pump subunit AcrA	acrA
13	20.761	249.96	4390	Alkyl hydroperoxide reductase subunit C	ahpC
4	53.994	19.758	5	AMP nucleosidase	amn
5	29.33	24.352	22	Methionine aminopeptidase	map

5	26.103	39.575	17	Class B acid phosphatase	aphA
2	12.866	7.9649	4	Gamma-glutamylcyclotransferase family protein YtfP	ytfP
8	17.634	65.142	55	Putative peroxiredoxin bcp	bcp
6	33.298	13.072	16	Magnesium and cobalt efflux protein CorC	corC
3	18.965	7.9522	10	Periplasmic protein CpxP	cpxP
7	26.312	52.556	42	Transcriptional regulatory protein CpxR	cpxR
3	17.108	60.683	15	Protein CreA	creA
4	87.982	12.466	10	Aerobic respiration control sensor protein ArcB	arcB
8	16.066	158.36	79	Universal stress protein A	uspA
1	23.892	9.8527	4	Response regulator UvrY	uvrY
5	41.269	10.782	9	Succinyl-diaminopimelate desuccinylase	dapE
5	19.787	12.35	15	Protein DcrB	dcrB
19	35.712	91.208	91	D-galactose-binding periplasmic protein	mglB
5	23.104	85.985	34	Thiol:disulfide interchange protein DsbA	dsbA
3	25.621	27.039	27	Thiol:disulfide interchange protein DsbC	dsbC
3	11.306	86.83	55	Protein ElaB	elaB
12	53.662	91.524	45	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase	miaB
10	25.56	221.05	109	3-oxoacyl-[acyl-carrier-protein] reductase FabG	fabG
7	27.864	64.122	28	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI	fabI
3	16.081	16.868	26	FKBP-type 16 kDa peptidyl-prolyl cis-trans isomerase	fkpB
3	29.039	4.6071	9	Cystine-binding periplasmic protein	fliY
8	32.942	26.185	37	UTP--glucose-1-phosphate uridylyltransferase	galU
10	27.19	57.649	65	Glutamine-binding periplasmic protein	glnH
12	96.962	58.95	32	DNA gyrase subunit A	gyrA
9	89.949	31.374	23	DNA gyrase subunit B	gyrB
4	11.857	124.25	73	Acid stress chaperone HdeA	hdeA
4	12.043	49.444	112	Acid stress chaperone HdeB	hdeB
7	26.778	39.98	22	7-alpha-hydroxysteroid dehydrogenase	hdhA
9	28.483	20.505	29	Histidine-binding periplasmic protein	hisJ
7	17.688	78.038	72	Chaperone protein Skp	skp
1	17.056	9.1665	4	Hydrogenase 3 maturation protease	hycI
3	33.755	16.704	12	1-phosphofructokinase	fruK

19	43.387	132.55	85	Maltose-binding periplasmic protein	malE
5	21.891	59.832	28	Modulator of drug activity B	mdaB
6	29.614	24.579	51	Septum site-determining protein MinD	minD
7	18.665	58.026	72	Molybdenum cofactor biosynthesis protein B	moaB
1	39.938	3.5961	4	Protein mrp	mrp
6	24.354	79.21	27	5-methylthioadenosine/S-adenosylhomocysteine nucleosidase	mtnN
4	40.949	7.4038	13	N-acetylglucosamine-6-phosphate deacetylase	nagA
3	27.163	15.854	6	Ribonucleotide monophosphatase NagD	nagD
5	26.449	39.381	19	Nitrate reductase molybdenum cofactor assembly chaperone NarJ	narJ
9	23.927	36.919	42	Nitrate/nitrite response regulator protein NarL	narL
5	9.6347	175.12	112	Cell division protein ZapB	zapB
1	15.656	14.43	1	UPF0047 protein YjbQ	yjbQ
4	13.519	12.771	22	Uncharacterized protein YjbR	yjbR
3	45.711	22.306	8	Guanine/hypoxanthine permease GhxP	ghxP
4	15.603	45.309	39	Regulator of ribonuclease activity B	rraB
4	13.611	99.162	135	2-iminobutanoate/2-iminopropanoate deaminase	ridA
10	16.865	169.94	169	Toxin-antitoxin biofilm protein TabA	tabA
2	16.457	13.208	9	NADH-quinone oxidoreductase subunit A	nuoA
4	25.056	12.343	13	NADH-quinone oxidoreductase subunit B	nuoB
6	18.59	16.378	11	NADH-quinone oxidoreductase subunit E	nuoE
3	20.538	7.5961	10	NADH-quinone oxidoreductase subunit I	nuoI
2	43.475	7.5852	5	Nucleoside permease NupC	nupC
19	54.87	323.31	165	Transcription termination/antitermination protein NusA	nusA
10	20.531	56.001	58	Transcription termination/antitermination protein NusG	nusG
20	105.06	143.97	82	2-oxoglutarate dehydrogenase E1 component	sucA
8	44.011	136.24	62	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	sucB
38	99.667	323.31	376	Pyruvate dehydrogenase E1 component	aceE
7	21.073	108.58	150	Osmotically-inducible protein Y	osmY
4	25.545	25.831	20	Pyridoxine/pyridoxamine 5-phosphate oxidase	pdxH
2	12.345	7.4139	23	Protein PhnA	phnA
1	53.388	1.4106	4	Low-affinity inorganic phosphate transporter 1	pitA

6	48.369	13.528	15	Metalloprotease PmbA	pmbA
5	20.431	43.442	41	Peptidyl-prolyl cis-trans isomerase A	ppiA
7	36.022	70.905	40	Glycine betaine-binding periplasmic protein	proX
8	25.493	68.251	48	Phage shock protein A	pspA
6	26.892	41.151	32	Putative GTP cyclohydrolase 1 type 2	ybgI
6	23.211	82.877	17	Uncharacterized protein YciO	yciO
4	23.445	10.276	12	Riboflavin synthase	ribC
3	14.927	14.17	16	Regulator of nucleoside diphosphate kinase	rnk
3	9.4795	18.038	14	Protein rof	rof
4	20.315	15.558	16	Negative modulator of initiation of replication	seqA
3	28.642	21.015	8	Protein SseB	sseB
3	18.262	20.858	10	Stringent starvation protein B	sspB
7	24.554	33.606	22	Ribulose-phosphate 3-epimerase	rpe
4	56.487	15.719	10	Amidophosphoribosyltransferase	purF
6	17.78	36.236	25	N5-carboxyaminoimidazole ribonucleotide mutase	purE
15	47.004	93.903	72	Transcription termination factor Rho	rho
7	14.364	77.702	175	50S ribosomal protein L17	rplQ
3	11.564	22.279	74	50S ribosomal protein L21	rplU
5	6.5417	13.616	11	50S ribosomal protein L30	rpmD
10	18.904	178.44	294	50S ribosomal protein L6	rplF
2	11.58	57.426	32	30S ribosomal protein S14	rpsN
2	9.7043	57.346	9	30S ribosomal protein S17	rpsQ
39	61.157	323.31	722	30S ribosomal protein S1	rpsA
4	31.488	7.1188	5	Uncharacterized oxidoreductase YghA	yghA
5	17.277	83.429	82	Protein-export protein SecB	secB
4	66.631	9.4362	8	Protein translocase subunit SecD	secD
2	35.382	9.9005	9	Protein translocase subunit SecF	secF
1	11.365	9.6843	6	Protein-export membrane protein SecG	secG
1	48.511	1.1808	3	Protein translocase subunit SecY	secY
2	21.695	8.1089	8	ECF RNA polymerase sigma-E factor	rpoE
3	73.352	2.7692	3	Soluble lytic murein transglycosylase	slt
3	17.681	3.5621	7	Superoxide dismutase [Cu-Zn]	sodC
12	21.265	57.889	41	Superoxide dismutase [Fe]	sodB

9	49.787	43.864	36	Signal recognition particle protein	ffh
5	18.975	43.668	19	Single-stranded DNA-binding protein	ssb
2	20.375	6.3939	12	Chromate reductase	chrR
9	29.777	27.777	44	Succinyl-CoA ligase [ADP-forming] subunit alpha	sucD
3	15.94	14.74	6	Sulfur acceptor protein CsdE	csdE
8	35.232	41.042	21	L-threonine dehydratase catabolic TdcB	tdcB
3	31.966	18.564	10	Acyl-CoA thioesterase 2	tesB
8	51.363	52.897	30	Metalloprotease TldD	tldD
2	33.451	9.3111	6	Ribose transport system permease protein RbsC	rbsC
7	37.784	36.207	21	Uncharacterized tRNA/rRNA methyltransferase YfiF	yfiF
5	47.526	39.332	24	Tyrosine--tRNA ligase	tyrS
1	10.784	3.9115	5	RNA-binding protein YhbY	yhbY
2	17.336	15.516	6	HTH-type transcriptional regulator IscR	iscR
3	14.007	35.867	37	Putative reactive intermediate deaminase TdcF	tdcF
3	12.769	67.386	35	50S ribosomal protein L18	rplR
2	10.234	5.4586	5	UPF0345 protein YaiE	yaiE
2	9.0714	3.7151	6	Antitoxin RelB	relB
6	15.088	24.033	30	Peroxiredoxin OsmC	osmC
2	54.845	3.0352	5	Proline/betaine transporter	proP
2	7.7315	2.7492	6	Protein IscX	iscX
2	23.335	9.6335	3	Ribosomal RNA large subunit methyltransferase E	rlmE
4	30.896	12.081	6	Small-conductance mechanosensitive channel	mscS
18	49.354	163.26	110	Periplasmic serine endoprotease DegP	degP
10	30.812	76.763	49	D-tagatose-1,6-bisphosphate aldolase subunit GatY	gatY
12	47.108	204.27	67	D-tagatose-1,6-bisphosphate aldolase subunit GatZ	gatZ
23	43.313	323.31	2351	Elongation factor Tu 2;Elongation factor Tu 1	tufB;tufA
5	23.67	22.141	21	Transcriptional regulatory protein RcsB	rcsB
12	54.513	66.111	36	Signal recognition particle receptor FtsY	ftsY
3	48.541	9.6444	5	Long-chain fatty acid transport protein	fadL
12	102.02	38.297	42	Protein translocase subunit SecA	secA
8	58.066	70.173	21	Respiratory nitrate reductase 1 beta chain	narH
3	46.161	5.8121	12	Cell division protein DamX	damX
10	64.682	102.76	43	Arginine--tRNA ligase	argS

6	44.067	35.34	35	Molybdopterin molybdenumtransferase	moeA
1	25.693	3.9902	5	Uracil-DNA glycosylase	ung
10	27.159	111.71	99	Uridine phosphorylase	udp
1	17.085	14.319	7	UPF0098 protein YbhB	ybhB
2	136	1.336	4	Methionine synthase	metH
5	34.746	7.4581	11	Protein FdhE	fdhE
25	80.023	223.91	222	Catalase-peroxidase	katG
4	45.357	2.7883	8	Anaerobic glycerol-3-phosphate dehydrogenase subunit B	glpB
15	56.75	74.283	52	Aerobic glycerol-3-phosphate dehydrogenase	glpD
6	37.971	40.004	17	RNA polymerase sigma factor RpoS	rpoS
12	63.636	37.73	28	Periplasmic trehalase	treA
12	60.105	8.4199	32	Fumarate hydratase class I, anaerobic	fumB
3	46.973	27.922	4	UDP-N-acetylmuramoylalanine--D-glutamate ligase	murD
12	49.815	49.916	49	Xaa-Pro aminopeptidase	pepP
21	141.4	67.838	54	Phosphoribosylformylglycinamidine synthase	purL
16	52.915	116.98	70	Cytosol non-specific dipeptidase	pepD
25	57.329	310.86	172	Bifunctional purine biosynthesis protein PurH;Phosphoribosylaminoimidazolecarboxamide formyltransferase;IMP cyclohydrolase	purH
13	45.94	147.31	94	Phosphoribosylamine--glycine ligase	purD
2	29.413	3.4037	7	Shikimate dehydrogenase (NADP(+))	aroE
3	48.906	9.8445	6	L-serine dehydratase 1	sdaA
1	64.98	5.7264	3	Formate hydrogenlyase subunit 5	hycE
6	36.687	32.639	26	Selenide, water dikinase	selD
20	63.692	231.93	113	Proline--tRNA ligase	proS
4	37.614	10.741	6	Thiosulfate-binding protein	cysP
5	26.8	26.379	16	Oxygen-insensitive NADPH nitroreductase	nfsA
12	66.894	51.636	26	Glutamine--fructose-6-phosphate aminotransferase [isomerizing]	glmS
2	63.997	1.5568	5	Sulfite reductase [NADPH] hemoprotein beta-component	cysI
3	26.555	6.1042	4	Ubiquinone biosynthesis O-methyltransferase	ubiG
2	43.766	1.5244	4	Acetylornithine/succinyldiaminopimelate aminotransferase	argD

2	17.352	6.5059	15	Uncharacterized protein YjjA	yjjA
12	90.398	24.809	27	Dimethyl sulfoxide reductase DmsA	dmsA
2	22.869	6.8356	5	Anaerobic dimethyl sulfoxide reductase chain B	dmsB
9	30.636	30.439	32	NH(3)-dependent NAD(+) synthetase	nadE
18	45.682	184.62	97	Glucose-1-phosphatase	agp
2	21.546	7.6078	7	Hydrogenase 1 maturation protease	hyaD
5	14.89	9.5422	4	Hydrogenase-1 operon protein HyaE	hyaE
2	57.518	10.426	4	PTS system fructose-specific EIIBC component;Fructose-specific phosphotransferase enzyme IIB component;Fructose permease IIC component	fruA
15	50.176	58.3	74	Xaa-Pro dipeptidase	pepQ
6	73.898	84.211	29	Biosynthetic arginine decarboxylase	speA
6	84.162	9.0613	24	Catalase HPII	katE
3	18.597	8.3083	15	Protein YciF	yciF
2	18.961	3.5507	5	Protein YciE	yciE
11	92.108	66.24	36	Ribonuclease R	rnr
6	49.914	39.007	15	ATP-dependent RNA helicase SrmB	srmB
21	118.2	238.21	86	Ribonuclease E	rne
19	51.357	228.65	151	Pyruvate kinase II	pykA
1	36.038	10.494	5	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase	lpxD
7	52.202	88.226	33	Cysteine--tRNA ligase	cysS
14	65.913	148.59	52	Aspartate--tRNA ligase	aspS
2	62.658	12.045	6	Asparagine synthetase B [glutamine-hydrolyzing]	asnB
6	53.343	10.227	16	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	murE
6	27.176	24.831	29	3(2),5-bisphosphate nucleotidase CysQ	cysQ
6	45.774	29.365	24	4-aminobutyrate aminotransferase GabT	gabT
22	59.643	253.11	172	Phosphoenolpyruvate carboxykinase [ATP]	pckA
6	36.397	13.863	16	Adenosine deaminase	add
5	170.23	5.5395	4	Chromosome partition protein MukB	mukB
4	26.974	11.526	10	Chromosome partition protein MukE	mukE
4	27.384	7.3367	8	5-keto-4-deoxy-D-glucarate aldolase	garL
13	39.783	86.503	53	Phosphoserine aminotransferase	serC

5	18.192	20.675	33	Ecotin	eco
9	25.535	41.744	54	Transcriptional regulatory protein PhoP	phoP
4	33.175	33.401	12	UPF0701 protein YicC	yicC
5	60.898	7.1093	5	Periplasmic oligopeptide-binding protein	oppA
4	11.475	29.346	33	Thiosulfate sulfurtransferase PspE	pspE
3	76.662	2.4157	7	Tail-specific protease	prc
7	18.153	45.341	47	Peptidyl-prolyl cis-trans isomerase B	ppiB
9	34.168	33.408	28	Methionyl-tRNA formyltransferase	fmt
5	45.366	20.269	8	Glutamate-1-semialdehyde 2,1-aminomutase	hemL
4	32.499	41.793	21	Fructokinase	mak
7	77.515	15.953	10	Peptidyl-dipeptidase dcp	dcp
2	13.6	6.1904	9	Protein YffB	yffB
7	49.32	28.693	31	Biotin carboxylase	accC
3	31.043	5.5295	10	Bifunctional protein FolD;Methylenetetrahydrofolate dehydrogenase;Methenyltetrahydrofolate cyclohydrolase	folD
11	44.838	80.154	70	Mannonate dehydratase	uxuA
4	43.867	11.766	9	Flavohemoprotein	hmp
5	39.359	16.836	35	S-(hydroxymethyl)glutathione dehydrogenase	frmA
13	97.676	43.406	29	Aconitate hydratase A	acnA
5	49.23	21.905	7	tRNA modification GTPase MnmE	mnmE
5	51.719	20.453	10	Succinate-semialdehyde dehydrogenase [NADP(+)] GabD	gabD
2	21.515	10.038	4	Maf-like protein YhdE	yhdE
2	40.338	5.1403	4	Riboflavin biosynthesis protein RibD;Diaminohydroxyphosphoribosylaminopyrimidine deaminase;5-amino-6-(5-phosphoribosylamino)uracil reductase	ribD
14	52.272	86.601	63	Lactaldehyde dehydrogenase	aldA
9	61.525	76.543	25	Membrane protein insertase YidC	yidC
4	14.259	105.07	31	Acidic protein MsyB	msyB
3	40.959	9.1163	8	tRNA-specific 2-thiouridylase MnmA	mnmA
6	45.068	105.59	42	NAD-dependent dihydropyrimidine dehydrogenase subunit PreA	preA
5	30.706	33.816	9	Putative oxidoreductase YdbC	ydbC
7	63.197	30.869	13	NAD-dependent malic enzyme	maeA

9	34.723	26.749	28	Probable acrylyl-CoA reductase AcuI	acuI
12	40.146	115.72	65	Aminomethyltransferase	gcvT
19	77.166	101.53	75	Oligopeptidase A	prlC
12	72.211	111.9	97	Transketolase 1	tktA
2	42.578	12.764	4	50S ribosomal protein L16 arginine hydroxylase	roxA
5	36.191	13.223	9	Cytoskeleton protein RodZ	rodZ
3	29.707	1.3523	1	Pyridoxal phosphate phosphatase YigL	yigL
3	35.172	18.25	18	Quinone oxidoreductase 1	qorA
5	27.751	32.868	27	Ferredoxin--NADP reductase	fpr
7	80.022	14.367	13	Anaerobic ribonucleoside-triphosphate reductase	nrdD
11	63.837	132.82	42	Trehalose-6-phosphate hydrolase	treC
2	35.793	3.1795	3	Cell division protein FtsN	ftsN
6	24.177	67.929	23	UPF0502 protein YceH	yceH
8	44.923	19.915	19	Peptidase T	pepT
5	32.762	13.379	7	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	nadC
4	35.048	15.098	15	Uncharacterized protein YbiB	ybiB
1	8.7578	21.187	7	Molybdopterin synthase sulfur carrier subunit	moaD
3	16.981	7.9965	6	Molybdopterin synthase catalytic subunit	moaE
7	26.929	18.6	17	Putative ABC transporter arginine-binding protein 2	artI
3	29.436	26.476	4	2,5-diketo-D-gluconic acid reductase B	dkgB
5	28.237	21.035	21	3-methyl-2-oxobutanoate hydroxymethyltransferase	panB
2	14.983	4.6259	3	Uncharacterized lipoprotein YedD	yedD
2	46.062	1.9462	5	Voltage-gated potassium channel Kch	kch
8	47.543	13.324	21	Phosphoglucosamine mutase	glmM
7	30.811	122.01	53	3-mercaptopyruvate sulfurtransferase	sseA
2	113.57	1.5488	3	Multidrug efflux pump subunit AcrB	acrB
2	46.186	4.4653	4	Uncharacterized protein YidR	yidR
2	26.076	4.8263	5	Uncharacterized HTH-type transcriptional regulator YieP	yieP
2	36.163	5.6047	3	Thiamine-binding periplasmic protein	thiB
10	31.19	142.75	30	Molecular chaperone Hsp31 and glyoxalase 3	hchA
8	31.597	44.317	25	Pantothenate synthetase	panC
4	26.689	10.315	12	Chaperone protein FimC	fimC

8	23.575	57.93	37	Nitrate/nitrite response regulator protein NarP	narP
10	49.292	57.878	37	NADH-quinone oxidoreductase subunit F	nuoF
4	52.728	12.057	4	Oxygen-independent coproporphyrinogen-III oxidase	hemN
6	67.355	43.863	16	GTP-binding protein TypA/BipA	typA
3	25.342	8.4446	6	Protein YiiM	yiiM
1	10.776	5.6837	4	UPF0381 protein YiiS	yiiS
5	27.389	31.335	16	Phosphoglycolate phosphatase	gph
2	22.612	5.5325	6	Uncharacterized protein YjaG	yjaG
14	57.912	96.826	85	Glucans biosynthesis protein G	mdoG
26	104.38	323.31	240	Glycine dehydrogenase (decarboxylating)	gcvP
2	12.962	6.6348	10	Protein YebF	yebF
9	42.433	49.595	27	Phosphoribosylglycinamide formyltransferase 2	purT
6	60.589	26.929	10	Putative acyl-CoA dehydrogenase AidB	aidB
2	17.973	2.9354	4	Uncharacterized protein YehS	yehS
1	34.424	1.5238	4	Putative osmoprotectant uptake system ATP-binding protein YehX	yehX
4	32.609	8.274	6	Putative osmoprotectant uptake system substrate-binding protein OsmF	osmF
6	83.459	13.848	15	Periplasmic beta-glucosidase	bglX
8	73.042	61.159	31	Transketolase 2	tktB
5	58.718	29.029	13	Nickel-binding periplasmic protein	nikA
14	68.235	51.321	43	NADH-quinone oxidoreductase subunit C/D	nuoC
15	100.3	173.02	82	NADH-quinone oxidoreductase subunit G	nuoG
3	37.121	9.5426	16	Ribosomal large subunit pseudouridine synthase D	rluD
4	37.822	12.407	17	Nucleoid-associated protein YejK	yejK
14	56.176	71.404	65	Alkyl hydroperoxide reductase subunit F	ahpF
11	34.455	69.073	38	Curved DNA-binding protein	cbpA
3	28.291	7.2893	7	Cell division protein ZapD	zapD
24	93.497	129.98	99	Aconitate hydratase B	acnB
2	34.593	6.2508	7	Probable HTH-type transcriptional regulator LrhA	lrhA
8	58.36	35.42	17	Phosphoglucomutase	pgm
4	31.934	29.507	26	Formyltetrahydrofolate deformylase	purU
19	46.18	232.46	107	Peptidase B	pepB
3	17.751	10.255	5	Hydrogenase 2 maturation protease	hybD

3	20.964	38.112	14	Outer membrane protein slp	slp
12	80.488	85.762	29	Malate synthase G	glcB
6	27.85	43.594	21	Oxidoreductase UcpA	ucpA
2	43.741	3.1161	7	Uncharacterized protein YhiN	yhiN
3	31.941	2.2158	5	Ribosomal RNA large subunit methyltransferase J	rlmJ
9	41.19	83.314	33	Multidrug resistance protein MdtE	mdtE
9	33.962	226.48	105	2-dehydro-3-deoxygluconokinase	kdgK
7	55.527	16.245	22	Protein YhjJ	yhjJ
3	35.395	9.9382	15	Glyoxylate/hydroxypyruvate reductase B	ghrB
10	56.306	35.762	22	Aldehyde dehydrogenase B	aldB
17	56.193	145.77	139	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	gpmI
4	32.693	13.038	6	Glucose-1-phosphate thymidylyltransferase 1	rmlA1
2	21.27	13.217	4	dTDP-4-dehydrorhamnose 3,5-epimerase	rfbC
7	42.965	49.137	26	UDP-galactopyranose mutase	glf
1	21.675	1.0552	3	Putative lipopolysaccharide biosynthesis O-acetyl transferase WbbJ	wbbJ
4	40.558	11.982	14	dTDP-glucose 4,6-dehydratase 1	rfbB
11	27.07	135.37	149	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase	kduD
2	49.874	1.9826	3	UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase	mpl
5	33.42	18.459	17	Glutamate/aspartate periplasmic-binding protein	gltI
4	16.016	29.925	33	Universal stress protein F	uspF
2	66.269	7.4122	4	Sulfite reductase [NADPH] flavoprotein alpha-component	cysJ
8	23.905	35.849	31	Oxygen-insensitive NAD(P)H nitroreductase	nfsB;NTR
4	12.88	18.972	26	Uncharacterized protein YggL	yggL
7	47.204	41.962	41	Periplasmic pH-dependent serine endoprotease DegQ	degQ
4	53.58	19.434	7	D-mannonate oxidoreductase	uxuB
3	33.777	3.6307	7	High-affinity zinc uptake system protein ZnuA	znuA
6	32.666	78.948	37	Putative glucose-6-phosphate 1-epimerase	yeaD
8	15.935	106.06	67	Universal stress protein G	uspG
15	106.82	182.86	42	Antigen 43;Antigen 43 alpha chain;Antigen 43 beta chain	flu
9	32.91	34.037	34	D-allose-binding periplasmic protein	alsB

1	38.749	9.5383	6	L-lysine 2,3-aminomutase	epmB
2	39.193	5.8247	2	Putative ribosome biogenesis GTPase RsgA	rsgA
4	34.344	20.089	29	ABC transporter periplasmic-binding protein YtfQ	ytfQ
3	41.083	11.906	4	Isoaspartyl dipeptidase	iadA
5	42.742	19.315	9	Uncharacterized protein YjiM	yjiM
4	37.624	18.702	10	Ribosomal RNA small subunit methyltransferase C	rsmC
5	27.249	78.279	36	NADP-dependent 3-hydroxy acid dehydrogenase YdfG	ydfG
2	62.757	3.8057	6	Glucans biosynthesis protein D	mdoD
6	40.902	27.08	19	N-methyl-L-tryptophan oxidase	solA
3	21.657	17	10	FMN-dependent NADH-azoreductase	azoR
1	12.315	1.4005	7	tRNA-binding protein YgjH	ygjH
14	85.935	35.262	39	PFL-like enzyme TdcE	tdcE
5	43.285	14.49	8	GTPase ObgE/CgtA	obgE
1	17.038	3.6052	3	Uncharacterized protein YhcH	yhcH
3	18.858	5.0249	4	Protein YhbO	yhbO
15	28.882	97.021	82	FKBP-type peptidyl-prolyl cis-trans isomerase FkpA	fkpA
6	25.892	20.525	27	RNA chaperone ProQ	proQ
2	20.767	6.1439	4	Threonylcarbamoyl-AMP synthase	tsaC
7	28.231	41.319	22	Uncharacterized protein YbgF	ybgF
9	85.119	19.418	14	Protein YhgF	yhgF
6	38.765	18.03	13	Uncharacterized oxidoreductase YhhX	yhhX
1	88.963	1.0677	3	Trimethylamine-N-oxide reductase 2	torZ
5	19.923	4.5542	10	Uncharacterized protein YaiL	yaiL
2	21.039	23.356	13	dITP/XTP pyrophosphatase	rdgB
2	36.534	5.8229	9	D-lactate dehydrogenase	ldhA
8	36.307	37.098	38	6-phosphogluconolactonase	pgl
4	34.877	16.276	5	Ribosomal RNA small subunit methyltransferase H	rsmH
11	29.86	135.21	234	50S ribosomal protein L2	rplB
9	22.243	96.274	171	50S ribosomal protein L3	rplC
4	23.592	18.045	8	Guanylate kinase	gmk
7	11.316	45.289	98	50S ribosomal protein L24	rplX
3	28.454	5.2383	6	Imidazole glycerol phosphate synthase subunit HisF	hisF
9	36.071	26.797	13	Lipoyl synthase	lipA

7	22.086	256.93	78	50S ribosomal protein L4	rplD
4	33.366	6.7237	5	ATP phosphoribosyltransferase	hisG
8	66.57	14.741	19	Elongation factor 4	lepA
9	47.029	21.907	12	Histidine--tRNA ligase	hisS
7	12.226	126.54	61	50S ribosomal protein L22	rplV
3	22.497	20.231	16	Outer-membrane lipoprotein carrier protein	lolA
5	25.096	39.458	26	Carbonic anhydrase 2	can
7	16.156	249.33	96	6,7-dimethyl-8-ribityllumazine synthase	ribE
18	32.337	308.92	435	Malate dehydrogenase	mdh
7	19.737	297.3	82	Flavodoxin-1	fldA
10	20.301	123.15	129	50S ribosomal protein L5	rplE
2	16.897	19.322	4	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	ispF
8	40.683	27.084	21	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	ispG
4	34.774	21.029	6	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	ispH
13	28.556	107.81	111	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	gpmA
7	15.096	49.768	45	UPF0325 protein YaeH	yaeH
5	20.997	96.234	76	Fe/S biogenesis protein NfuA	nfuA
2	26.556	4.0705	3	23S rRNA (guanosine-2-O-)-methyltransferase RlmB	rlmB
7	20.815	77.92	50	Phosphoheptose isomerase	gmhA
2	21.294	9.1075	11	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase	gmhB
2	55.076	10.323	5	Probable glutamate/gamma-aminobutyrate antiporter	gadC
40	95.584	323.31	287	Chaperone protein ClpB	clpB
2	18.534	6.3409	7	Uncharacterized N-acetyltransferase YhbS	yhbS
2	31.46	2.3932	3	Uncharacterized protein YpfJ	ypfJ
2	8.7998	5.0526	7	Uncharacterized protein YdcY	ydcY
3	8.4165	6.907	7	OriC-binding nucleoid-associated protein	cnu
3	17.85	24.626	6	Uncharacterized protein YeaK	yeaK
1	10.835	5.3685	5	Uncharacterized protein YoaC	yoaC
1	8.7527	39.947	3	Uncharacterized protein YebV	yebV
2	12.378	4.4622	9	Uncharacterized protein YebY	yebY
3	8.5791	24.19	11	Uncharacterized protein YodD	yodD

3	12.466	13.746	10	Nickel/cobalt homeostasis protein RcnB	rcnB
2	9.9999	6.2407	2	Uncharacterized protein YfcL	yfcL
2	10.547	10.357	6	Antitoxin CptB	cptB
4	11.051	40.785	20	Uncharacterized protein YqjD	yqjD
4	23.401	34.836	11	Transcriptional regulator YqjI	yqjI
3	20.028	5.7914	8	Uncharacterized protein YraP	yraP
1	10.68	2.1537	1	Probable phospholipid ABC transporter-binding protein MlaB	mlaB
1	19.576	1.7945	4	Probable phospholipid ABC transporter-binding protein MlaD	mlaD
1	26.82	1.1058	4	Uncharacterized protein YheO	yheO
1	8.1741	2.6759	3	Uncharacterized lipoprotein YgdI	ygdI
3	16.274	21.024	19	Uncharacterized protein YqcA	yqcA
3	20.376	49.007	12	Uncharacterized Nudix hydrolase YfcD	yfcD
3	25.787	8.8776	5	UPF0001 protein YggS	yggS
3	20.122	8.5998	8	Phosphodiesterase YfcE	yfcE
2	6.5542	3.1607	5	UPF0181 protein YoaH	yoaH
3	11.905	3.5813	7	UPF0267 protein YqfB	yqfB
12	34.893	76.44	69	ADP-L-glycero-D-manno-heptose-6-epimerase	hldD
15	14.284	323.31	543	Autonomous glycyl radical cofactor	grcA
2	40.99	13.056	18	Maltose/maltodextrin import ATP-binding protein MalK	malK
4	8.3252	109.31	23	UPF0337 protein YjbJ	yjbJ
3	8.4999	13.697	9	30S ribosomal protein S21	rpsU
13	54.879	229.9	80	Cytosol aminopeptidase	pepA
6	10.693	60.165	102	50S ribosomal protein L25	rplY
4	8.2495	136.1	41	Translation initiation factor IF-1	infA
5	27.656	13.436	9	Transcriptional regulatory protein BaeR	baeR
3	14.163	7.8947	4	Outer membrane lipoprotein RcsF	rcsF
2	9.6639	2.9051	3	Sec-independent protein translocase protein TatA	tatA
12	23.586	142.54	167	Adenylate kinase	adk
7	19.859	31.489	32	Adenine phosphoribosyltransferase	apt
3	24.882	5.1181	10	Iron-sulfur cluster repair protein YtfE	ytfE
2	8.3234	32.283	27	Major outer membrane lipoprotein Lpp	lpp

11	18.251	111.32	168	Glucose-specific phosphotransferase enzyme IIA component	crr
4	50.676	10.775	11	PTS system glucose-specific EIICB component;Glucose permease IIC component;Glucose-specific phosphotransferase enzyme IIB component	ptsG
1	11.426	2.7119	4	N,N-diacetylchitobiose-specific phosphotransferase enzyme IIB component	chbB
12	35.047	121.85	102	PTS system mannose-specific EIIAB component;Mannose-specific phosphotransferase enzyme IIA component;Mannose-specific phosphotransferase enzyme IIB component	manX
9	39.647	97.995	49	Multiphosphoryl transfer protein;Phosphocarrier protein HPr;Fructose-specific phosphotransferase enzyme IIA component	fruB
3	16.907	17.504	21	Galactitol-specific phosphotransferase enzyme IIA component	gatA
5	17.959	33.074	12	Nitrogen regulatory protein	ptsN
19	52.668	323.31	192	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA
1	6.8558	1.2109	12	Carbon storage regulator	csrA
7	64.976	21.685	15	L-fucose isomerase	fucI
13	43.517	98.845	41	Ribonucleoside-diphosphate reductase 1 subunit beta	nrdB
3	28.437	8.4539	12	Esterase YbfF	ybfF
5	23.784	12.709	33	Uncharacterized protein YcbL	ycbL
6	14.701	32.189	26	Uncharacterized protein YccU	yccU
3	44.356	1.9226	3	Ribosomal RNA large subunit methyltransferase I	rlmI
6	35.343	13.276	21	Glyoxylate/hydroxypyruvate reductase A	ghrA
5	26.89	7.9823	9	Probable phosphatase YcdX	ycdX
2	25.438	7.7971	6	Lipoprotein-releasing system ATP-binding protein LolD	lolD
2	26.716	7.2192	5	NAD-dependent protein deacylase	cobB
7	22.632	34.798	19	PTS-dependent dihydroxyacetone kinase, ADP-binding subunit DhaL	dhaL
7	38.215	32.154	26	PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK	dhaK
1	17.663	15.206	9	Rac prophage repressor	racR
7	37.609	13.354	11	NADPH-dependent curcumin reductase	curA
8	38.612	45.451	35	Uncharacterized protein YncE	yncE
2	12.909	7.1482	2	UPF0482 protein YnfB	ynfB

15	33.903	153.39	82	Protein YdgH	ydgH
3	33.675	16.468	14	Oxidoreductase YdhF	ydhF
7	36.082	33.379	16	Probable L,D-transpeptidase YnhG	ynhG
3	13.386	10.945	10	Uncharacterized protein YeaO	yeaO
1	18.121	13.741	7	Free methionine-R-sulfoxide reductase	msrC
2	30.279	9.068	6	Protein MtfA	mtfA
2	51.193	14.446	5	Uncharacterized protease YegQ	yegQ
12	44.329	62.489	62	NAD-dependent dihydropyrimidine dehydrogenase subunit PreT	preT
3	31.22	7.1327	7	N-acetylmuramic acid 6-phosphate etherase	murQ
9	33.052	47.66	21	Probable deferrochelatase/peroxidase YfeX	yfeX
13	82.416	97.177	44	NADP-dependent malic enzyme	maeB
2	24.855	11.475	3	Esterase YpfH	ypfH
3	13.398	63.412	33	Uncharacterized protein YfgD	yfgD
4	22.176	8.2414	20	UPF0070 protein YfgM	yfgM
5	51.05	10.352	14	Bifunctional protein HldE;D-beta-D-heptose 7-phosphate kinase;D-beta-D-heptose 1-phosphate adenylyltransferase	hldE
4	27.495	8.7257	8	Thiol:disulfide interchange protein DsbG	dsbG
4	24.33	7.242	9	2-deoxyglucose-6-phosphate phosphatase	yniC
7	39.516	18.818	13	N-ethylmaleimide reductase	nemA
15	62.801	231.84	122	Uncharacterized sulfatase YdeN	ydeN
4	127.21	3.3561	4	Mechanosensitive channel MscK	mscK
2	59.899	3.6244	2	Periplasmic murein peptide-binding protein	mppA
10	31.791	153.09	63	Uncharacterized protein YbbN	ybbN
2	32.903	30.905	10	Glutaminase 1	glsA1
3	42.876	7.6999	4	Uncharacterized protein YdhQ	ydhQ
7	23.007	30.543	17	Sugar phosphatase YfbT	yfbT
1	13.3	3.5449	5	Protein SufA	sufA
3	6.5476	14.88	24	Protein GnsB	gnsB
2	19.431	27.654	12	Uncharacterized lipoprotein YbaY	ybaY
3	54.972	5.7341	8	tRNA sulfurtransferase	thiI
4	32.458	23.984	18	Uncharacterized protein YniA	yniA
5	41.887	5.1296	5	Outer membrane protein assembly factor BamB	bamB

11	89.986	36.667	17	Probable dimethyl sulfoxide reductase chain YnfF	ynfF
15	54.688	38.769	38	Protein YdgA	ydgA
4	44.225	16.113	15	NMN amidohydrolase-like protein YfaY	yfaY
8	48.228	23.95	26	Thiosulfate sulfurtransferase YnjE	ynjE
3	7.9952	25.335	25	Uncharacterized protein YibT	yibT
5	55.36	12.365	13	6-phospho-beta-glucosidase BglA	bglA
3	32.391	4.6754	10	Disulfide-bond oxidoreductase YghU	yghU
12	38.832	149.53	90	L-glyceraldehyde 3-phosphate reductase	gpr
9	42.097	117.07	72	Alcohol dehydrogenase YqhD	yqhD
7	31.109	56.531	39	2,5-diketo-D-gluconic acid reductase A	dkgA
11	31.076	233.83	181	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase	kduI
2	20.777	8.0028	19	Probable protein deglycase	yajL
1	39.091	9.7569	5	tRNA pseudouridine synthase D	truD
3	87.872	1.71	4	Copper-exporting P-type ATPase A	copA

Table 3.2S The Cys-containing Peptides Identified in the experiment using MaxQuant platform along with sequence window.

Score	Position within the protein	Protein names	Gene names	Sequence window
222.27	260	Trehalose-6-phosphate hydrolase	treC	GLMTVGEMSSTSLEHCQRYAALTGSELSMTF
195.79	294	Tryptophanase	tnaA	LLCMKDDSFFDVYTECRTLCVVQEGFPTYGG
182.18	114	Elongation factor G	fusA	VERSMRVLDGAVMVYCAVGGVQPQSETVWRQ
183.7	566	Glycine dehydrogenase (decarboxylating)	gcvP	QLADWLVKLTGYDAVCMQPNSGAQGEYAGLL
185.33	82	Protein CreA	creA	GGLGLAEDTSDAAISCQQVGPIELSDRIKNG
223.07	249	Phosphoglycerate kinase	pgk	YEADLVDEAKRLLTTCNIPVPSDVRVATEFS
153.09	63	Ribokinase	rbsK	AVAAGRSGANIAFIACTGDDSIGESVRQQLA
152.99	117	Bifunctional purine biosynthesis protein PurH;Phosphoribosylaminoimidazole carboxamide formyltransferase;IMP cyclohydrolase	purH	VVNLYPFAQTVAREGCSLEDAVENIDIGGPT
149.82	166	Alkyl hydroperoxide reductase subunit C	ahpC	KIKAAQYVASHPGEVCPAKWKEGEATLAPSL
165.11	125	Phosphoenolpyruvate carboxykinase [ATP]	pckA	TRQLSGKRLFVVDAFCGANPDTRLSVRFITE
184.76	155	Periplasmic AppA protein;Phosphoanhydride phosphohydrolase;4-phytase	appA	TSSPDPLFNPLKTGVCQLDNANVTDAILSRA
237.21	84	50S ribosomal protein L14	rplN	VRRPDGSVIRFDGNACVLLNNNSEQPIGTRI
192.4	180	KHG/KDPG aldolase;4-hydroxy-2-oxoglutarate aldolase;2-dehydro-3-deoxy-phosphogluconate aldolase	eda	SPANYRDYLALKSVLCIGGSWLVPADALEAG
199.4	95	Thiol peroxidase	tpx	VLCISADLPFAQSRFCGAEGLNNVITLSTFR
185.12	115	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	mtnN	ADVTAFGYEYGQLPGCPAGFKADDKLIAAAE
130.4	148	Tryptophanase	tnaA	NYFFDTTQGHSQINGCTVRNVYIKEAFDTGV
130.27	68	Osmotically-inducible lipoprotein E	osmE	GKPSSEVSMIHARGTCQTYILGQRDGKAETY
194.53	44	Type-1 fimbrial protein, A chain	fimA	NGGTVHFKGEVVNAACAVDAGSVDQTVQLGQ
120.77	612	Pyruvate dehydrogenase E1 component	aceE	YIYYSMFGFQRIGDLCWAAGDQQARGFLIGG
121.61	25	Dihydrolipoyl dehydrogenase	lpdA	VVLGAGPAGYSAAFRCADLGLETVIVERYNT
114.89	139	Nucleoside diphosphate kinase	ndk	SAAREIAYFFGEGEVCPRTR_____
185.95	65	Translation initiation factor IF-3;Translation initiation factor IF-3, N-terminally processed;Translation initiation factor IF-3S	infC	GVDLVEISPNAEPPVCRIMDYGKFLYEKSKS

219.97	88	Flavodoxin-1	fldA	LEEIDFNGKLVALFGCGDQEDYAEYFCDALG
219.97	99	Flavodoxin-1	fldA	ALFGCGDQEDYAEYFCDALGTIRDIIEPRGA
165.11	209	UDP-N-acetylmuramoylalanine--D-glutamate ligase	murD	QYRAAKLRIYENAKVCVVNADDALTMPIRGA
111.64	251	Malate dehydrogenase	mdh	LSLVRALQGEQGVVECAYVEGDGQYARFFSQ
111.61	125	50S ribosomal protein L6	rplF	SHPVDHQLPAGITAECPTQTEIVLKGADKQV
154.15	239	Transaldolase B	talB	GASFRNIGEILELAGCDRLTIAPALLKELAE
141	328	Cysteine desulfurase IscS	iscS	LIMALKDLAVSSGSACTSASLEPSYVLRALG
131.82	285	Phosphoenolpyruvate carboxykinase [ATP]	pckA	EHGWDDDGVFNFEGGCYAKTIKLSKEAEPEI
110.55	39	Aerobic glycerol-3-phosphate dehydrogenase	glpD	GRGLSVLMLEAQDLACATSSASSKLIHGGLR
131.83	286	Acetate kinase	ackA	TKESGLLGLTEVTSDCRYVEDNYATKEDAKR
109.11	138	60 kDa chaperonin	groL	AVTAAVEELKALSVPCSDSKAIAQVGTISAN
126.83	64	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	ELYLDGNARQNLATFCQTWDDENVHKLMDLS
183	36	Uncharacterized lipoprotein YbjP	ybjP	TVTPAYKDNGTRSGPCVEGGPDNVAQQFYDY
161.17	87	30S ribosomal protein S2	rpsB	TKRAASEAVKDAALSCDQFFVNHRWLGGMLT
105.65	14	Aminomethyltransferase	gcvT	__MAQQTPLYEQHTLCGARMVDFHGWMMPLH
118.52	37	Iron-sulfur cluster assembly scaffold protein IscU	iscU	NNDENVGSGMVGAPACGDVMKLQIKVNDEGI
119.4	110	Cytosol aminopeptidase	pepA	KTINTLNDTGSMEAVCFLTELHVKGRNNYWK
169.58	398	Elongation factor G	fusA	AAAIGLKDVTTGDTLCDPDAPIILERMEFPE
103.22	144	Aerobic glycerol-3-phosphate dehydrogenase	glpD	SVLKPEIKRGFEYSDCWVDDARLVLANAQMV
153.48	126	Triosephosphate isomerase	tpiA	KKFAVLKEQGLTPVLCIGETEAENEAGKTEE
199.76	159	Methionine--tRNA ligase	metG	GTCPKCKSPDQYGDNCEVCGATYSPTELIEP
199.76	162	Methionine--tRNA ligase	metG	PKCKSPDQYGDNCEVCGATYSPTELIEPKSV
115.83	247	Enolase	eno	AAGYELGKDITLAMDCAASEFYKDGKYVLAG
100.04	289	Glyceraldehyde-3-phosphate dehydrogenase A	gapA	YTEDDVVSTDFNGEVCTSVFDAKAGIALNDN
138.54	256	Elongation factor Tu 2;Elongation factor Tu 1	tufB;tufA	EEVEIVGIKETQKSTCTGVEMFRKLLDEGRA
99.011	124	Succinyl-CoA ligase [ADP-forming] subunit alpha	sucD	KVKLDEAGVRMIGPNCPGVITPGECKIGIQP
99.011	124	Succinyl-CoA ligase [ADP-forming] subunit alpha	sucD	KVKLDEAGVRMIGPNCPGVITPGECKIGIQP
99.011	133	Succinyl-CoA ligase [ADP-forming] subunit alpha	sucD	RMIGPNCPGVITPGECKIGIQPGHIHKPGKV
153.24	43	UPF0033 protein YedF	yedF	MPQLKKGEILEVVSDCPQSINNIPLDARNHG

143.31	282	2-hydroxy-3-oxopropionate reductase	garR	LRADGLGTADHSALACYYEKLAKVEVTR___
97.235	17	UPF0033 protein YedF	yedF	KNIVPDYRLDMVGEPCPYPAVATLEAMPQLK
103.5	513	Succinate dehydrogenase flavoprotein subunit	sdhA	RLDDTSSEFNTQRVECLELDNLMETAYATAV
96.982	152	6-phosphogluconolactonase	pgl	DNRTLWVPALKQDRICLFTVSDDGHLVAQDP
158.89	15	Chaperone protein DnaK	dnaK	_MGKIIGIDLGTTNSCVAIMDGTTPRVLENA
111.58	135	Alcohol dehydrogenase YqhD	yqhD	LQTGGKEIKSAIPMGCVLTLPATGSESNAGA
94.465	55	Glutamate synthase [NADPH] large chain	gltB	QHRGAILADGKTGDGCGLLLQKPDRFFRIVA
94.302	454	2-oxoglutarate dehydrogenase E1 component	sucA	DFRNTFKRDVFIDLVCYRRHGHNEADEPSAT
93.598	128	Leucine--tRNA ligase	leuS	KMLGFGYDWSRELATCTPEYYRWEQKFFTEL
185.14	209	Glycerophosphoryl diester phosphodiesterase	glpQ	KKYGYTGKDDKVYLQCFDADELKRIKNELEP
107.45	157	Methionyl-tRNA formyltransferase	fmt	MDVGLDTGDMLYKLSCPITAEDTSGTLYDKL
125.75	212	Ferredoxin--NADP reductase	fpr	TIGLPMNKETSHVMLCGNPQMVRDTQQLLKE
90.653	100	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	IDKEEYPQSAAIDLRCVNMVADLWHAPAPKN
90.434	90	Glutamine synthetase	glnA	AVIDPFFADSTLIIRCDILEPGTLQGYDRDP
90.259	367	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	YLADEIAKLGPYEFICTGRPDEGIPAVCFKL
90.259	379	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	EFICTGRPDEGIPAVCFKLKDGEDPGYTLYD
142.35	70	Tryptophan--tRNA ligase	trpS	KLRKATLDTLALYLACGIDPEKSTIFVQSHV
92.906	45	Dihydrolipoyl dehydrogenase	lpdA	LETVIVERYNTLGGVCLNVGCIPSKALLHVA
92.906	50	Dihydrolipoyl dehydrogenase	lpdA	VERYNTLGGVCLNVGCIPSKALLHVAKVIEE
109.81	324	Rod shape-determining protein MreB	mreB	EETGIPVVVAEDPLTCVARGGGKALEMIDMH
239.38	47	ATP synthase subunit alpha	atpA	VSVSDGVIRIHGLADCMQGEMISLPGNRYAI
88.021	259	Respiratory nitrate reductase 1 beta chain	narH	CIFCYPRIEAGQPTVCSETCVGRIRYLGVLL
88.021	263	Respiratory nitrate reductase 1 beta chain	narH	YPRIEAGQPTVCSETCVGRIRYLGVLLYDAD
149.16	136	Citrate synthase	gltA	LFHAFRRDSHPMAVMCGITGALAAFYHDSLD
135.86	207	Molecular chaperone Hsp31 and glyoxalase 3	hchA	LALRHGDNPLNGYSICAFPDAADKQTPEIGY
127.42	47	Deoxyribose-phosphate aldolase	deoC	LCHQAKTPVGNTAAICIYPRFIPIARKTLKE
160.04	234	2-dehydro-3-deoxygluconokinase	kdgK	NAGVKEVVVKRGADSCLVSIAGEGLVDVPAV

172.92	841	Aldehyde-alcohol dehydrogenase;Alcohol dehydrogenase;Acetaldehyde dehydrogenase [acetylating];Pyruvate-formate-lyase deactivase	adhE	LANVDKLSEDAFDDQCTGANPRYPLISELKQ
86.879	60	Succinate dehydrogenase iron-sulfur subunit	sdhB	KDPSLSFRRSCREGVCGSDGLNMNGKNGLAC
85.745	118	Glucosamine-6-phosphate deaminase	nagB	NINLLNGNAPDIDAECRQYEEKIRSYGKIHL
113.63	132	NADH-quinone oxidoreductase subunit E	nuoE	PGQTTFDGRFTLLPTCCLGNCDKGPNMMIDE
113.63	133	NADH-quinone oxidoreductase subunit E	nuoE	GQTTFDGRFTLLPTCCLGNCDKGPNMMIDED
113.63	137	NADH-quinone oxidoreductase subunit E	nuoE	FDGRFTLLPTCCLGNCDKGPNMMIDEDTHAH
88.294	113	Protein-export protein SecB	secB	LGAYCPNILFPYARECITSMVSRGTFPQLNL
84.198	239	Glucosamine-6-phosphate deaminase	nagB	WTISCLQLHPKAIMVCDEPSTMELKVKTLRY
129.42	36	Uncharacterized protein YmgD	ymgD	DYNIKNGLPSETYITCAEANEMAKTDSAQVA
95.195	64	Peptide methionine sulfoxide reductase MsrB	msrB	CDAPLFHSQTKYDSGCGWPSFYEPVSEESIR
83.54	139	33 kDa chaperonin	hslO	RYQGVVGLEGDTLAACLEDYFMRSEQLPTRL
83	11	Acetate kinase	ackA	_____MSSKLVLVLNCGSSSLKFAIIDAVNG
92.935	391	Glucose-1-phosphatase	agp	TLQAPAQRVTLELSGCPIDADGFCPMDKFDS
92.935	399	Glucose-1-phosphatase	agp	VTLELSGCPIDADGFCPMDKFDSVLNEAVK_
146.36	180	NADH-quinone oxidoreductase subunit F	nuoF	FDFELFVHTGAGRYICGEETALINSLEGRRA
87.138	130	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	NGQAVGTNTIGSSEACMLGGMAMKWRWRKRM
80.706	436	Glutamine--tRNA ligase	glnS	ERVEKDAEGNITTIFCTYDADTLSKDPADGR
83.729	20	1-phosphofructokinase	fruK	VATITLNPAYDLVGFCPEIERGEVNLVKTTG
119.45	167	Transaldolase A;Transaldolase B	talA;talB	NCNLTLLFSFAQARACAEAGVFLISPFVGRI
92.457	99	Porphobilinogen deaminase	hemC	DVPVEFPQGLGLVTICEREDPRDAFVSNNYD
79.974	343	Peptidase T	pepT	EHPHILDIAQQAMRDCDIEPELKPIRGGTDG
160.57	87	Acid stress chaperone HdeA	hdeA	DVQGIATVTPAIVQACTQDKQANFKDKVKGE
118.51	150	Glyceraldehyde-3-phosphate dehydrogenase A	gapA	NFDKYAGQDIVSNASCTTNCLAPLAKVINDN
118.51	150	Glyceraldehyde-3-phosphate dehydrogenase A	gapA	NFDKYAGQDIVSNASCTTNCLAPLAKVINDN
118.51	154	Glyceraldehyde-3-phosphate dehydrogenase A	gapA	YAGQDIVSNASCTTNCLAPLAKVINDNFGII
77.305	147	Pantothenate synthetase	panC	TIVSKLFNLVQPDIACFGEKDFQQLALIRKM
76.713	519	60 kDa chaperonin	groL	LQYAASVAGLMITTECMVTDLPKNDAADLGA

76.588	369	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	FRQIEESGADLVVTDCETCKWQIEMSTSLRC
76.588	372	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	IEESGADLVVTDCETCKWQIEMSTSLRCEHP
197.3	27	D-ribose pyranase	rbsD	SSVISRLGHTDTLVVCDAGLPIPKSTTRIDM
93.959	89	Succinate dehydrogenase flavoprotein subunit	sdhA	KGSDYIGDQDAIEYMCKTGPEAILELEHMGL
124.25	84	Gamma-glutamyl phosphate reductase	proA	TPARLKGIADDVRQVCNLADPVGQVIDGGVL
75.416	79	Dihydroorotase	pyrC	LDAVPAGHDFTPLMTCYLTDSLDPNELERGF
75.197	42	2Fe-2S ferredoxin	fdx	ILDAALRNGIEIEHACEKSCACTTCHCIVRE
74.76	543	Fumarate hydratase class I, aerobic	fumA	DDKGNDFFQQIQLTQCTRCVK_____
106.39	132	Toxin-antitoxin biofilm protein TabA	tabA	GDFVVFYPGEVHKPLCAVGAPAQVRKAVVKM
119.54	36	Triosephosphate isomerase	tpiA	ELVSNLRKELAGVAGCAVAIAPPEMYIDMAK
76.198	36	Acetate kinase	ackA	IDAVNGEEYLSGLAECFHLPEARIKWKMDGN
160.75	39	Acid stress chaperone HdeB	hdeB	AQSALAANESAKDMTCQEFIDLNPKAMTPVA
76.823	101	Osmotically-inducible lipoprotein E	osmE	ALDDTGHVINSGYQTCAEYDTDPQAAK___
114.87	266	Elongation factor G	fusA	ALRQRVLNNEIILVTCGSAFKNKGVQAMLDA
71.451	27	30S ribosomal protein S12	rpsL	RARKVAKSNVPALEACPQKRGVCTRVYTTTP
71.268	480	NADH-quinone oxidoreductase subunit C/D	nuoC	LKVEELRQSLRILEQCLNNMPEGPFKADHPL
71.267	547	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	aceF	KARDGKLTAGEMQGGCFTISSIGGLGTTHFA
156.8	68	Serine hydroxymethyltransferase	glyA	NKYAEGYPGKRYYGGCEYVDIVEQLAIDRAK
94.614	330	Glycerol dehydrogenase	gldA	KEDVPAKMRIVAEAACAEGETIHNMPGGATP
121.5	360	Cytosol aminopeptidase	pepA	TVEVLNTDAEGRLVLCDVLTYVERFEPEAVI
148.48	458	60 kDa chaperonin	groL	ALRAMEAPLRQIVLNCGEEPSVVANTVKGGD
139.48	61	Thiol peroxidase	tpx	KRKVLNIFPSIDTGVCAASVRKFNQLATEID
89.629	85	UPF0482 protein YnfB	ynfB	KEWDKADAAFDNRDKCEQSANINAYWEPNTL
68.83	512	Glycine dehydrogenase (decarboxylating)	gcvP	RKDLALNQAMIPLGSCTMKLNAAAEMIPITW
229.9	387	Serine hydroxymethyltransferase	glyA	RGFKEAEAKELAGWMCDVLDSINDEAVIERI
85.377	289	Glutathione synthetase	gshB	IIGDRLTEINVTSPTCIREIEAEFPVSITGM
96.303	169	6-phosphogluconate dehydrogenase, decarboxylating	gnd	PILTKIAAVAEDGEPCVTYIGADGAGHYVKM
92.589	83	Periplasmic serine endoprotease DegP	degP	MPRNFQQFFGDDSPFCQEGSPFQSSPFCQGG
92.589	95	Periplasmic serine endoprotease DegP	degP	SPFCQEGSPFQSSPFCQGGQGGNGGGQQQKF

64.984	130	Periplasmic AppA protein;Phosphoanhydride phosphohydrolase;4-phytase	appA	TRKTGEAFAAGLAPDCAITVHTQADTSSPDP
99.174	359	Serine--tRNA ligase	serS	EVWIPAQNTYREISSCSNVWDFQARRMQARC
73.156	525	Formate acetyltransferase 1	pflB	SLMALHDRDVIRTMACGIAGLSVAADSLSAI
107.45	162	Phosphoglycerate kinase	pgk	AQASTHGIGKFADVACAGPLLAAELDALGKA
99.566	198	Cytosol aminopeptidase	pepA	GIKAAKDLGNMPPNICNAAYLASQARQLADS
64.685	124	Glycerol dehydrogenase	gldA	VPVAIAPTIASTDAPCSALSVIYTDEGEFDR
62.924	288	ATP-dependent protease ATPase subunit HslU	hslU	SREGVQRDLLPLVEGCTVSTKHGMVKTDHIL
62.594	182	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	dapD	VLEPLQANPTIIEDNCFIGARSEVVEGVIVE
73.616	186	Trehalose-6-phosphate hydrolase	treC	LNWENPAVRAELKKVCEFWADRGVDGLRLDV
61.691	278	D-tagatose-1,6-bisphosphate aldolase subunit GatY	gatY	AKSAMRDVVSKVIADCGCEGRA_____
61.691	280	D-tagatose-1,6-bisphosphate aldolase subunit GatY	gatY	SAMRDVVSKVIADCGCEGRA_____
96.331	281	Tryptophanase	tnaA	MSAKKDAMVPMGGLLCMKDDSFFDVYTECRT
94.09	71	Glutaminase 1	glsA1	GDSDYRFALESISKVCTLALALEDVGPQAVQ
162.61	383	Tryptophanase	tnaA	LPHIPADQFPAQALACELYKVAGIRAVEIGS
140.17	176	DNA-directed RNA polymerase subunit alpha	rpoA	SEEDERPIGRLLVDACYSPVERIAYNVEAAR
147.2	323	Outer membrane protein A	ompA	GNTCDNVKQRAALIDCLAPDRRVEIEVKGIK
107.65	138	Elongation factor Tu 2;Elongation factor Tu 1	tufB;tufA	GRQVGVPYIIVFLNKCDMVDDEELLELVEME
79.458	13	2Fe-2S ferredoxin	fdx	___MPKIVILPHQDLCPDGAVLEANSGETIL
154.55	243	Delta-aminolevulinic acid dehydratase	hemB	EAIRESLLDEAQGADCLMVKPAGAYLDIVRE
108.7	21	DNA-binding protein H-NS	hns	KILNNIRTLRAQARECTLETLEEMLEKLEVV
109.66	256	Glycerol kinase	glpK	SGIAGDQQAALFGQLCVKEGMAKNTYGTGCF
57.278	95	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase	kduI	VINIGGAGTITVDGQCYEIGHRDALYVGKGA
82.426	135	RNA polymerase-binding transcription factor DksA	dksA	EIGIRRLEARPTADLCIDCKTLAEIREKQMA
82.426	138	RNA polymerase-binding transcription factor DksA	dksA	IRRLEARPTADLCIDCKTLAEIREKQMAG__
66.267	11	2-dehydro-3-deoxygluconokinase	kdgK	_____MSKKIAVIGECMIELSEKGADVKRGF
76.228	140	Lysine--tRNA ligase	lysS	KGKLFKTKTGELSIHCTELRLLTKALRPLPD
74.415	125	Uncharacterized protein YccU	yccU	AVLARDAGLNVVMDRCPAIEIPRLGLAK___
173.03	39	Acid stress chaperone HdeA	hdeA	AQKAADNKKPVNSWTCEDFLAVDESFQPTAV

96.342	340	Aspartate ammonia-lyase	aspA	PAKVNPVVPEVVNQVCFKVIGNDTTVTMAAE
92.551	323	Cytosol aminopeptidase	pepA	AELQLPINVIGVLAGCENMPGGRAYRPGDVL
61.11	388	NADH-quinone oxidoreductase subunit F	nuoF	RGEGQPGDIETLEQLCRFLGPGKTFCAHAPG
149.94	18	50S ribosomal protein L31	rpmE	KDIHPKYEEITASCSCGNVMKIRSTVGHDLN
149.94	16	50S ribosomal protein L31	rpmE	MKKDIHPKYEEITASCSCGNVMKIRSTVGHD
149.94	18	50S ribosomal protein L31	rpmE	KDIHPKYEEITASCSCGNVMKIRSTVGHDLN
53.967	30	ATP-dependent 6-phosphofructokinase isozyme 2	pfkB	SATITPQIYPEGKLRCTAPVFEPGGGGINVA
89.247	384	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	CETCKWQIEMSTSLRCEHPITLLAQALA___
115.77	79	Transcriptional repressor MprA	mprA	LESQENHSIQPSELSCALGSSRTNATRIADE
151.41	37	Fructose-bisphosphate aldolase class 2	fbaA	FQVAKENNFALPAVNCVGTDSINAVLETAAK
63.908	463	Isoleucine--tRNA ligase	ileS	GQARIESMVANRPDWCISRQRTWGVPMSLFV
93.29	65	Uridine phosphorylase	udp	FTTWRAELDGKPVIVCSTGIGGPSTSIAVEE
99.305	61	50S ribosomal protein L35	rpmI	PKAMVSKGDLGLVIACLPYA_____
69.451	502	Phosphoenolpyruvate-protein phosphotransferase	ptsI	QVIDASHAEGKWTGMCGELAGDERATLLLLG
51.286	73	Transaldolase A	talA	GKKNGKTQEQQVVAACDKLAVNFGAEILKIV
51.211	216	Ribose-phosphate pyrophosphokinase	prs	VSQVMHIIGDVAGRDCVLVDDMIDTGGTLCK
96.23	171	Putative GTP cyclohydrolase 1 type 2	ybgI	WCGDTGPEVVQRVAWCTGGGQSFIDSAARFG
66.267	285	Lactaldehyde dehydrogenase	aldA	VKAIVDSRVINSGQVCNCAERVYVQKGIYDQ
66.267	287	Lactaldehyde dehydrogenase	aldA	AIVDSRVINSGQVCNCAERVYVQKGIYDQFV
162.49	37	50S ribosomal protein L31	rpmE	MKIRSTVGHDLNLDVCSKCHPFFTGKQRDVA
151.7	121	30S ribosomal protein S11	rpsK	RITNITDVTPIPHNGCRPPKKRRV_____
71.888	109	Universal stress protein A	uspA	VLVDAIKKYDMDLVVCGHHQDFWSKLMSSAR
63.29	50	Methionine--tRNA ligase	metG	WVRYQRMRGHEVNFICADDAHGTPIMLKAQQ
167.28	209	Acetate kinase	ackA	AKMLNKPVEELNIITCHLGNGGSVSAIRNGK
60.209	104	Adenylosuccinate synthetase	purA	DRGIPVRERLLLSEACPLILDYHVALDNARE
47.849	249	Lactaldehyde dehydrogenase	aldA	GEKIMATAAKNITKVCLELGGKAPAIVMDDA
66.664	82	Elongation factor Tu 2;Elongation factor Tu 1	tufB;tufA	HVEYDTPTRHYAHVDCPGHADYVKNMITGAA
68.49	452	Cytosol aminopeptidase	pepA	MANIGGRPGGAITAGCFLSRFTRKYNWAHLD
64.27	244	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	AVGVKGIPVIATSSTCTFALRDEYPEVLNVD
45.864	225	Acetate kinase	ackA	HLGNGGSVSAIRNGKCVDTSMGLTPLEGLVM
60.474	78	Elongation factor Ts	tsf	KTKIDGNYGIILEVNCQTDFVAKDAGFQAFA

74.047	113	Malate dehydrogenase	mdh	VKNLVQQVAKTCPKACIGIITNPVNTTVAIA
86.756	38	Phosphopentomutase	deoB	FGDVGADTLGHIAEACAKGEADNGRKGPLNL
58.477	54	DNA-directed RNA polymerase subunit alpha	rpoA	LGNALRRILLSSMPGCAVTEVEIDGVLHEYS
104.15	132	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	mtnN	AGFKADDKLIAAAEACIAELNLNAVRGLIVS
84.658	219	Aminomethyltransferase	gcvT	AADFWRALVEAGVKPCGLGARDTLRLEAGMN
85.355	12	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	____MNDTSFENCIKCTVCTTACPVSRVNPG
85.355	15	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	_MNDTSFENCIKCTVCTTACPVSRVNPGYPG
85.355	19	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	glpC	TSFENCIKCTVCTTACPVSRVNPGYPGPKQA
44.367	144	Chaperone protein DnaJ	dnaJ	VRGVTKEIRIPTLEECDVCHGSGAKPGTQPQ
44.367	147	Chaperone protein DnaJ	dnaJ	VTKEIRIPTLEECDVCHGSGAKPGTQPQTCP
166.28	352	Tryptophanase	tnaA	QVQYLVDGLEEIGVVCQQAGGHAAFVDAGKL
122.01	147	Malonyl CoA-acyl carrier protein transacylase	fabD	MAAIIGLDDASIAKACEEAAEGQVVSPVNFN
97.452	140	Lysine--tRNA ligase, heat inducible	lysU	RGTLFKTQTGELSIHCTELRLLTKALRPLPD
49.448	137	Glutathione S-transferase GstB	gstB	PEERDQAAIDASCKECDALFALLDAELAKVK
138.78	413	Tryptophanase	tnaA	SFLLGRDPKTGKQLPCPAELLRLTIPRATYT
142.43	260	Pyruvate dehydrogenase E1 component	aceE	IATREKLDNLVFVINCNLQRLDGPVTGNGKI
290.37	298	Tryptophanase	tnaA	KDDSFFDVYTECRTLCVVQEGFPTYGGLEGG
196.45	129	D-tagatose-1,6-bisphosphate aldolase subunit GatZ	gatZ	VRAGFSKIHLDASMSCAGDPIPLAPETVAER
94.309	71	50S ribosomal protein L10	rplJ	RNTLLRRAVEGTPFECLKDAFVGPTLIAYSM
62.183	198	Glutamate decarboxylase beta;Glutamate decarboxylase alpha	gadB;gadA	RPGQLFMDPKRMIEACDENTIGVVPTFGVTY

APPENDIX II

STEP BY STEP PROTEOMICS DATA ANALYSIS INSTRUCTIONS

Data conversion instructions using MSconvertGUI

1- Download the tool through the website:

http://proteowizard.sourceforge.net/download.html

2- Navigate to the MSconvert app and start it.

3- Browse and select the raw files (.d) and select all the files you want to convert.

4- Select the output format (mzML, mzXML, mgf) depending on the downstream PSM tool intended to be used.

5- Use Peak Picking filter: Vendor. MS levels 1-2.

6- Chose Binary encoding precision: 32-bit.

7- Use the recommended options as follow: write index: yes, TPP compatibility: yes.

8- Click start. The output files should be at the exact location where the raw files are.

SearchGUI/Peptideshaker platform

1- Download the apps through the website:

http://compomics.github.io/projects/searchgui.html and

http://compomics.github.io/projects/peptide-shaker.html

2- Start SearchGUI and import the mass spec raw files converted to mgf file format.

3- Create a new search setting: i) select the database file to search against. This file needs to be in FASTA file format. ii) add Carbamidomethylation of C and Oxidation of M as variable modifications. iii) set the precursor and fragment m/z tolerances to 20 PPM.

4- Select the output location

5- Specify the search engine(s) needed for the analysis. Usually X!Tandem is great. You may want to include more than one search engine for specialized tasks like PTMs.

6- Enable PeptideShaker and name the output.

7- Start the search.

MetaMorpheus platform

Download MetaMorpheus from the following link:

https://github.com/smith-chem-wisc/MetaMorpheus

1- After the installation, start the application.

2- Select the database FASTA file to search against. You can add multiple database files to search against and also you can add a contaminants database (optional).

3- Add the mass spec raw files in mzML file format. You can add multiple file.

4- The first task should be the Calibration task. This task is re-calibrate the raw files.

5- The second task is the Discover PTMs. This task is to search perform the open-search. In this search, a big precursor tolerance is used to find potential PTMs for every peptide.

6- The third task is the main search. The output location will be at the same location as the raw files.

MaxQuant platform

1- Download MaxQuant from the following website:

https://maxquant.org/

2- Start the application.

3- Load the raw files - no data conversion needed.

4- After loading the data, go ahead and specify a name for every tested file by specifying an experiment name (e.g. rep01, rep02 … etc.)

5- Under Group-specific parameters:

a. Select the following modification as variable: Carbamidomethyl (C) and Oxidation (M).

b. Select LFQ under Label-free quantification. This is optional if protein quantification is required.

6- Under Global parameters: add the database FASTA file and specify the identify rule according the FASTA file source (Uniprot or NCBI).

7- Set the number of processor depending on the computer using specification - usually 8 – 12 processors are selected.

8- Start the analysis.