Exploring the Effects of Human Cytomegalovirus Infection on Host Cell Metabolism: A Proteomic Approach

Zachary S. Predmore

Submitted in Partial Fulfillment
of the Requirements of the
Degree of
Bachelor of Arts

To the Department of Chemistry
of
Princeton University

April 18, 2011

© Copyright by Zachary Scott Predmore, 2011. All rights reserved.

I hereby declare that I am the sole author of this thesis, and that this thesis represents my own work in accordance with University regulations.

_____

Signature

Abstract

Human cytomegalovirus (HCMV) is a major public health concern; it is estimated to latently infect a majority of the world's population. Though infections are usually minor, HCMV can cause serious problems in the immunocompromised and is a leading cause of birth defects. There are several antivirals approved to treat HCMV currently on the market, but viral strains resistant to all of the existing drugs have been reported. Since all of the drugs target some aspect of nucleic acid metabolism, future antiviral development could benefit from targeting other aspects of the host-pathogen interaction.

This study is the first comprehensive analysis of the proteome of cells infected with HCMV. A SILAC-based proteomic strategy was used to identify proteins that change in abundance upon HCMV infection, with the goal of examining metabolic enzymes. Metabolic enzymes generally increased by between 1.5- and 4-fold under infection.

Changes in protein abundance were compared to data obtained previously detailing HCMV infection's impact on mRNA transcript levels and changes in metabolic flux. While no correlation was found between mRNA transcript level and protein level, changes in flux through glycolysis suggest that HCMV increases flux in this pathway at least in part by increasing the levels of the enzymes of the pathway. However, flux changes in the TCA cycle cannot be accounted for based solely on changes in protein level, suggesting the virus modifies these metabolic reactions in a different manner. Analysis suggests increased flux through the pyruvate dehydrogenase complex can be partially explained by a combination of increased abundance of enzyme and altered phosphorylation patterns resulting from infection.

Executive Summary (adapted from this paper's abstract)

Human cytomegalovirus (HCMV) is a major public health concern; it is estimated to latently infect a majority of the world's population. Though infections are usually minor, HCMV can cause serious problems in the immunocompromised and is a leading cause of birth defects. When HCMV infects a human cell, it effectively turns the cell into a "virus-producing factory", taking control of host cell metabolic processes and diverting host cell resources towards the production of more virus. Since proteins are responsible for the chemical changes of normal metabolism, the altered metabolism seen in cells infected with HCMV likely has its basis in altered levels of various metabolic proteins, or enzymes. This study seeks to identify these changes in protein level and link them to corresponding changes in metabolite flux, with the ultimate aim of developing new treatment strategies for HCMV.

There are several antivirals approved to treat HCMV currently on the market, but viral strains resistant to all of the existing drugs have been reported. Since all of these drugs target some aspect of nucleic acid metabolism, future antiviral development could benefit from targeting other metabolic aspects of the host-pathogen interaction, potentially arming doctors with combination therapies similar to those used to treat HIV/AIDS.

This study is the first comprehensive analysis of the proteome of cells infected with HCMV. A SILAC-based proteomic strategy was used to identify host cell proteins that change in abundance upon HCMV infection, with the goal of examining the changes in metabolic enzymes as a result of HCMV infection. Many metabolic proteins were quantified and generally increased by between 1.5- and 4-fold under infection, suggesting that the metabolic changes seen in cells infected with HCMV have some basis in changes in enzyme level.

Changes in protein abundance were compared to data obtained previously detailing HCMV infection's impact on mRNA transcript levels and changes in metabolic flux. While no correlation was found between mRNA transcript level and protein level, changes in flux through glycolysis suggest that HCMV increases flux in this pathway at least in part by inducing protein synthesis and increasing the levels of the enzymes of the pathway. However, flux changes in the TCA cycle cannot be accounted for based solely on changes in protein level, suggesting the virus modifies these metabolic reactions in a different manner. Additional analysis accomplished through the use of Western blots suggests increased flux through one specific enzyme, the pyruvate dehydrogenase complex, can be partially explained by a combination of increased abundance of enzyme and altered phosphorylation patterns resulting from infection.

This study documents the changes in the proteome that accompany HCMV infection, and explores a novel avenue for HCMV treatment, identifying several possible routes for future research. It also highlights the importance of scientific research in improving global health – every pharmaceutical development has its roots in laboratory science.

Table of Contents

Abstract iv

Executive Summary v

Table of Contents vii

List of Figures ix

List of Tables x

Introduction 1

- Introduction to human cytomegalovirus 1
- Current antiviral therapies 2
- Virus Infection and Metabolism 5
- What we already know about the -omics of HCMV infection 6
- Protein mass spectrometry 12
- Mass spectrometry-based proteome studies of viruses 15
- Proteomics and HCMV 18
- This study 21
- Mass Spectrometry Sample Preparation and Data Acquisition 22
  - Biological reagents and stable isotope labeling 22
  - Infections 22
  - Lysis 22
  - Sample processing 23
  - SCX Fractionation 24
  - Chromatography 24
  - Ionization and MS/MS 24
  - Data Analysis 26
- Gene Ontology and Enrichment Analysis 27
- SDS PAGE and Western Blots 27
  - Biological reagents 27
  - Infections 27
  - Protein harvesting and Western blotting 28
  - Antibodies 29

Results 30
Results of the SILAC screen 30
Quantification of sample peptides 32
Changes in protein level across the proteome 35
Changes in the levels of enzymes in key pathways 40
Changes in levels of subunits of multisubunit complexes 43
Gene Ontology Enrichment Analysis 44
Western blots probing different phosphorylation states of pyruvate dehydrogenase 48
Discussion 50
Comparison of protein quantifications with mRNA transcript level 51
Comparison of protein quantifications with flux changes 54
Limitations of quantification with MaxQuant ("All-or-nothing" peptides) 57
Proteins in multisubunit complexes 62
Gene Ontology and Enrichment Analysis 63
Comparing the results of this study with those for other viruses 65
Future Work 66
Conclusion 68
Appendix 69
Relationship between posterior error probability and signal intensity 69
Metabolic proteins identified by the SILAC screen 70

List of Figures

Figure 1 – Labeling Cells and Preparing Peptides (D. Perlman) 23

Figure 2 – Peptide-Level SCX Chromatography and LC-MS (D. Perlman) 25

Figure 3 – Quantitative SILAC MS Data Analysis (D. Perlman) 26

Figure 4 – Number of proteins quantified for each peptide count 31

Figure 5 – Total ion chromatogram for a peptide from pyruvate kinase 33

Figure 6 – Total ion chromatograms for a peptide used to quantify AHNAK 34

Figure 7 – Changes in abundance of host cell proteins as a result of HCMV infection 35

Figure 8 – Relationship of L/H ratios to intensity for proteins detected by the screen 36

Figure 9 – Virus proteins identified by the screen 37

Figure 10 – Metabolic enzymes identified by the screen 38

Figure 11 – Changes in abundance of glycolytic enzymes and glycolytic metabolites (from Munger et al. 2006) (increased abundance under infection indicated by brighter red color) 41

Figure 12 – Changes in abundance of enzymes and of the TCA cycle, fatty acid synthesis, and several anapleurotic reactions. Changes in metabolite concentrations from Munger et al. 2006. 42

Figure 13 – Western blot for phosphorylated pyruvate dehydrogenase 48

Figure 14 – Changes in mRNA transcript level compared to changes in protein level 52

Figure 15 – Ratio of changes in flux to change in enzyme level during infection. 54

Figure 16 – Intensities of the heavy and light signals for proteins using an incorrect mass shift for arginine (Δ6 instead of Δ10) 58

Figure 17 – Intensities of the heavy and light signals for proteins using the correct mass shift for arginine (Δ10) 59

Figure 18 – Quantification of a peptide from pp65 (a viral structural protein) through SILAC .. 61

Figure 19 – Relationship between Posterior Error Probability (PEP) and Intensities of Heavy and Light Signals 69

List of Tables

Table 1 – Ten human proteins with greatest observed increases in abundance under infection .. 39

Table 2 – Ten human proteins with greatest observed decreases in abundance under infection.. 39

Table 3 - Changes in abundance of the subunits of ATP synthase 43

Table 4 - Changes in the abundance of the subunits of isocitrate dehydrogenase 44

Table 5 – GO Terms enriched in the upregulated fraction 46

Table 6 – GO Terms enriched in the downregulated fraction 47

Introduction

Introduction to human cytomegalovirus

Human cytomegalovirus (HCMV) is a human beta herpesvirus[1]. Epidemiological studies estimate that HCMV latently infects between 50-90% of people worldwide, with the fraction of the population infected increasing with age[1, 2]. Several recent studies have suggested the percentage of those infected is actually lower in certain regions of the world[1]. The virus's mode of transmission is currently unknown, though it is believed to be transmitted through bodily fluids. Primary infection by HCMV is typically mild; infected persons often present with a fever and swollen glands, if symptomatic at all[3]. HCMV infections have two phases: lytic and latent. After the primary infection, the virus establishes latency in cells of the myeloid lineage through a poorly understood mechanism[1]. The virus can remain latent for years, lying dormant in a host cell. While latent, the virus produces very few viral gene products and releases no virions. The exact mechanism of reactivation from latency is also not known; studies have suggested that immunosuppression, infection, stress, or inflammation can trigger the process[4-6].

Reactivation of a latent HCMV infection can be serious. In immunocomptent people, reactivation can cause mononucleosis and can possibly lead to atherosclerosis[7]. However, HCMV reactivation is particularly dangerous for the immunocompromised; the virus can cause retinitis, colitis, hepatitis, and several other dangerous conditions[1]. As a result, serious infection resulting from HCMV reactivation is a frequent complication encountered by organ transplant patients taking immunosuppressive drugs or patients also infected with HIV[1]. Additionally, the virus is of concern to pregnant women and can cause birth defects such as hearing loss if transmitted to the developing fetus[8].

The virus is structurally similar to all other herpesviruses. At its core, the virion consists of a double stranded DNA genome inside an icosahedral protein capsid[1]. A layer of protein tegument surrounds the capsid and a lipid envelope encloses the tegument[9].

The HCMV life cycle has been studied and many of its mechanisms have been elucidated to a great degree of detail. First, the virus enters a host cell through one of two distinct mechanisms. In some cases, the virion binds host cell receptors and fuses its envelope with the host cell membrane, pushing the capsid inside the cell[10]. However, in certain cell types, the virus enters when it is endocytosed by the host cell[10]. After entering the cell, the viral capsid travels to the nucleus and releases its DNA. HCMV replication is a complex and coordinated process with many temporally well-defined steps. The nearly 200 proteins produced by the virus fall into three distinct classes: immediate-early (IE), early, and late[9]. Within two hours of viral entry, transcription of mRNA encoding immediate early proteins begins[11]. The immediate early proteins are all transactivators and coordinate the transcription of the early and late HCMV genes[11]. Early genes are expressed without viral DNA replication and generally encode proteins that modify host cell functions[12]. Late genes begin to be expressed after early genes since late gene expression requires viral DNA synthesis[12]. Typically first seen 24-48 hours post infection, proteins encoded by late genes have a variety of functions but include many structural proteins involved in virion assembly[12]. After assembling, the newly-created virion exits the cell and is ready to infect other cells, typically between 72-96 hours post infection[12].

Current antiviral therapies

Given HCMV's high prevalence and damaging effects on immunocompromised persons, several HCMV therapeutics are approved by the FDA. Most of these antiviral agents are

nucleoside analogues, a pattern also observed for many other viruses. Nucleoside analogues have been particularly effective as antiviral agents due to their ability to specifically inhibit viral DNA polymerases or reverse transcriptases, allowing these drugs to discriminate between enzymes essential for virus replication and those human enzymes necessary for normal metabolic function.

The first drug approved to treat HCMV infection, the guanine analogue ganciclovir, is currently the first line of treatment for the disease[13]. Once the drug has entered the bloodstream, it can cross host cell membrane and be absorbed into a cell. Inside an infected cell, ganciclovir is triply phosphorylated by viral and cellular kinases[13]. The newly phosphorylated compound can then be incorporated into the viral genome. Once ganciclovir is added to a growing strand of DNA, the strand is not able to react any further, halting DNA chain elongation and limiting virion production. However, pharmacokinetic studies revealed the low oral bioavailability of ganciclovir (5-8%) and prompted the development of valganciclovir, the valyl ester of ganciclovir[14]. This prodrug, which is metabolized into ganciclovir by nonspecific enzymes in the liver, similarly inhibits the viral DNA polymerase but has a much higher oral bioavailablity (60%)[14].

Foscarnet is another drug approved for treatment of HCMV, initially approved to treat HCMV retinitis in HIV/AIDS patients[15]. A much smaller molecule than ganciclovir or valganciclovir, foscarnet is not a nucleoside analogue, but instead is a pyrophosphate analogue and inhibits the viral DNA polymerase by preventing the cleavage of diphosphate from nucleoside triphosphates[16]. This cleavage is necessary for continued polymerization of the chain. Foscarnet is more toxic than ganciclovir, and as a result is used primarily in patients with ganciclovir-resistant HCMV[15].

A fourth antiviral drug, cidofovir, is approved to treat HCMV retinitis in HIV/AIDS patients[17]. Cidofovir is a nucleoside analogue and also inhibits the viral DNA polymerase, though unlike other nucleoside analogues, while it needs to be phosphorylated by cellular kinases it does not need to be phosphorylated by viral enzymes[17]. However, as with foscarnet, cidofovir has shown considerable toxicity in clinical trials, and is typically reserved for the treatment of strains of the virus that are resistant to the other, more easily tolerated antivirals[17].

The final antiviral drug approved for treatment of HCMV is fomivirsen[18]. Unlike several of the other drugs discussed, fomivirsen is not a nucleoside analogue, but rather is an antisense oligonucleotide 21-mer with phosphothionate linkages[18]. Instead of inhibiting the viral DNA polymerase, fomivirsen binds mRNA transcribed from an immediate early gene, preventing ribosomes from translating the sequences into protein[18]. This interrupts the cascade of regulation by the immediate early gene products and prevents viral replication. The phosphothionate linkages prevent degradation of the drug, giving it a longer half-life within the cell than a typical oligonucleotide, which would be digested by nucleases[18].

Though these drugs have proven effective in treating HCMV and many of its different symptoms, antiviral-resistant strains of HCMV are common and have been observed following all forms of organ transplantation[15]. Multidrug resistance makes certain strains of HCMV particularly difficult to treat, and necessitates the constant need for the development of new antiviral agents. However, all of the currently approved HCMV drugs target some aspect of virus nucleic acid metabolism, whether they inhibit the viral DNA polymerase or prevent mRNA from being translated into protein. This narrowly focused approach to antiviral drug development limits the scope of drug discoverers and decreases the size of the pool of possible antiviral agents.

Though the drug discovery industry frequently makes use of high-throughput techniques to rapidly screen large libraries of molecules for potential drug leads, there is greater potential for the high-throughput approaches of systems biology in drug development. By studying the host-virus interaction using a combination of transcriptomics, metabolomics, and proteomics, drug developers may be able to go beyond simply uncovering new ways to inhibit the viral DNA polymerase and instead identify novel targets for antiviral molecules. Antiviral treatments targeted against multiple viral enzymes or processes may be used in combination with existing drugs that target other aspects of viral infection to create highly effective combination therapies and greatly reduce the possibility of the emergence of drug resistant virus strains.

Virus Infection and Metabolism

Virus infection represents an ideal subject for these types of systems biology experiments, specifically metabolic profiling experiments. Viruses are obligate parasites; they cannot reproduce without a host cell since all of the biomolecules required to produce virions are taken from the host cell. As a result, many viruses have evolved complex control mechanisms that allow them to hijack host cellular processes to use them for their own nefarious ends[19]. Metabolism is a major host cell process, and is no stranger to viral hijacking. When a virus infects a host cell, it drastically alters cellular metabolism in an effort to maximize the production of biomolecules necessary for the production of new infectious virions[19]. As a result, viral infection can cause measurable changes in metabolic flux and drastically alter levels of a host cell's biomolecules. If it were possible to identify key aspects of the virus's control over host cell metabolism, antiviral therapies that interfere with this control program could be used as antiviral agents. Some drug developers have started to incorporate these systems biology approaches to

target selection and have seen success. For example, several fatty acid elongases, which convert 16-carbon fatty acids produced by fatty acid synthase into long chain fatty acids with 18 or more carbon molecules, were implicated as important for virion production[20, 21]. As a result, inhibitors of these elongases have recently been patented and are being tested as potential antiviral agents in humans.

This treatment strategy targeting fatty acid elongases represents a radical departure from the traditional nucleoside analogue drugs that dominate the market for HCMV treatment. Further studies into the biochemistry of HCMV infection and the effect of HCMV infection on host cell metabolic enzymes could potentially lead to the development of novel antivirals.

What we already know about the -omics of HCMV infection

Large-scale "omics" techniques allow for rapid characterization of an organism's biological profile. Genomics, or the analysis of whole genomes, has its origins in phage DNA sequencing in the 1970s[22]; since its introduction, many organisms have had their entire genomes sequenced. Metabolomics also has its origins in the 1970s, when gas chromatography mass spectrometry was first applied to study the small molecules present in urine and human tissues[23]. Lipidomics has only recently emerged as a potential new avenue for biological research due to advances in mass spectrometry instrumentation[24]. Taken together, these techniques allow the production of large data sets documenting system-wide biological profiles. While high throughput approaches lack the specificity of targeted experiments, data generated through these high throughput approaches can often reveal macroscopic trends and patterns of thousands of different biological molecules under different conditions.

Metabolomic analysis has been used to probe how viruses alter the metabolism of their host cell. Previous work by Munger et al. documented metabolic changes in human fibroblasts infected by HCMV[21, 25]. Many metabolite pool sizes increased; these metabolites with increased pool sizes were found across many different biochemical pathways, including glycolysis, the tricarboxylic acid cycle, and pyrimidine synthesis[25]. Among all of the metabolites monitored, acetyl-CoA showed the largest increase, with a 35-fold greater pool size in infected cells[25].

Additionally, microarrays were used to monitor changes in mRNA transcript levels under infection[25]. The changes in mRNA levels for key metabolic enzymes often matched with the changes in pool sizes of the reactants and products of the reactions they catalyze[25]. For example, in addition to the increased level of acetyl-CoA observed in infected cells, HCMV-infected cells also demonstrated an increase in the mRNA transcript level of the various components of the pyruvate dehydrogenase complex, the enzyme responsible for the production of acetyl-CoA from pyruvate[25]. With more pyruvate dehydrogenase to convert pyruvate to acetyl-CoA, an increased pool of acetyl-CoA is expected. Additionally, phosphoenolpyruvate (PEP) pool sizes increased in HCMV-infected fibroblasts[25]. This increase was consistent with decreased levels of the mRNA transcript of pyruvate kinase, the enzyme that catalyzes the conversion of PEP to pyruvate[25]. With less pyruvate kinase to convert PEP to pyruvate, it follows that PEP will accumulate.

However, while changes in metabolite pool sizes can suggest changes in metabolic flux, pool sizes can only tell part of the story. An increase in the size of a given metabolite pool can be the result of either an increase in flux through a reaction producing the metabolite or a decrease in flux through a reaction consuming the metabolite. Therefore, seeing an increase in acetyl-CoA pool size does not necessitate an increase in flux to acetyl-CoA through pyruvate dehydrogenase.

This increase could be accounted for by a decrease in flux through reactions that consume acetyl-CoA, such as fatty acid synthesis. To address this issue, initial experiments were followed by kinetic flux profiling to probe the effect of HCMV infection on cellular metabolic flux[21].

A profile of the effect of HCMV infection on metabolic flux was developed using several distinct assays: directly measuring metabolite uptake and excretion, kinetic flux profiling with ^{13}C-labeled glucose and glutamine, using [1,2 ^{13}C] glucose to monitor the pentose phosphate pathway, and using [3-^{13}C] glucose to probe pyruvate metabolism[21]. Differences in HCMV-infected and mock-infected cells were seen using all of these assays; HCMV-infected cells showed both increased rates of glucose and glutamine uptake as well as increased rates of glutamate and lactate excretion[21].

Also, by switching HCMV-infected fibroblasts into media containing ^{13}C-labeled glucose and glutamine, the flow of ^{13}C through the metabolic network of HCMV-infected fibroblasts was shown[21]. The ^{13}C label appeared in metabolites "downstream" in the glycolytic pathway earlier in infected cells than in mock infected cells, indicating that glycolytic flux is increased under infection[21].

Flux from acetyl-CoA to citrate increased more than twenty-fold in infected fibroblasts[21]. However, the near complete ^{13}C-labeling of citrate fifteen minutes after the addition of ^{13}C-labeled glucose was accompanied by only thirty percent ^{13}C-labeling of malate, a compound downstream in the TCA cycle[21]. When ^{13}C-labeled glutamine was used instead of ^{13}C-labeled glucose, the TCA cycle compounds citrate and malate were labeled at similar levels[21]. This result suggests that the citrate produced from ^{13}C-labeled glucose via acetyl-CoA exits the TCA cycle to be used for another purpose. This increased pool of acetyl-CoA was accompanied by an increase in flux to fatty acid synthesis[21]. Since fatty acids are critical for the production of the

lipids that make up the outer envelope of the HCMV virion, it seemed likely that increased fatty acid synthesis was necessary for increased virion production. To test this hypothesis, inhibitors of fatty acid synthase and acetyl-CoA carboxylase, (fatty acid synthase synthesizes a 16-carbon fatty acid chain and acetyl-CoA carboxylase produces malonyl-CoA, a starting material for fatty acid synthesis) were added to HCMV-infected fibroblasts[21]. Cells treated with these inhibitors showed greatly decreased viral titers, indicating that fatty acid synthesis is important for virus replication and virion production[21].

Chambers et al. followed up on these experiments and explored the altered TCA cycle metabolism suggested by the results of the ^{13}C-labeling experiments, specifically the increased uptake of glutamine to provide carbon for the TCA cycle downstream of citrate[26]. Consistent with the flux analysis, cells grown in glutamine-free media failed to produce infectious virions[26]. Furthermore, virus titers in infected cells grown in glutamine-free conditions could be restored to pre-glutamine starvation levels through the addition of other TCA cycle intermediates, such as alpha-ketoglutarate and oxaloacetate[26]. These data further demonstrate that glutamine is entering the TCA cycle anapleurotically and is responsible for energy production through NADH generation.

Consistent with this hypothesis, increases in the activities of glutaminase and glutamate dehydrogenase were observed[26]. Glutaminase catalyzes the transformation of glutamine to glutamate, and glutamate dehydrogenase converts glutamate to alpha-ketoglutarate[26]. These two enzymes allow glutamine to enter the TCA cycle anapleurotically and be oxidized to produce NADH, which is consumed to produce ATP through oxidative phosphorylation. This suggests that fibroblasts use glucose primarily as a source of biosynthetic carbon, while they derive energy when glutamine and other compounds enter the TCA cycle anapleurotically[26].

Flux through pyruvate dehydrogenase was 84-fold greater in virus-infected cells than in mock-infected cells[21]. Pyruvate dehydrogenase catalyzes the transformation of pyruvate and coenzyme A into acetyl-CoA, a particularly relevant biological molecule. The pyruvate dehydrogenase complex is a multisubunit complex embedded within the inner membrane of mitochondria. The complex is made of multiple copies of three distinct subunits. The first subunit, pyruvate dehydrogenase (E1), decarboxylates pyruvate and frees a molecule of CO_2[27]. The second subunit, dihydrolipoyl transacetylase (E2), reduces lipoate and produces a molecule of acetyl-CoA from the previously decarboxylated pyruvate[27]. The third subunit, dihydrolipoyl dehydrogenase (E3), oxidizes the newly created dihydrolipoate back to lipoate, readying the E2 subunit to accept another pair of electrons and producing a molecule of NADH[27].

The enzyme is inhibited by both of its products: NADH and acetyl-CoA. NADH competes with NAD+ in the active site of E3 and acetyl-CoA competes with CoA at the active site of the E2 subunit[28]. The activity of the complex is also tightly regulated by phosphorylation; in cells, the complex is bound to both a kinase and a phosphatase[28]. The bound pyruvate dehydrogenase kinase can inactivate the enzyme by phosphorylating at any of three phosphorylation sites on the E1 subunit (Serine 232, Serine 292, and Serine 300)[28]. Alternatively, the bound pyruvate dehydrogenase phosphatase removes these phosphate groups, activating the enzyme[28]. In addition, NADH and acetyl-CoA activate the kinase, deactivating the complex when the concentrations of its products are high enough. As an additional level of regulation, calcium ions activate one of the phosphatases, which activates the enzyme complex[29]. Work by Sharon-Friling et al. determined that HCMV infection results in the release of calcium ions from the endoplasmic reticulum. These calcium ions were responsible for the actin reorganization of the cell and the characteristic cell swelling that gives human cytomegalovirus

its name[30]. However, it is also conceivable that these calcium ions could also enter the mitochondria and activate pyruvate dehydrogenase phosphatase. This increased phosphatase activity would increase the fraction of the E1 subunit present in its nonphosphorylated, active state, increasing flux from pyruvate to acetyl-CoA.

Fatty acid metabolism was shown to be essential for cytomegalovirus replication; fibroblasts treated with fatty acid synthesis inhibitors yielded reduced virus titers after infection, indicating that synthesis of fatty acids is an important step in viral replication[21]. Lipidomic profiling has revealed that very long chain fatty acids are present in particularly high numbers in infected cells, likely as a part of the viral envelope.

Given the importance of lipids for virus replication, recent experiments have examined the role of acetyl-CoA carboxylase in HCMV infection and determined that inhibition of the enzyme leads to decreased virus titers[31]. Acetyl-CoA carboxylase, which converts acetyl-CoA into malonyl-CoA as the first committed step of fatty acid synthesis, increased in abundance as a result of HCMV infection[31]. Additionally, the enzyme was observed to be more active in infected cells, consistent with the increased flux seen to fatty acid synthesis observed previously[21, 31].

Many aspects of the effect of HCMV infection on host cell metabolism in human fibroblasts have been demonstrated using a combination of transcriptomic, metabolomic, and lipidomic methods. However despite all of this information about various classes of biomolecules, the lack of proteomic analysis is glaring. Advances in protein mass spectrometry have made rapid quantification of the proteome possible; these techniques have already been applied to study several other types of viruses, documenting the effects of these parasites on host cell proteomes.

Protein mass spectrometry

Two of the main goals of protein mass spectrometry are the identification and quantification of peptides and proteins. Two major strategies for protein mass spectrometry have emerged: bottom-up analysis, where proteins are digested to peptides by a protease prior to analysis by mass spectrometry, and top-down analysis, where whole proteins are analyzed by the spectrometer[32]. In bottom-up protein mass spectrometry, peptides are fragmented and the resultant fragment ions are analyzed and quantified[32]. These fragment ions are reconstructed into a peptide sequence, which is then matched to a database to identify the protein from which the peptide arose. In top down proteomics, entire proteins are ionized out of solution and fragmented into peptides[32]. These peptides are detected by the mass spectrometer and reconstructed into a protein[32].

A typical mass spectrometer used for proteomic analysis contains three major instruments: an ion source, a mass analyzer, and a detector[33]. Most detectors are functionally identical; they usually have high sensitivity as they use electron multiplier tubes to enhance signals[33]. However, there are many possible combinations of ion sources and mass analyzers, each of which can be used to produce a mass spectrometer tuned for slightly different methods of analysis.

Ion sources have traditionally ionized peptides in one of two main ways, either through electrospray ionization (ESI) or through matrix-assisted laser desorption / ionization (MALDI)[33]. In ESI, a high voltage is applied to a stream of liquid, loading the molecules in solution with charge[33]. These charged species repel each other, forming an aerosol[33]. The aerosol is then filtered into a stream of charged, fast-moving ions[33]. MALDI uses a different approach to ionization; a crystalline matrix is loaded with the sample to be analyzed[33]. A laser is then shot at

the matrix, exciting the matrix molecules and causing the sample molecules to gain or lose a proton and leave the matrix[33]. However, ESI has emerged as the more readily usable technique; it ionizes peptides out of solution, so it can be coupled to a liquid chromatography (LC) column for an easy transition from separation to ionization[33]. On the other hand, ionization through MALDI requires the preparation of a crystalline structure, which does not allow for easy analysis post-LC separation[33].

After peptides have been ionized, they enter the mass analyzer. Central to the mass spectrometer, mass analyzers come in many different forms. Uniting all mass analyzers is their use of electromagnetic fields to deflect charged species and discriminate ions based on their mass-to-charge ratio. Several different types of mass analyzers may be used for proteomic analysis, including the quadrupole, time of flight, Fourier transform ion cyclotron resonance, and the Orbitrap[33].

The quadrupole consist of four metal rods arranged around a central chamber. During analysis, a voltage is applied to the poles, causing ions in the quadrupole to oscillate with a frequency dependent on their mass-to-charge ratio[33]. The current is adjusted to filter out all ions that do not oscillate within a certain range; these ions collide with the metal rods and are destroyed[33]. As a result, the only ions that pass through the quadrupole are those with a set mass-to-charge ratio, allowing the quadrupole to function as a mass filter[33]. The triple quadrupole is a particularly common equipment set up for protein mass spectrometry. In this style of analysis, three quadrupole analyzers are arranged linearly, the first acts as a mass filter and isolates a peptide of interest, flinging out the rest[33]. The second quadrupole is filled with an inert gas and serves as a collision cell, where peptides collide with the gas and undergo collision-induced dissociation, fragmenting along their peptide backbone[33]. These resultant fragment peptides are

then fed into the third quadrupole, which scans the entire m/z dynamic range and detects fragment ions of all sizes[33]. The fragment ion spectra are then analyzed and matched to peptides, which are mapped to proteins based off of amino acid sequences in protein databases[33].

The time-of-flight mass analyzer (TOF) is another type of mass analyzer. In TOF, ionized peptides are accelerated using a high voltage gap[33]. The ions fly towards a detector; ions with higher mass-to-charge ratios travel at slower speeds and take a longer amount of time to reach a detector[33].

Fourier transform ion cyclotron resonance (FT-ICR) is yet another example of a mass analyzer. Ions are trapped by electromagnetic forces and rotate around in a cyclotron[33]. These rotating ions will oscillate with a frequency dependent on their mass-to-charge ratio and produce a measurable current in the charged metal plates surrounding the chamber[33]. The frequency of the current can be broken down into component frequencies using a Fourier transform[33]. These component frequencies indicate the mass-to-charge ratios of the various trapped ions with a very high degree of mass accuracy.

The Orbitrap, a recently developed mass analyzer, relies on similar principles[34]. Ions are trapped in a cylindrical field surrounding a metallic rod[34]. The rotating ions also produce a current in the rod, which can be broken down into component frequencies to determining the mass-to-charge ratios of the various rotating ions[34]. The Orbitrap provides a very high degree of mass accuracy, a large dynamic range, and excellent resolving power[34]. These attributes have made the Orbitrap, and its recent upgrade, the Velos, two of the most powerful mass analyzers in protein mass spectrometry[34, 35].

The mass spectrometers used for mass spectrometry-based proteomics vary greatly. This variation allows researchers to customize experiments to best suit the mass spectrometers

available; different combinations of ionizers and mass analyzers are used for different purposes, depending on the goals of any particular analysis.

Mass spectrometry-based proteome studies of viruses

Even though the discipline is relatively new, proteomics analysis through the use of mass spectrometry has already been applied to virology. Many studies have used protein mass spectrometry to identify proteins present in released virus particles and to elucidate mechanisms of various parts of virus lifecycle including virion assembly. Some studies detailing changes in host cell proteome under infection by different infectious agents have been completed. The goals of these studies differ slightly; some look at changes in metabolic proteins under infection while others sought to identify host cell factors involved in viral replication.

One study of herpes simplex virus type-1 (HSV-1) used 2-D gel electrophoresis to track changes in host cell proteomes resulting from HSV-1 infection[36]. Use of the 2-D gel electrophoresis technique makes the quantification of large numbers of proteins difficult, in this case roughly 100 proteins were quantified (63 upregulated and 40 downregulated)[36]. Several proteins previously observed as important for HSV-1 replication were seen to be upregulated as a result of infection[36]. However, the majority of proteins identified were involved in either DNA and RNA processes or the ER stress response; very few metabolic proteins were identified in the screen[36].

Temporal proteomic analysis of liver hepatocytes under infection with Hepatitis C virus revealed the virus altered the abundances of the enzymes of the host cell's metabolic pathways to support viral replication[37]. Enzymes involved in fatty acid metabolism, such as citrate synthase and fatty acid binding protein 1, showed marked upregulation within 24 hours of infection[37].

Complementing this proteome data, lipidomic analysis revealed increased abundance of fatty acids during infection[37]. Additionally, most glycolytic enzymes showed increases of between 1.5- and 3-fold 24 hours post infection[37]. However, these same glycolytic enzymes also showed a relative decrease in abundance 48 hours post infection, indicating that the viral moderation of host cell energetic pathways is more complex than a simple up or down regulation of enzymes[37]. These changes in enzyme levels with time suggest distinct phases of virus infection, a biosynthetic phase followed by an energy producing phase that maintains energy levels during virus replication[37].

A similar proteomic analysis of cells infected with coronavirus, a member of the family of viruses that includes SARS virus, revealed large changes in the global proteome as a result of infection[38]. Instead of looking at metabolic proteins, the authors used Gene Ontology (GO) annotations to identify organelles containing high percentages of up or down regulated proteins[38]. In particular, all of the proteins with changes in abundance in the Golgi apparatus and endoplasmic reticulum showed downregulation; observing changes in protein levels in these organelles is consistent with the massive membrane disruption observed in virus-infected cells, suggesting the virus has a particularly large effect on these two membrane-enclosed organelles[38]. Additionally, none of the detected ribosomal proteins were depleted, but many were upregulated, consistent with the idea that virus infection increases protein synthesis[38].

Finally, though not technically a virus, Nelson et al. completed a proteomic analysis detailing the effect of an intracellular parasite, the protist *Toxoplasma gondii*, on the human fibroblast proteome[39]. The plurality of proteins with observed changes in abundance were metabolic proteins[39]. Furthermore, most glycolytic proteins increased in abundance under infection by between 1.5 fold and 3 fold, falling in line with a trend that has been observed for

other parasites, including HCV[37, 39]. Additionally, one third of the proteins affected by the parasite localized to the mitochondria, suggesting that the parasite's metabolic reorganization of the host cell is focused on energy metabolism[39].

These studies have elucidated different aspects of the impact of virus infection on cellular metabolism. As a general trend, upon entering the cell, viruses and other parasites modulate host cell metabolism to maximize the production of biomolecules and produce enough energy to enable replication. Additionally, changes in protein level upon infection are generally 2 or 3 fold increases; these observed changes in enzyme levels are relatively small compared to changes in changes in metabolic flux observed as a result of infection by the same parasites.

Several studies have compared proteomic data to metabolic flux, in an attempt to understand the relationship between abundances of various enzymes and flux through the reactions they catalyze. Costenoble et al. studied changes in protein levels in response to different environmental conditions (ex. aerobic vs. anaerobic, glucose-containing media vs. galactose-containing media) in *Saccharomyces cerevisiae*[40]. A targeted proteomic approach termed selected reaction monitoring was used to analyze proteins previously identified as metabolic enzymes[40]. They had previously observed changes in metabolic flux under different conditions, but found that these changes in flux did not correlate with changes in the amount of protein present, and a change in protein level was not required to change flux through a process[40]. However, it was observed that when a reaction went from having no observed flux in a certain condition to a non-zero flux in another, this change was often accompanied by an increase in protein level, suggesting that the addition of new fluxes is controlled by synthesizing protein while flux modifications do not require protein synthesis[40].

Additionally, previous studies have found little in terms of correlation between levels of mRNA transcript and corresponding protein levels in both yeast and human cells[41-43]. mRNA transcripts have different stabilities; some degrade quickly and produce very little protein, while others are very stable and produce large quantities of protein. The protein products of these mRNAs also have varying half-lives and as a result persist in the cell for varying amounts of time. Thus, while mRNA transcripts can provide clues as to a protein's abundance, they are not a direct indicator of abundance. In order to truly determine the effect of HCMV on the abundances of host cell proteins, these proteins must be examined directly.

Proteomics and HCMV

Despite the wealth of information about HCMV's effect on fibroblast metabolism, a more complete understanding of the effect of cytomegalovirus infection on mammalian cells requires more modalities of data. As discussed previously, genomic and transcriptomic data detailing the virus's effect on nucleic acids synthesis, specifically mRNA, suggest possible alterations to protein concentrations. However, only looking at one class of biomolecules will provide an inherently incomplete picture, as protein levels and mRNA transcript level often do not show strong correlation. Also, genomic techniques cannot capture translational effects or post-translational modifications. After translation, proteins can be modified in many different ways that alter their biological activity. The activities of many proteins are controlled by their acetylation or phosphorylation state, and some even need to be cleaved to become active.

Metabolic enzymes are no strangers to these types of post-translational modification. Most enzymes catalyzing functionally irreversible reactions have several modalities of control; they may have multiple phosphorylation sites or allosteric inhibitors. For example, the E1

subunit of the pyruvate dehydrogenase complex is inactivated by phosphorylation[28]. Therefore, an increase in the level of the mRNA transcript encoding this enzyme cannot be used to conclusively show an increase in the activity of the enzyme, as even if there is more mRNA present under a certain condition, a higher fraction of the enzyme may be phosphorylated and inactive, resulting in less active enzyme despite more mRNA. Other proteins may be produced in an inactive form and must be modified in some way (through acetylation or cleavage) in order to become active. These simple examples of post-translational modifications are only a few of many ways cells can alter the functions of proteins that cannot be observed using an mRNA transcript-based approach. The many possible fates of the protein arising from a piece of mRNA illustrate the need for a protein-level approach in order to have a more complete understanding of complex cellular processes.

One of the aims of systems biology is to integrate different modalities of data, to combine the results of transcriptomic and metabolomic analysis with proteomic data to achieve a more complete understanding of the physical basis of life. Application of these three approaches used in tandem to study the virus-host interaction allows for an integrated analysis of the mechanisms of infection. However, this analysis is inherently incomplete without data at the protein level, especially since proteins are responsible for carrying out many biological processes such as cellular metabolism.

Relative concentrations of proteins in cells can be achieved using stable isotope labeling by amino acids in cell culture (SILAC) [44]. SILAC is a relatively new method for determining the relative quantifications of proteins in cells grown in cell culture via mass spectrometry[44]. In this technique, cells are grown for several generations in media containing key amino acids with atoms of either heavy or light isotopes (if trypsin is to be used for proteolysis, arginine and lysine

containing heavy isotopes are used) until a high percentage of the cellular protein contains only these heavy or light amino acids[44]. The cells are then subjected to different experimental conditions. Cells are lysed and the protein concentrations of the lysates are determined using the SDS-compatible BCA protein content assay[44]. Equal amounts of heavy and light protein are added to a single sample tube, allowing any error in any further sample processing steps to propagate equally in the samples[44].

Proteins are then digested using a highly purified protease[44]. If arginine and lysine are used as the heavy labeled amino acids, trypsin is used to digest proteins, as it will cleave after each arginine or lysine, resulting in peptides containing one and only one arginine or lysine[44]. Since samples from different conditions contain arginine and lysine made with different isotopes of carbon, this cleavage allows peptides from different samples to be distinguished based on mass shifts (except for C-terminal peptides, which likely contain no arginine or lysine)[44]. These peptides are then quantified based upon the relative abundance of each peptide identified as part of a heavy-light isotopically labeled pair[44, 45].

Since peptides ionize with different efficiencies depending on their amino acid sequences, mass-spectrometry based techniques cannot quantify a protein's absolute abundance without the use of an external standard. However, a SILAC protocol circumvents this inability to establish absolute quantification by determining the relative peptide signals of the two samples. This relative quantification is possible because the presence of heavy isotopes does not affect a peptide's ionization efficiency; peptides with the same amino acid sequence in the heavy and light samples ionize equally well[44]. This equal chance of ionization allows the samples to be directly compared, using ion counts to establish a ratio of protein in the light and heavy conditions[44]. Each peptide is then matched back to a protein database to identify the protein

whose digestion would produce the peptide identified[44, 45]. Many scans are completed, making quantification of any individual protein the result of the quantification of many distinct peptides.

This study

In this study, we complete the first quantitative proteomic analysis to identify changes in the human fibroblast proteome resulting from HCMV infection. To do this, we use a human fibroblast model of infection. The virus can infect many different cell types including, but not limited to, endothelial cells, epithelial cells, and fibroblasts[1]. Fibroblasts represent a convenient medium in which to study HCMV infection; they are easy to culture and the virus infects these cells without difficulty. In fibroblasts, the virus begins producing new viral structural proteins within 48 hours of infection[1].

A SILAC labeling experiment was completed in an effort to obtain a general sense of the effect cytomegalovirus has on the host cell proteome. This proteomic analysis was compared to microarray data for infected cells to determine the relationship between mRNA transcript levels and protein levels. Also, the changes in levels of key metabolic enzymes were compared to changes in flux through the reactions catalyzed by those enzymes to determine the method by which HCMV alters flux through these enzymes. This coupling of different modalities of analysis allows us to view proteomic analysis as a link between changes in mRNA transcript levels and changes in metabolic flux, and probe the biological basis of these changes.

Additionally, Western blots were performed for two purposes, both to verify the quantification determined by the SILAC mass spectrometry-based analysis, and to examine in greater detail specific metabolic enzymes with large discrepancies between their observed change in protein level and the change in flux of the reactions they catalyze.

Mass Spectrometry Sample Preparation and Data Acquisition

Adapted from a Materials and Methods section by D. Perlman

Biological reagents and stable isotope labeling

Human foreskin fibroblast cells (HFFs) (passage 8) were grown to confluence in 10 cm plates using Dulbecco's modified Eagle medium (DMEM) containing 10% dialyzed fetal bovine serum (FBS), penicillin/streptomycin and maintained at 37 °C in the presence of 5% CO_2. HFFs were grown for >7 generations in unlabeled (light) (DMEM) or labeled (heavy) DMEM reconstituted with U-$^{13}C_6$ lysine (Δ6) and U-$^{13}C_6$, U-$^{15}N_4$ arginine (Δ10). A viral stock of wild-type HCMV strain AD169 purified by centrifugation through 20% sorbitol was used to infect HFFs.

Infections

Confluent plates of cells were serum starved for 24 hours prior to infection and were infected with HCMV AD169 at a multiplicity of infection of 3.0 pfu/cell (light cells) or mock infected (heavy cells). After 1 hour, the infection media was replaced with serum-free DMEM. At a time of 48 hours post infection, cells were lysed *in situ.*

Lysis

Cells were lysed using a solution of 4% SDS, 100 mM Tris pH 7.4, 5 mM DTT, and protease and phosphatase inhibitors. The samples were boiled for approximately 10 min with frequent vortexing to help reduce viscosity by shearing DNA prior to centrifugation to clarify the lysate. Protein quantification was performed on the lysate using the reducing-agent compatible

BCA assay (BCA RAC, Pierce). After quantification, equal amounts of protein from each condition were pooled. Samples were stored at -80 °C until processing.

Sample processing

Pooled protein solution (containing 1 mg protein) was subjected to buffer exchange, thiol reduction and alkylation, and trypsin digestion by the FASP procedure[46]. Peptides were desalted by capillary reversed-phase chromatography (500um id x 20 cm, POROS 10R2 C18 resin) using a Harvard syringe pump, and eluted directly onto a home-packed capillary strong cation exchange column (500um id x 45 cm, POROS SCX resin), which was connected to the outlet of the reversed-phase column to minimize sample loss.

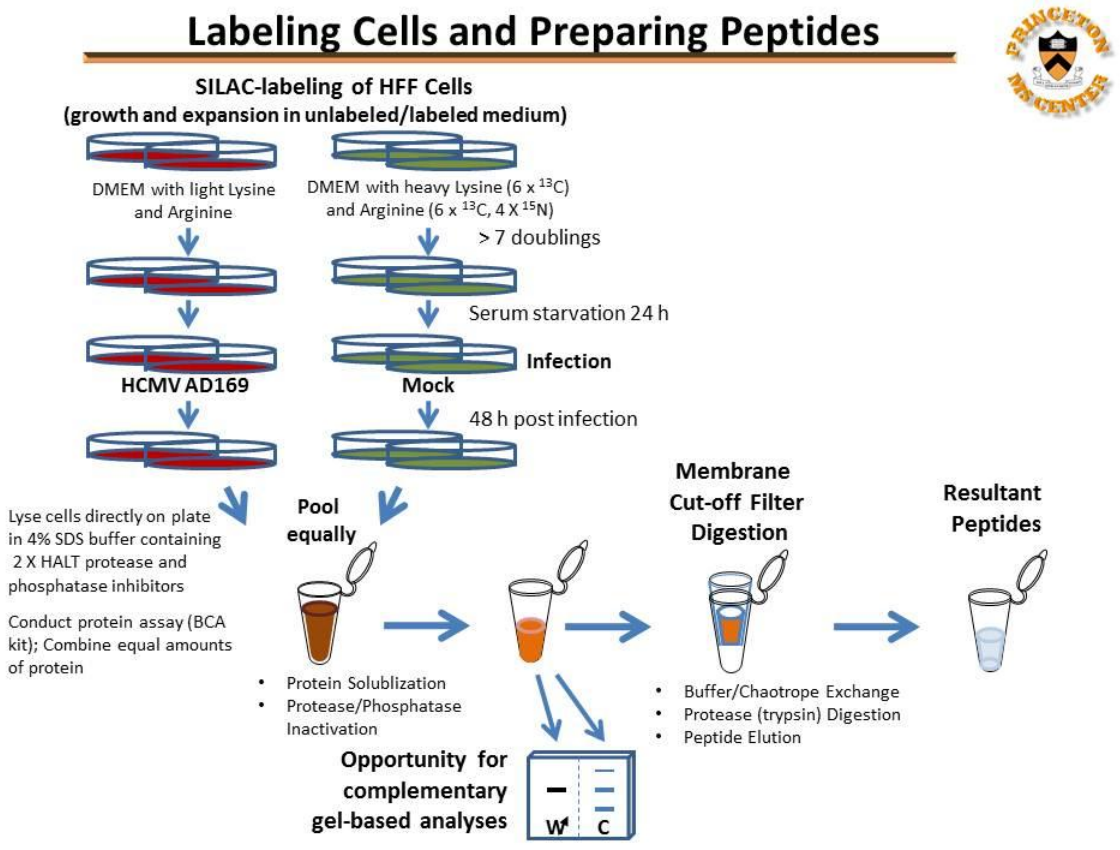


Figure 1 – Labeling Cells and Preparing Peptides (D. Perlman)

SCX Fractionation

SCX fractionation of peptides was conducted using a Dionex Ultimate NanoLC capillary HPLC system (Dionex, Sunnyvale, CA), using a gradient from a 75%:25% mix of buffers A:B to 100% buffer B (buffer A: 7mM KH2PO4, pH 2.65 , 30% ACN; buffer B: 7mM KH2PO4, 350mM KCl, pH 2.65 , 30% ACN) over an 80 min period at flow rate of 10ul/min, followed by column stripping and reconditioning for 10 min in buffer C (50mM K2HPO4, 500mM NaCl, pH 7.5) and water. SCX fractions were collected every 5 min and were pooled into 14 fractions of roughly equivalent peptide abundance according to the integration of their UV absorbance (λ = 214 nm). An additional 15th fraction was added consisting of the SCX column flow through.

Chromatography

Fractions were desalted using StageTip micro-scale reversed-phase chromatography[47]. The desalted fractions were then subjected to reversed-phase nano-LC-MS and MS/MS performed on a nano-flow capillary high pressure HPLC system (Eksigent, Dublin, CA) coupled to an LTQ-Orbitrap™ hybrid mass spectrometer (ThermoFisher Scientific, San Jose, CA), outfitted with an NanoMate ion source robot (Advion, Ithaca, NY). Sample concentration and desalting were performed online using a trapping capillary column (200 μm x ca. 30 mm, packed with 5 μm, 100 Å Magic AQ C18 material, Michrom, Auburn, CA) at a flow rate of 7μL/min for 3.5 min, while separation was achieved using an analytical capillary column (75 μm x ca. 20 cm, packed with 3 μm, 100 Å Magic AQ C18 material, Michrom), under a linear gradient of A and B buffers (buffer A: 3% ACN/ 0.1% FA; buffer B: 97% ACN/ 0.1% FA) over 180 min at a flow rate of approximately 0.5 μL/ min.

Ionization and MS/MS

Electrospray ionization was carried out using the NanoMate at 1.74 kV, with the LTQ heated capillary set to 200 °C. Full-scan mass spectra were acquired in the Orbitrap in the positive-ion mode over the range m/z 300–1800 at a resolution of 60,000. MS/MS spectra were simultaneously acquired using the LTQ for the seven most abundant multiply charged species in the full-scan spectrum having signal intensities of >1000 NL. Dynamic exclusion was set such that MS/MS was acquired only once for each species over a period of 120 s. All spectra were recorded in profile mode. Chromatography and LC-MS procedures are outlined in Figure 2.

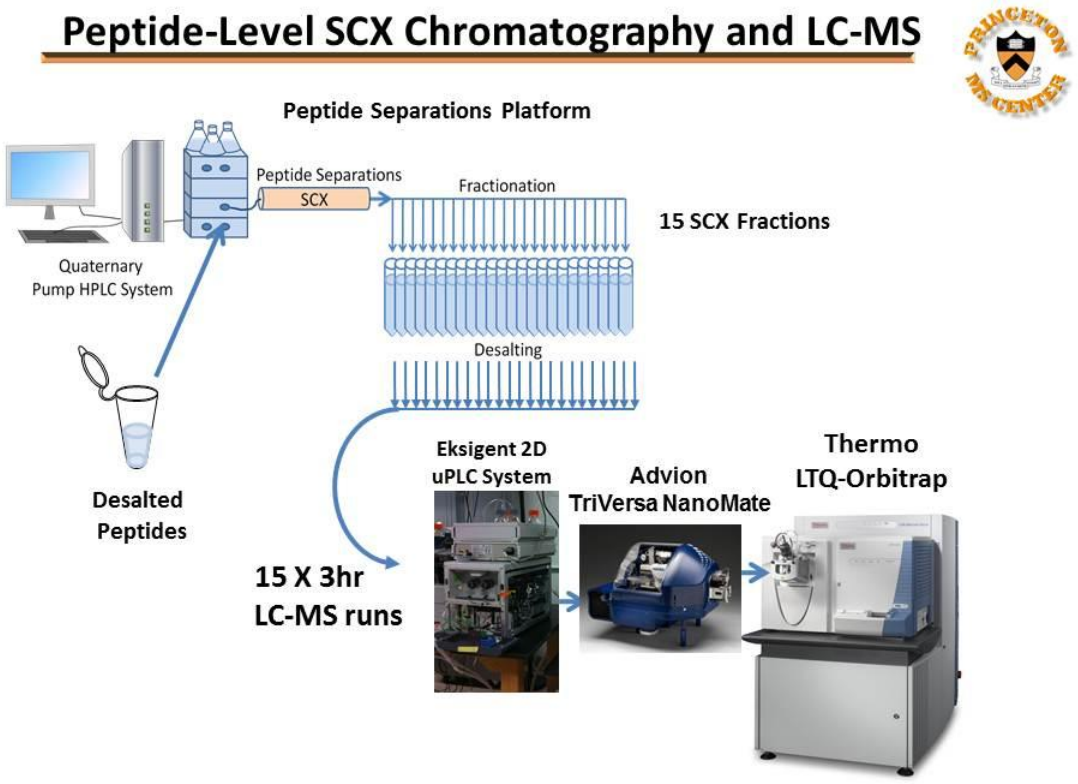


Figure 2 – Peptide-Level SCX Chromatography and LC-MS (D. Perlman)

Data Analysis

Data analysis, including quantitation of the SILAC pairs in the MS and identification of these pairs by their MS/MS, was performed using MaxQuant[48] with the embedded Andromeda search engine[49], using the following critical parameters: the search was conducted against the SwissProt human database concatenated with the HCMV AD169 sequences from Trembl; trypsin with up to two missed cleavages was used with an initial error tolerance of 15 ppm for the first search before internal recalibration; oxidation of methionine, *N*-terminal protein acetylation, pyroglutamate from *C*-terminal glutamate or glutamine, and phosphorylation at S,T, and Y residues were permitted as variable modifications, while cysteine carbamidomethylation was required as fixed; both protein and peptide false-discovery rates were set to 1%, and proteins were quantified based on unique peptides only.

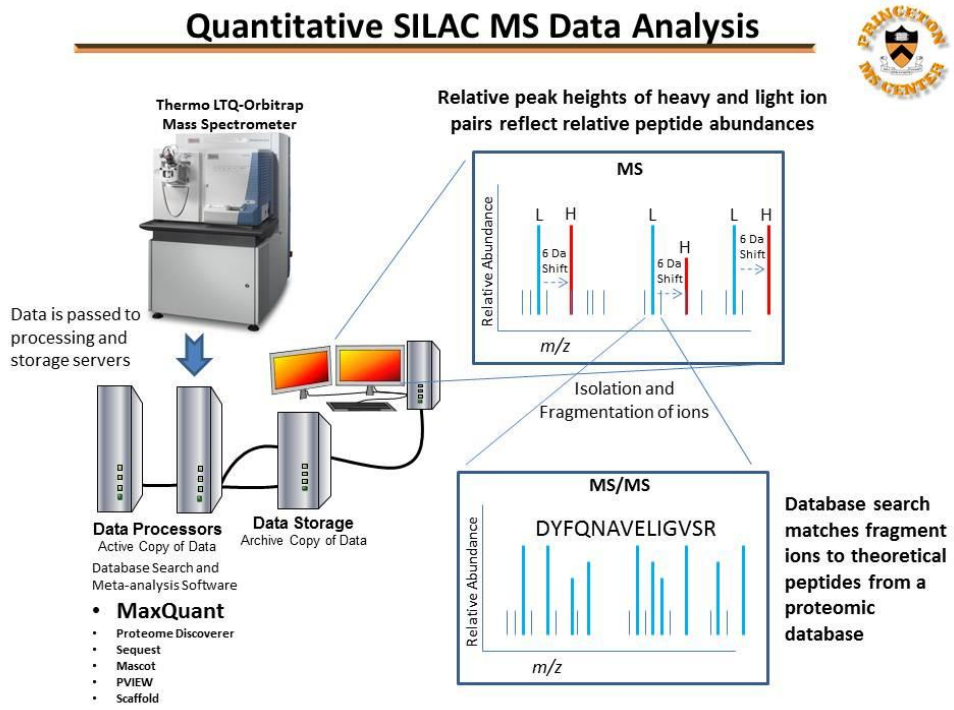


Figure 3 – Quantitative SILAC MS Data Analysis (D. Perlman)

Gene Ontology and Enrichment Analysis

Gene Ontology and Enrichment Analysis were completed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) [50]. Proteins were classified into three groups: those that were upregulated under infection (those with normalized Heavy/Light (H/L) ratios of 0.66 or less), those that were downregulated under infection (those with normalized H/L ratios of 1.5 or greater), and those that showed no change under infection (normalized H/L ratios of between 0.66 and 1.5). 880 proteins were in the upregulated group, and 598 were in the downregulated group. A 1.5-fold change in either direction normalized to a median H/L ratio of 1 was chosen as the cutoff point because it produced up and down regulated groups of roughly equal size, with a large group of proteins classified as not changing. Gene Ontology (GO) term enrichment was performed using these proteins lists, using DAVID's Functional Annotation Clustering function. For a more thorough explanation of the clustering refer to the Results section of this thesis.

SDS PAGE and Western Blots

Biological reagents

HFFs (passage 12) were grown to confluence in 10 cm plates using DMEM containing 10% FBS. Cells were infected with the wild-type HCMV strain AD169.

Infections

Confluent plates of cells were serum starved for 24 hours and either mock-infected or infected with HCMV AD169 at a multiplicity of infection of 3.0 pfu/cell. After 1 hour, the

infection media was replaced with serum-free DMEM. Cells were harvested for protein analysis at 4, 24, 48, 72, and 96 hours post-infection.

Protein harvesting and Western blotting

Cells were washed twice with warm serum-free DMEM, scraped, and pelleted by centrifugation. Media was aspirated and the cells were resuspended in 100μL of RIPA buffer containing 20 mM Tris pH 7.2, 0.1% SDS, 1% Tx-100, 1mM EDTA, 0.5% NaDOC, 150 mM NaCl, and Roche Protease Inhibitor. Cells were resuspended in a small volume of lysis buffer because one of the antibodies used for Western blotting was specific to a certain phosphorylation state of an enzyme, so a large amount of protein may have been needed to load onto a gel. Cells were sonicated until no pellets were visible. Samples were then boiled for 5 minutes, vortexed for 1 minute, boiled for 3 minutes, vortexed for 1 minute, boiled for 3 minutes, and vortexed for 1 minute (similar to the lysis conditions used for proteomic analysis).

Protein concentrations of the samples were quantified using a BioRad protein assay. A standard curve was developed using BSA varying in concentration from 0 mg/mL to 9 mg/mL and showed a linear relationship over the given range. All samples were found to have protein concentrations within this linear range. Equal amounts (30 μg) of protein of each sample were loaded onto a 10% SDS-PAGE gel and run for two hours at 25 milliamps per gel. Proteins were then transferred to nitrocellulose paper (ProTran) with a current of 140 milliamps per gel for 90 minutes. The nitrocellulose paper was blocked with 5% milk for 45 minutes. Primary antibodies were added to the nitrocellulose paper at dilutions ranging from 1:2500 to 1:5000 in TBS-T (20 mM Tris pH 7.6,135 mM NaCl, 0.5% Tween-20) and 0.5% milk. After rocking overnight at 4 °C, nitrocellulose paper was rinsed 3 times with TBS-T for 5-15 minutes. Secondary antibodies conjugated to horseradish peroxidase were added at a 1:5000 dilution also in TBS-T and 0.5%

milk. After rocking for 45 minutes at 4 °C, blots were rinsed 4 times with TBS-T for 5-15 minutes. Proteins were visualized using ECL reagent and exposed to film (Kodak) for 30 seconds.

Antibodies

Antibodies specific to the phosphorylated serine at position 300 of the pyruvate dehydrogenase E1α subunit were purchased from CalBioChem. Antibodies to α-tubulin were purchased from Sigma. Antibodies specific to pp28, an HCMV protein, were the gift of the Shenk lab. Secondary antibodies produced in goat and conjugated to horseradish peroxidase (HRP) were used. These antibodies were also the gift of the Shenk lab.

Results

Results of the SILAC screen

The SILAC-based proteomic screen identified many proteins and quantified their changes as a result of infection. Across the entire set of 15 SCX fractions each subjected to 3 hour LC-MS runs, 13,281,884 peaks were detected. 1,209,838 of these detected peaks had recognizable peptide-like isotope patterns; among those, 275,406 were recognized as members of a SILAC pair. While 226,876 MS/MS were achieved on the fly, 80,769 of them were a member of a SILAC pair. Of these SILAC pairs, 45,372 were identified by database searching with a 1% peptide false discovery rate (FDR). These peptides map back into 3,591 proteins (also 67 proteins identified as contaminants, and 36 reverse proteins) at a 1% protein FDR. 74 proteins (approximately 2% of all proteins identified) were HCMV proteins, indicating the protein products of a large fraction of the HCMV genome were detected. We were able to achieve some kind of quantification for these 3,591 proteins, giving an average of more than 12 peptides identified and quantified per protein. However, the median number of peptides quantified per protein was 5, indicating that despite an average of 12 peptides per protein, at least half of the quantified proteins had quantifications based off of 5 or fewer peptides. 951 proteins, or 26% of the 3,591 proteins, were quantified based off of 12 peptides or more. Additionally, peptides of a given sequence were often quantified multiple times; each distinct peptide underwent an average of 2.35 MS/MS events.

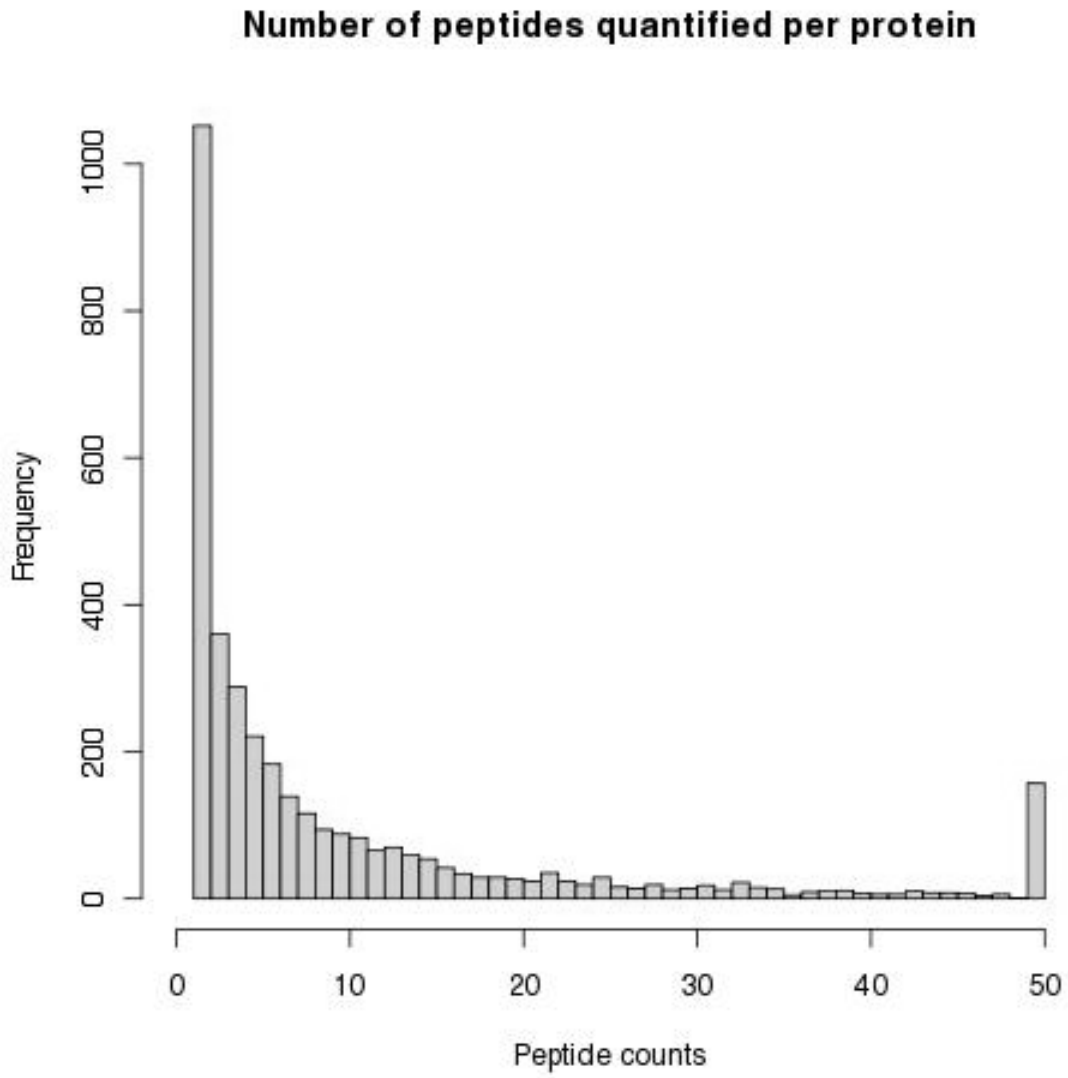


Figure 4 – Number of proteins quantified for each peptide count

Figure 4 depicts the number of proteins for which a certain number of peptides were quantified. The rightmost bar represents all proteins with quantifications based off of 50 or more peptides. As is evidenced by the height of the first bar, the quantification for the plurality of proteins was based off of 1 peptide. However, the median peptide count of 5 indicates that the quantifications of at least 50% of the proteins were based off of 5 or more peptides.

Unfortunately, only one replicate of the SILAC labeling experiment could be completed at this time. A second set of plates were grown and infected, reversing the SILAC labels by infecting fibroblasts labeled with heavy arginine and lysine and mock infecting fibroblasts grown in light, normal media. If this pair of samples could be analyzed, the data produced would provide both validation of the previously determined values for specific proteins and confirm that the choice of heavy or light arginine and lysine labels did not influence cellular growth, infection dynamics, or the quantification by mass spectrometry. However, due to various circumstances beyond our control (circuit-boards catching on fire), only one of the replicates was analyzed via mass spectrometry. As a result, any calculations of error for protein quantification are based on the previously mentioned average of 12 peptides quantified per protein.

The relative quantification for each protein is often based on the quantifications of multiple peptides. Therefore, even though only one biological replicate was completed, many measurements were taken to quantify the relative amount of heavy and light peptides present in the sample.

Quantification of sample peptides

In order to demonstrate the quantification process for a typical peptide, several total ion chromatograms are found below, each illustrating a different aspect of the quantification procedure.

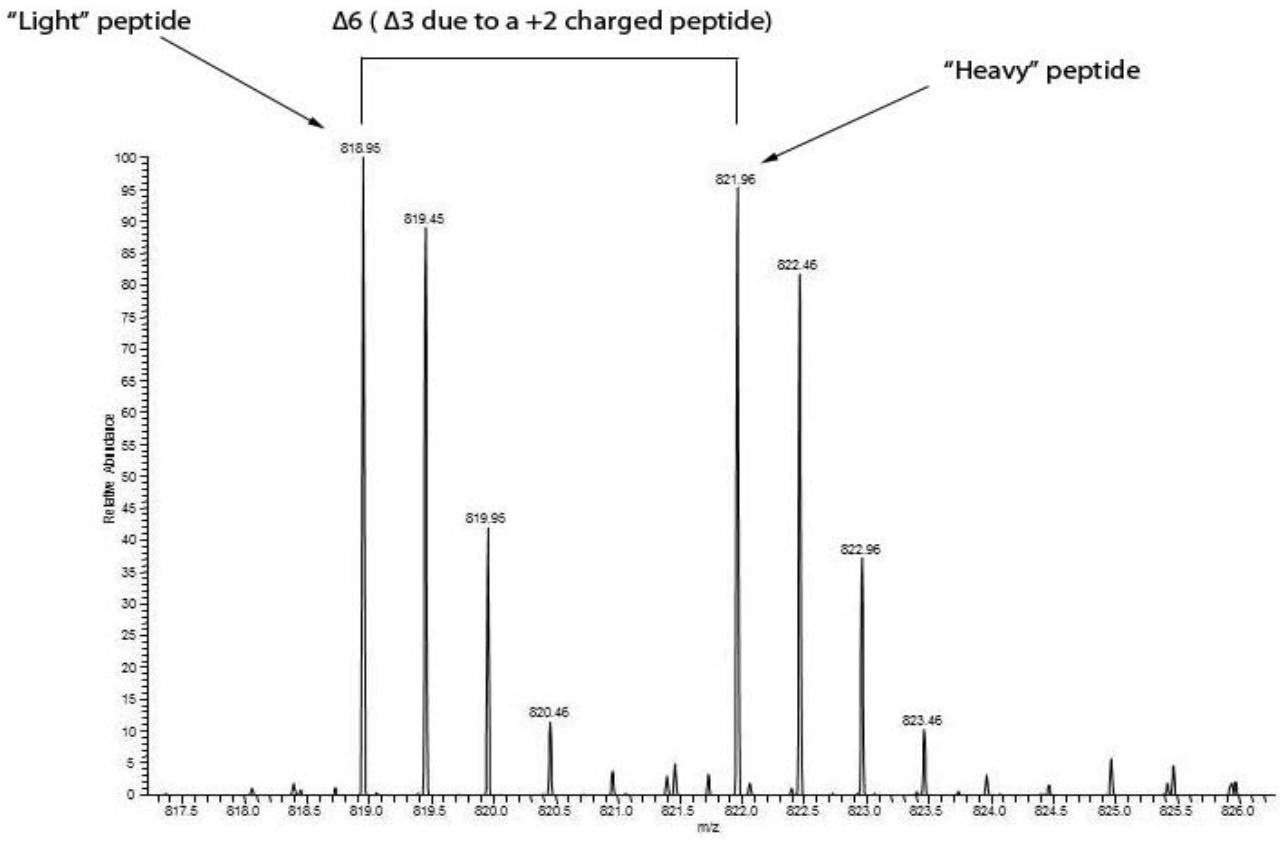


Figure 5 – Total ion chromatogram for a peptide from pyruvate kinase

Figure 5 shows the spectra used for quantification of one peptide of pyruvate kinase. The peptide sequence was determined through MS/MS analysis, where the peptide is fragmented and the resultant ions of various sizes are used to determine the amino acid sequence of the peptide (not shown). In Figure 5, the peaks corresponding to the heavy and light isotope-containing versions of the peptide are clearly visible. In this example, the peptide contains a heavy-labeled lysine so one would expect to see a shift of 6 Da. However, because the peptide is doubly charged (in the +2 state), the gap between the masses of the heavy and light peptides is only 3 Da. Ultimately, pyruvate kinase was determined to have a H/L ratio of 0.961, indicating it is only slightly more abundant in the infected cells. This can be seen in Figure 5, as the light peptide peak is just noticeably higher than the heavy peak.

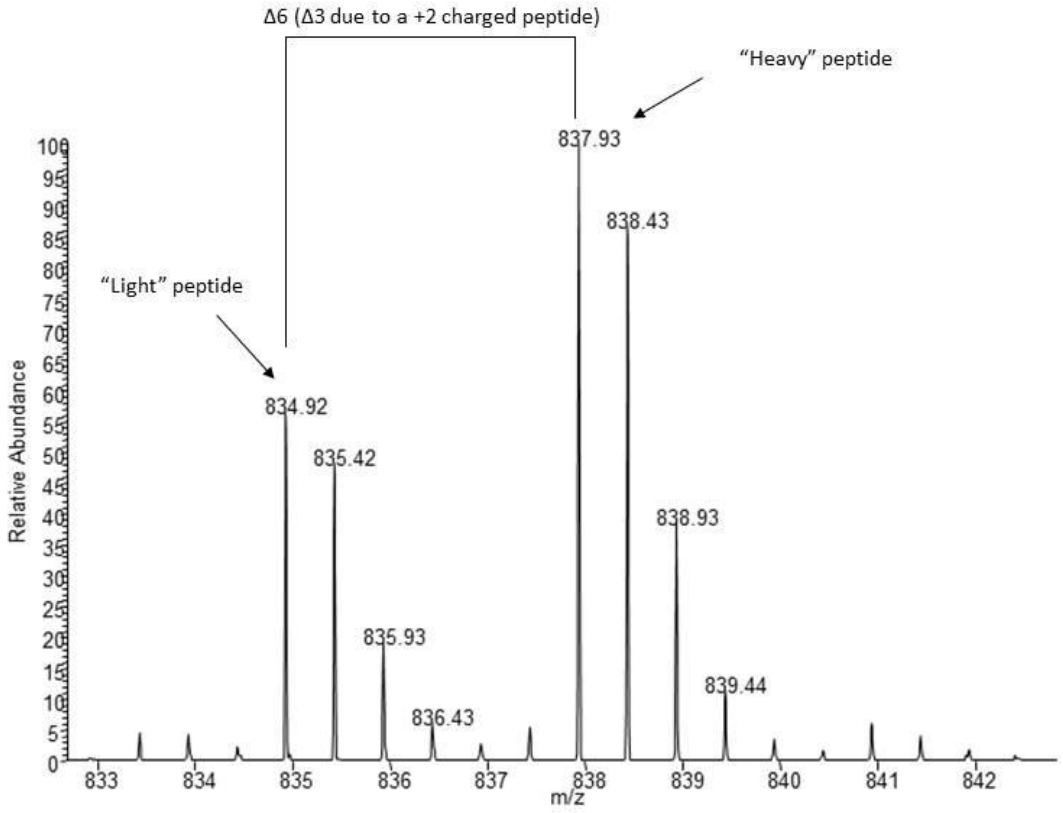


Figure 6 – Total ion chromatograms for a peptide used to quantify AHNAK

Quantification of a sample protein with a change in abundance as a result of infection is shown in Figure 6. This protein, the neuroblast differentiation-associated protein AHNAK was the protein for which the most peptides were quantified (667 quantification events for 276 distinct peptides). AHNAK, which translates to "giant" from Hebrew, is a massive, 5643 amino acid protein that interacts with calcium transport channels in the heart[51]. Additionally, nothing is currently known about the role of AHNAK in HCMV replication. In this case, the heavy-labeled version of the peptide produces a peak nearly twice as large as the peak produced by the light version. This is reflected in the ultimate quantification of the protein as it was determined to have an H/L ratio of 1.72.

Changes in protein level across the proteome

Proteins were then quantified by aggregating the individual peptide measurements for each protein. The value obtained for protein quantification is the median value for the set of peptide quantifications.

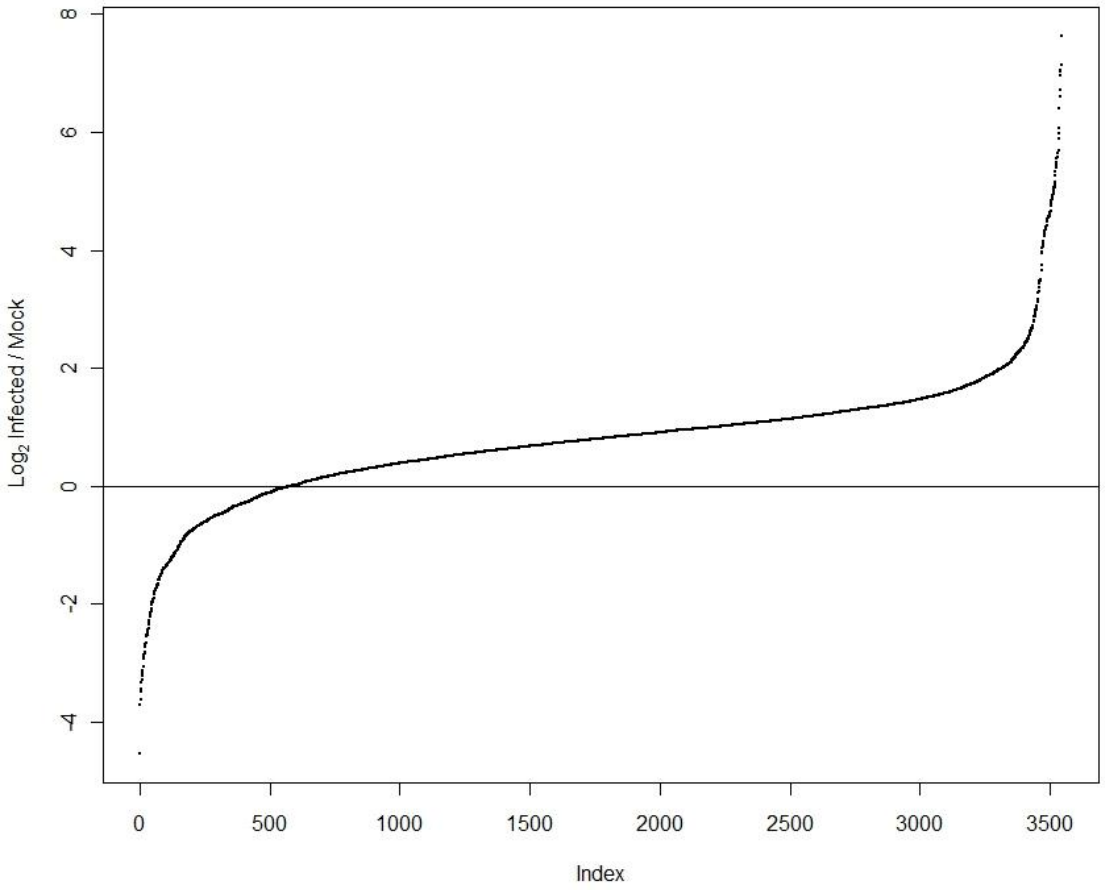


Figure 7 – Changes in abundance of host cell proteins as a result of HCMV infection

Changes in the abundances of all of the proteins identified in the screen are displayed in Figure 7. Proteins are indexed based on the ratio of their change in abundance resulting from virus infection. Those with the lowest indices are seen at lower abundances in infected cells compared to mock infected cells and those with higher indices are found at greater abundance in infected cells compared to mock infected cells. Proteins were detected with infected / mock

ratios ranging from 0.043 to 197.77, demonstrating an ability to determine changes in abundance that range from a 25 fold decrease to a near 200 fold increase in protein level.

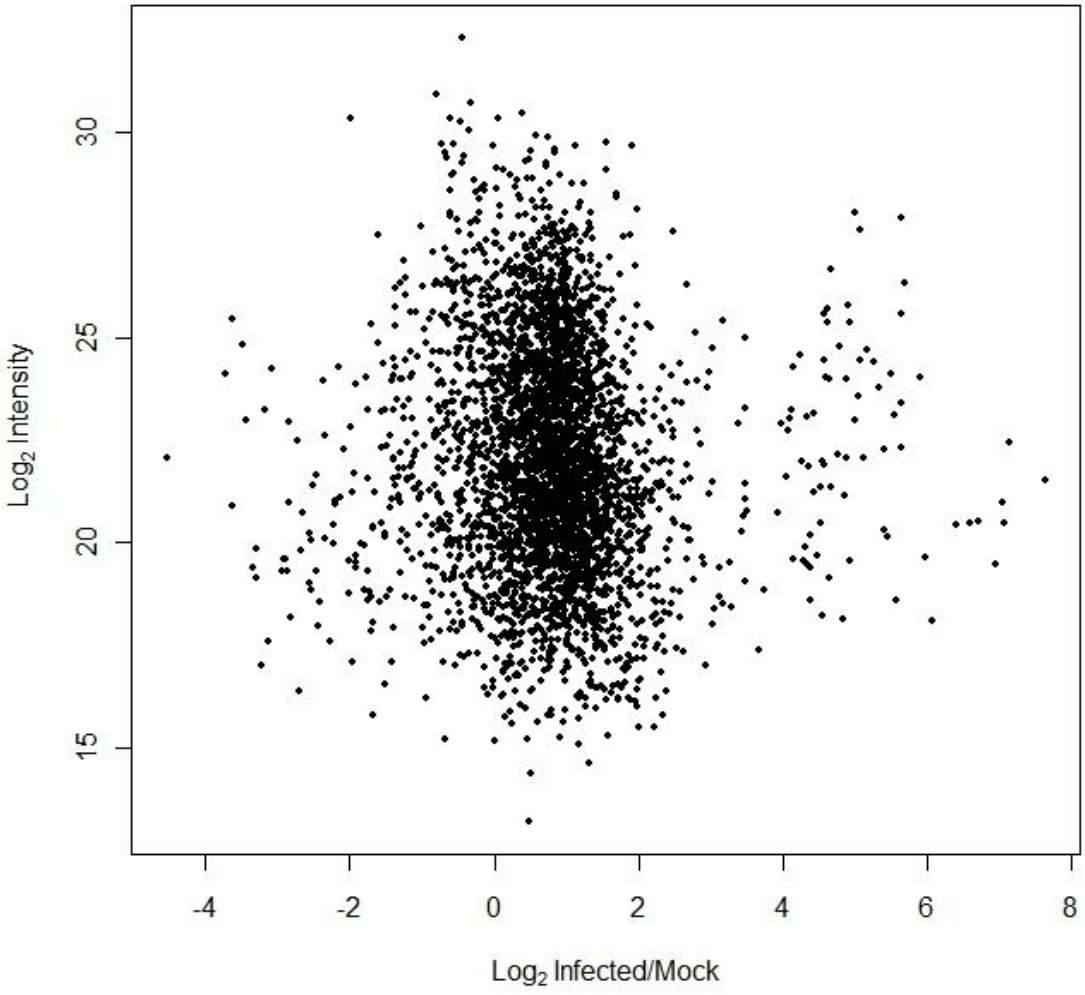


Figure 8 – Relationship of L/H ratios to intensity for proteins detected by the screen

Figure 8 shows the ratio of protein in infected fibroblasts to that in mock infected fibroblasts. There is no correlation (r = 0) between the ratio of protein in infected and mock-infected cells and the intensity of the total signal, indicating that intensity of the total ion signal does not affect the quantification of the protein.

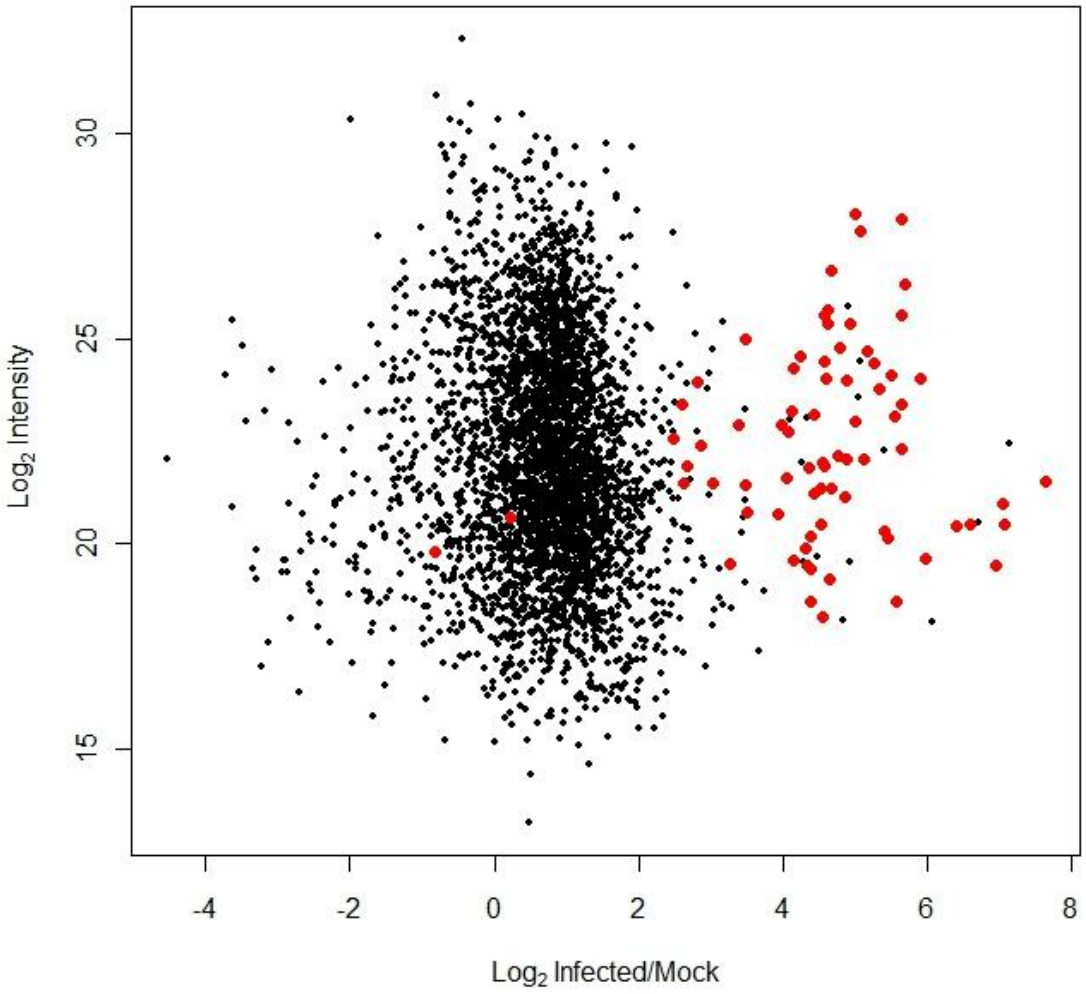


Figure 9 – Virus proteins identified by the screen

Figure 9 highlights in red the viral proteins identified by the MS-based screen. Of the 74 proteins identified, all but 2 show marked increases upon infection. Each of these two proteins was quantified using only 1 peptide, making misidentification a likely explanation for the observed decreased abundance. Additionally, most of the proteins found with marked upregulation under infection were virus proteins. Furthermore, there was no observed linear correlation between protein ratio and intensity of the total signal ($r = 0$), suggesting there is no relationship between signal intensity and quantification. All three temporal expression classes of viral proteins (immediate-early, early, and late) were represented in the 72 viral proteins

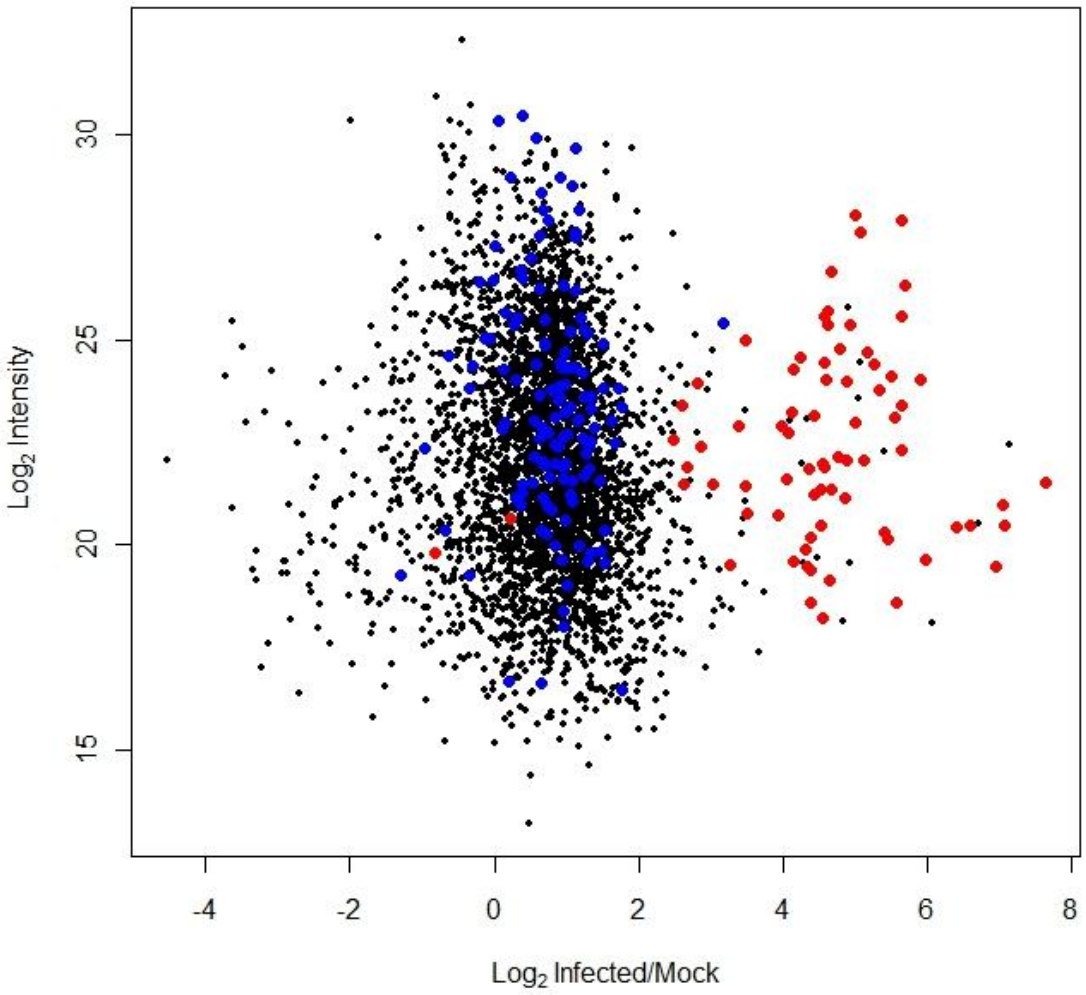


Figure 10 – Metabolic enzymes identified by the screen

In Figure 10 proteins associated with metabolic activities are highlighted in blue. The screen identified 111 metabolic enzymes, their quantifications are found in the appendix. All but 5 of these enzymes were increased as a result of infection. These metabolic proteins generally fall within the range (less than a 1.5 fold change) seen for the majority of the proteins identified in the screen.

Table 1 highlights the ten proteins with the greatest fold increases under virus infection. Only one metabolic protein, glutathione peroxidase 3, was observed in this subset of proteins; however, glutathione peroxidase 3 also has a large posterior error probability (PEP), which

represents the chance that a protein assignment is incorrect. As described earlier, the plurality of proteins have quantifications based solely off of one peptide. Some of these quantifications were likely very reliable and some may be questionable. PEP is an effective measure of the validity of quantification; those proteins with large PEPs are more likely to have incorrect quantifications than those with very small PEPs.

Table 1 – Ten human proteins with greatest observed increases in abundance under infection

Gene Name	Protein description	Peptides	PEP	Ratio L/H
CUX1_HUMAN	Homeobox protein cut-like 1	1	6.73×10^{-6}	141.38
GP101_HUMAN	Probable G-protein coupled receptor 101	1	6.44×10^{-4}	104.50
EPIPL_HUMAN	Epiplakin	1	4.53×10^{-41}	66.77
TMTC2_HUMAN	Transmembrane and TPR repeat-containing protein 2	4	1.68×10^{-2}	41.95
ISG15_HUMAN	Ubiquitin-like protein ISG15	15	1.65×10^{-15}	33.53
CHM4A_HUMAN	Charged multivesicular body protein 4a	6	4.64×10^{-3}	33.13
GPX3_HUMAN	Glutathione peroxidase 3	1	2.08×10^{-2}	30.66
MX1_HUMAN	Interferon-induced GTP-binding protein Mx1	35	3.38×10^{-57}	29.81
PLSI_HUMAN	Plastin-1	3	1.78×10^{-8}	28.43
EF1A2_HUMAN	Elongation factor 1-alpha 2	7	1.83×10^{-37}	25.02

Additionally, the screen identified many proteins downregulated as a result of HCMV infection. The top ten most downregulated proteins are shown in Table 2. Interestingly, three different collagen proteins were observed to be greatly downregulated in HCMV-infected fibroblasts. Also, tumor necrosis factor receptor is downregulated in infected cells, suggesting a possible regulation of apoptosis pathways by the virus.

Table 2 – Ten human proteins with greatest observed decreases in abundance under infection

Gene Name	Protein description	Peptides	PEP	Ratio L/H
VASN_HUMAN	Vasorin	5	1.49×10^{-8}	0.043
PTK7_HUMAN	Tyrosine-protein kinase-like 7	12	1.31×10^{-56}	0.077
CO1A1_HUMAN	Collagen alpha-1(I) chain	23	6.00×10^{-77}	0.081
MYOV2_HUMAN	Myeloma-overexpressed gene 2	1	8.53×10^{-5}	0.082

	protein			
CO1A2_HUMAN	Collagen alpha-2(I) chain	11	1.11×10^{-19}	0.089
CTHR1_HUMAN	Collagen triple helix repeat-containing protein 1	5	4.14×10^{-3}	0.093
ELN_HUMAN	Elastin	2	3.44×10^{-6}	0.099
SDC3_HUMAN	Syndecan-3	1	1.98×10^{-4}	0.102
PCDGM_HUMAN	Protocadherin gamma-C5	1	8.80×10^{-3}	0.103
TR11B_HUMAN	Tumor necrosis factor receptor superfamily member 11B	1	2.83×10^{-4}	0.108

Changes in the levels of enzymes in key pathways

In addition to the proteome-wide analysis made possible by viewing the abundance changes as a whole, quantification for individual metabolic enzymes can be used to determine the effect of virus infection on specific enzymes and pathways. To facilitate a more focused analysis of the screen, metabolic proteins involved in key pathways (i.e. those showing the greatest alterations in metabolic flux upon HCMV infection) were identified.

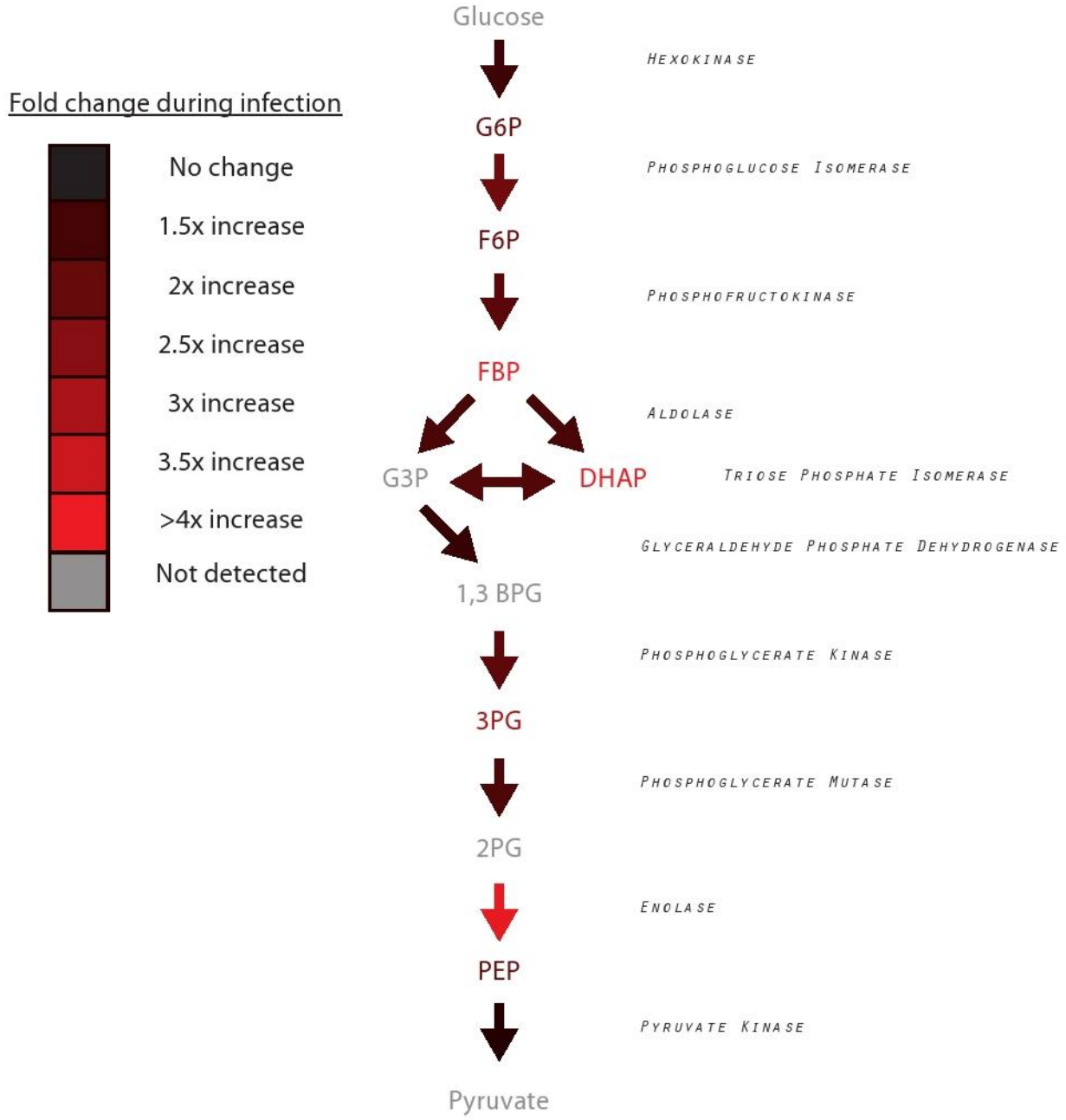


Figure 11 – Changes in abundance of glycolytic enzymes and glycolytic metabolites (from Munger et al. 2006) (increased abundance under infection indicated by brighter red color)

All of the enzymes in the glycolytic pathway were detected in the mass spectrometry-based screen of the HCMV-infected fibroblast proteome. If a protein contains multiple subunits, or multiple isozymes were detected (for example, three forms of phosphofructokinase were

quantified), fold changes in Figure 11 are the average fold change of the subunits or isozymes. Additionally, all showed at least a slightly increased abundance as a result of HCMV infection. These changes in abundance ranged from a 1.04 fold increase in the level of pyruvate kinase to an 8.98 fold increase in the abundance of gamma-enolase, one of the isozymes of the enzyme enolase.

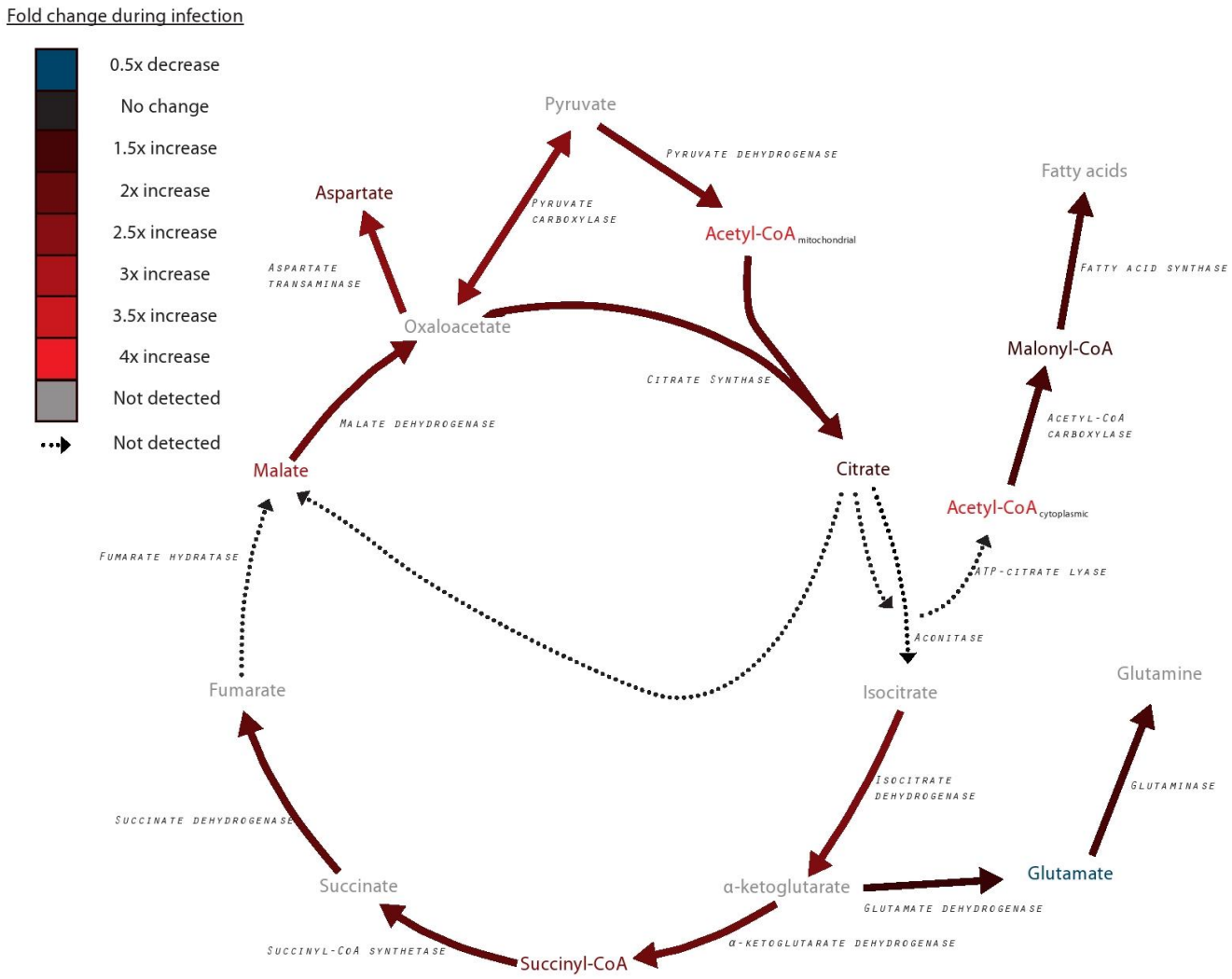


Figure 12 – Changes in abundance of enzymes and of the TCA cycle, fatty acid synthesis, and several anapleurotic reactions. Changes in metabolite concentrations from Munger et al. 2006.

As Figure 12 demonstrates, many TCA cycle enzymes showed increased abundance as a result of virus infection. Unfortunately, unlike the glycolytic pathway, the screen was not able to

identify several key enzymes. The screen did not detect any peptides originating from either aconitase or ATP-citrate lyase, so quantification of these proteins was impossible.

Changes in levels of subunits of multisubunit complexes

In addition to observed changes in levels of enzymes in key metabolic pathways, changes in levels of proteins making up multisubunit complexes were also of interest. If it is true that the virus increases the synthesis of specific proteins, it would seem likely that enzymes involved in multisubunit complexes will show changes consistent in direction and magnitude. 13 of the proteins comprising the protein complex ATP synthase were identified in the screen. All of the identified proteins showed consistent levels of increase in abundance (between 1.80- and 2.62-fold). This is a narrow range of increases compared to the wide range of observed changes in protein level (0.043- to 191.77-fold) seen as a result of infection suggests that changes in ATP synthase subunits may be controlled by HCMV in a precise manner.

Table 3 - Changes in abundance of the subunits of ATP synthase

Gene Name	Protein description	Peptides	PEP	Ratio L/H
ATP8_HUMAN	ATP synthase protein 8	2	1.34×10^{-4}	2.52
ATP6_HUMAN	ATP synthase subunit a	1	3.42×10^{-5}	1.80
ATPA_HUMAN	ATP synthase subunit alpha	27	1.18×10^{-188}	2.13
AT5F1_HUMAN	ATP synthase subunit b	10	4.92×10^{-23}	2.34
ATPB_HUMAN	ATP synthase subunit beta	27	0	2.20
ATP5H_HUMAN	ATP synthase subunit d	8	1.92×10^{-23}	2.36
ATPD_HUMAN	ATP synthase subunit delta	2	3.58×10^{-10}	2.24
ATP5I_HUMAN	ATP synthase subunit e	3	3.22×10^{-5}	2.62
ATPK_HUMAN	ATP synthase subunit f	2	$6.65x\ 10^{-12}$	2.09
ATP5L_HUMAN	ATP synthase subunit g	5	2.29×10^{-14}	2.84
ATPG_HUMAN	ATP synthase subunit gamma	8	3.48×10^{-36}	2.07
ATPO_HUMAN	ATP synthase subunit O	8	1.27×10^{-25}	2.28
ATP5S_HUMAN	ATP synthase subunit s	1	3.96×10^{-4}	1.95

Additionally, all of the components of both isozymes of isocitrate dehydrogenase (NAD^+-dependent and $NADP^+$-dependent) identified by the screen showed increases of between 1.12

fold and 3.32 fold under infection. Of note is the observation that the subunits of each isozymes demonstrated similar levels of changes, suggesting a tightly controlled regulation of multisubunit complexes by HCMV, with differential regulation even among different isozymes of enzymes.

Table 4 - Changes in the abundance of the subunits of isocitrate dehydrogenase

Gene Name	Protein description	Peptides	PEP	Ratio L/H
IDH3A_HUMAN	Isocitrate dehydrogenase [NAD] subunit alpha	9	1.24×10^{-14}	3.17
IDH3B_HUMAN	Isocitrate dehydrogenase [NAD] subunit beta	12	1.21×10^{-21}	3.32
IDH3G_HUMAN	Isocitrate dehydrogenase [NAD] subunit gamma	2	1.52×10^{-6}	2.87
IDHC_HUMAN	Isocitrate dehydrogenase [NADP] cytoplasmic	32	7.05×10^{-64}	1.12
IDHP_HUMAN	Isocitrate dehydrogenase [NADP]	26	7.96×10^{-41}	1.25

Quantifications for the observed 111 metabolic proteins are provided in the appendix. While most of these proteins showed increased abundance as a result of infection, some showed downregulation in HCMV-infected cells.

Gene Ontology Enrichment Analysis

The Gene Ontology (GO) project is a collaborative effort to classify genes and their protein products. Ultimately, the goal of the project is to generate a database of protein entries and annotate each entry with terms corresponding to different attributes of the protein. These terms range from the cellular component (i.e. subcellular localization of protein), to the biological processes in which they function, to the general molecular function of the protein. Examples of terms include “nucleolus”, corresponding to proteins that localize in the nucleolus, “membrane fusion”, for any proteins that are involved in membrane fusion, and “NAD^+ binding”, for any proteins with binding sites for NAD^+. GO terms allow for the easy sorting and

manipulation of large proteomic and genomic data sets to identify common patterns. To this end, GO analysis provides an intermediate level of detail; the patterns of change revealed are more specific and focused than proteome-wide patterns, such as those in Figure 7, but also less targeted than those looking at individual proteins, such as those in Figure 11. For example, if a condition alters secretory pathways, one would expect to observe changes in the levels of a high fraction of proteins with GO terms like "Golgi apparatus", "endoplasmic reticulum", or "ER to Golgi vesicle-mediated transport".

To analyze patterns of up or down regulation in the proteins identified in this study we used the Database for Annotation, Visualization and Integrated Discovery (DAVID). DAVID is one of several Internet-based applications for GO term enrichment, connecting user-generated lists of proteins to the current GO term database. Lists of UniProt Accession numbers were generated for proteins with normalized (so the median H/L ratio was 1) increases in abundance greater than 1.5-fold (880 total), proteins with normalized decreases in abundance of 1.5-fold or greater (598 total), and those proteins with less than a 1.5-fold change in abundance in either direction (1993 total). The choice of a normalized 1.5-fold change in either direction was an arbitrary cut-off producing groups of the desired size. The lists of accession numbers were submitted to DAVID, which linked the accession numbers to GO annotations.

The first major GO category is the cellular compartment in which the protein is localized. Molecular functions, another category, describe the major activities of a protein inside a cell. Examples of molecular functions include "binding" and "enzyme activity".
Biological process GO terms are broader terms that capture the protein's role in a cell. For example, a biological process GO term could be "cell cycle" or "metabolic processes"

Functional Annotation Charts were generated, listing GO terms with particularly significant presence in the proteins identified in our screen. So-called "fat" GO terms were used, which cover many different levels of GO term specificity. To test for significantly increased clusters of proteins, DAVID uses the human proteome as background and determines the fold-enrichment of certain terms in the list of proteins provided. For example, if proteins with the GO term "ATP binding" made up 5% of the provided proteins but only 2.5% of the background list, the term "ATP binding" would have a 2-fold enrichment in the sample. We generated values for fold-enrichment for various GO terms found in each of our three sets of samples: upregulated, downregulated, and no change. The enrichment values for upregulated and downregulated proteins were then compared to the values for no change to determine the enrichment of certain terms in the upregulated or downregulated fractions.

GO terms with significant differences in fold change compared to the unchanging proteins are listed below in Table 5 and Table 6. While there may seem to be some repetition of terms, since "fat" GO terms were used, these seemingly repetitive terms represent different levels of specificity. For example, "helicase activity" and "helicase, superfamily 1 and 2, ATP binding" may seem to be the same, but actually refer to different levels of specificity.

Table 5 – GO Terms enriched in the upregulated fraction

GO Term	Enrichment in upregulated fraction
Helicase C-terminal	4.12
Helicase, superfamily 1 and 2, ATP-binding	3.90
helicase activity	3.64
HELICc	3.63
mitochondrial membrane organization	3.23
helicase	3.20
TPR	3.19
Q motif	3.16
ATPase activity	3.10
ATP-dependent helicase activity	3.04
purine NTP-dependent helicase activity	3.04

RNA helicase, ATP-dependent, DEAD-box, conserved site	3.02
Nucleic acid-binding, OB-fold	3.02
RNA helicase, DEAD-box type, Q motif	2.88
protein targeting to mitochondrion	2.87
protein localization in mitochondrion	2.87
nucleoid	2.85
mitochondrial nucleoid	2.85
ATPase activity, coupled	2.81
DNA/RNA helicase, DEAD/DEAH box type, N-terminal	2.69
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	2.64
unfolded protein binding	2.59
Chaperone	2.53
mitochondrial transport	2.49
ribosome biogenesis	2.44
mitochondrion organization	2.43

Table 6 – GO Terms enriched in the downregulated fraction

GO Term	Enrichment in downregulated fraction
actin-binding	4.17
amino-acid biosynthesis	3.89
actin binding	3.82
regulation of protein complex disassembly	3.77
cell cortex	3.62
membrane protein	3.59
regulation of cytoskeleton organization	3.32
actin cytoskeleton	3.30
regulation of actin filament-based process	3.25
regulation of actin cytoskeleton organization	3.25
calcium-binding region:2	3.16
cytoskeleton organization	3.09
cytoskeletal protein binding	3.01
calcium-binding region:1	3.00
peroxidase	2.86
response to inorganic substance	2.84
cellular amino acid biosynthetic process	2.83
regulation of actin filament length	2.81
regulation of actin polymerization or depolymerization	2.81
Actin-binding, cofilin/tropomyosin type	2.71
regulation of protein polymerization	2.59
ruffle organization	2.59
endonexin fold	2.50
prenylated cysteine	2.38

response to oxidative stress	2.34
negative regulation of cellular component organization	2.30
regulation of actin filament polymerization	2.30
regulation of organelle organization	2.30
cytoskeleton	2.26

Western blots probing different phosphorylation states of pyruvate dehydrogenase

Of particular interest was the large discrepancy between the increase in metabolic flux through pyruvate dehydrogenase under HCMV infection and the increase in amount of the enzyme present in infected cells. While Munger et al. observed an 84-fold increase in flux from pyruvate to acetyl-CoA, only a 1.88- to 2.80-fold increase in enzyme was observed in the SILAC screen. This large increase in flux paired with a small increase in protein abundance suggests that not only is the virus upregulating the synthesis of the different components of the pyruvate dehydrogenase complex, but it is also increasing the activity of the complex. As discussed earlier, pyruvate dehydrogenase activity is tightly regulated by phosphorylation; the phosphorylated form of the enzyme is inactive and the nonphosphorylated form is active.

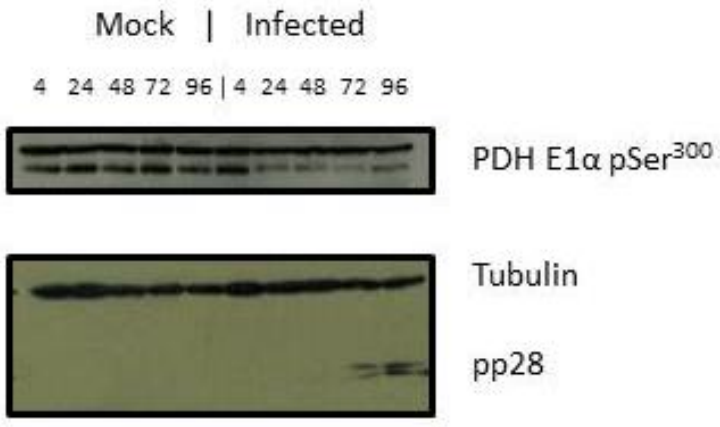


Figure 13 – Western blot for phosphorylated pyruvate dehydrogenase

Western blots shown in Figure 13 suggest a decreased level of the inactive, phosphorylated version of the E1 subunit of the pyruvate dehydrogenase complex. The antibody specific to the phosphorylated serine residue at position 300 produced a significantly smaller band in the 24, 48, 72, and 96 hours post infection timepoints, with tubulin serving as proof of equal protein loading and the HCMV late protein pp28 demonstrating successful infection. However, the band for the phosphorylated serine residue ran as a band with an unexpected mass. The pyruvate dehydrogenase E1 alpha subunit is an approximately 42 kDa protein. Unfortunately, the antibody specific to the phosphorylated E1 subunit produced a band that ran as a protein with a size of 52 kDa. Since phosphorylation can slow a protein's migration through SDS-PAGE by decreasing SDS binding, it is possible that this apparent 52 kDa band is actually the phosphorylated E1 subunit. The antibody also recognized an additional band at ~60-65 kDa that was unaltered by virus infection. This band may represent unaltered PDH protein, doubly or triply phosphorylated PDH protein, or may be a contaminant band of unknown origin.

Discussion

The main goal of this study was to monitor changes in the host cell proteome of fibroblasts infected with HCMV. This proteomic data can be compared with what is already known about the metabolomics and fluxomics of HCMV infection to better understand the virus's mechanisms of metabolic control. Previous studies have indicated that flux through glycolysis, the TCA cycle, and fatty acid synthesis increase as a result of infection with HCMV. Since metabolic reactions are catalyzed by enzymes, a true understanding of how the virus alters metabolic flux requires knowledge of the changes in enzyme activities.

Flux through a metabolic pathway can be controlled by many factors. Firstly, the concentration of the enzymes of the pathway directly affects the flux through the pathway, as a greater concentration of enzyme will often result in greater flux. A proteomic study such as this one can help elucidate changes in the levels of various important enzymes. However, there are other mechanisms of altering metabolic flux. Some enzyme activities are regulated allosterically when effector molecules bind the protein at sites other than the active site, altering their activity. Additionally, post-translational modifications can alter an enzyme's activity and the metabolic flux through the reaction is catalyzes. Adding even more complexity, subcellular localization can affect metabolic flux – if an enzyme is unable to get to the same cellular compartment as its substrate, it does not matter how abundant the enzyme or substrate is, flux through the pathway will be zero.

This study hopes to determine changes in the abundance of various host cell proteins during HCMV infection. As mentioned above, protein abundance is one of several factors that affect metabolic flux. By determining changes in protein abundance resulting from infection, we

can move step closer to understanding how the virus alters flux through host cell metabolic processes to achieve its goal of successful replication.

This study is the first large-scale effort to quantify changes in the host cell proteome resulting from HCMV infection. A large number of proteins were both identified and quantified, covering a broad range of cell processes and systems. This study both provided validation of phenomenon observed through metabolomics and transcriptomics and highlighted several new areas of potential research. For example, the increased pool of phosphoenolpyruvate observed in HCMV-infected fibroblasts can be explained by examining the changes in abundance of enzymes that catalyze reactions forming and consuming phosphoenolpyruvate (Figure 11). The three isoforms of enolase, which produce phosphoenolpyruvate from 2-phosphoglycerate, show an average nearly 4-fold increase as a result of infection. Additionally, cellular levels of pyruvate kinase, which consumes phosphoenolpyruvate to produce pyruvate, do not change as a result of infection. This increase in enolase and static level of pyruvate kinase are likely responsible for the increased pool of phosphoenolpyruvate observed in HCMV-infected fibroblasts.

Comparison of protein quantifications with mRNA transcript level

One of the goals of systems biology is to connect different modalities of data to create more complete pictures of various biological processes. As mentioned earlier, the transcriptome for fibroblasts infected with HCMV was previously documented by Munger et al[21, 25]. This data detailing log-transformed changes in mRNA transcript abundance was stored in the Princeton Microarray Database (PUMAdb). We matched gene IDs from this data to UniProt proteins identifications, allowing those changes in mRNA transcript levels to be easily compared to the

changes in protein level observed in this study. This analysis found little in terms of correlation between changes in protein abundance and changes in mRNA level as a result of infection.

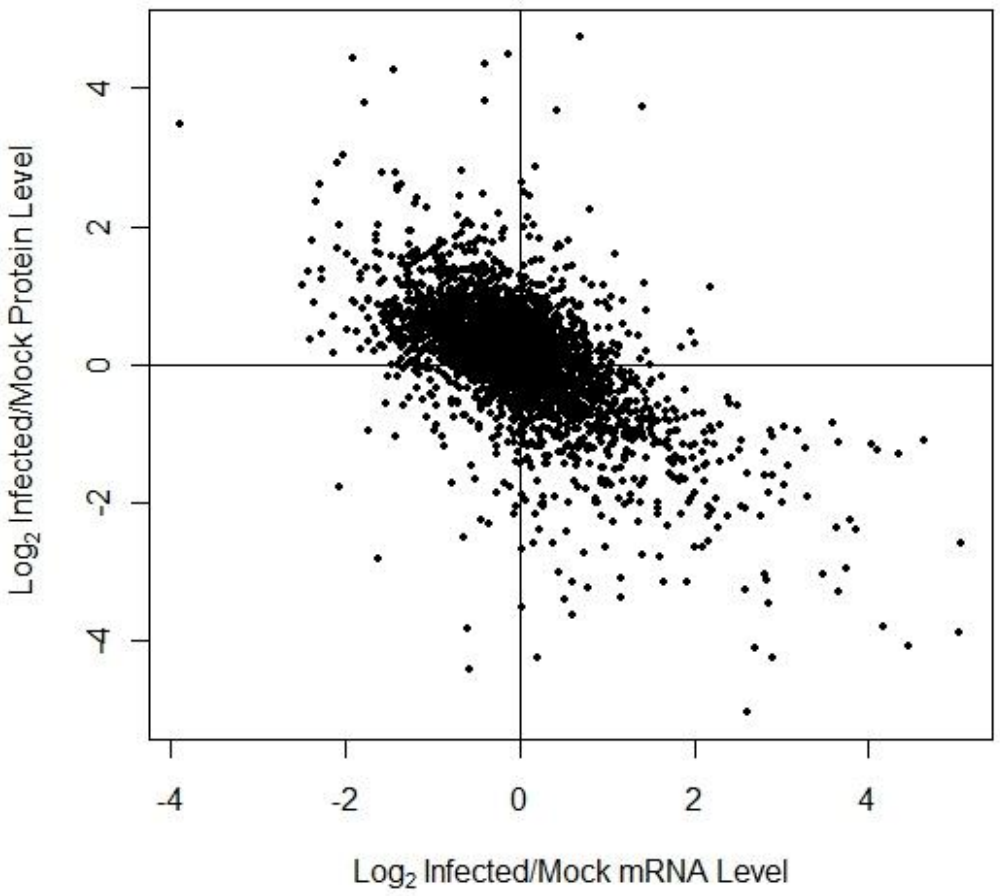


Figure 14 – Changes in mRNA transcript level compared to changes in protein level

Of the 2697 proteins with both protein abundance data and mRNA transcript level, 286 showed both a decrease in protein level and mRNA level. 1172 proteins showed decreases in mRNA level but increases in protein level and 766 had increased mRNA transcript level and decreased protein levels. Finally, both the mRNA level and protein level increased upon infection in the remaining 473 proteins. These proteins with codirectional changes in infection correspond to 28.1% of proteins identified, the remaining 71.9% of mRNA-protein matches showing opposite directions of change. A simple least-squares regression produced a correlation

(r) of -0.5906 and an R^2 of 0.3488. Though these correlation values are weak, the negative sign indicates that in our data there is an indirect relationship between mRNA transcript level and protein level. However, the small value for R^2 reveals the correlation is not strong and visual inspection of Figure 14 reveals a swarm-like relationship between transcript level and protein level, suggesting the indirect relationship is not significant. Due to the suspect nature of this result, we verified that our script for matching genes with proteins produced the correct output for several sample proteins. Since there were no errors in this matching process, it is possible that the data stored in PUMAdb is incorrect and does not truly reflect changes in the transcriptome during HCMV infection.

While this first pass analysis suggests there seems to be no direct relationship between change in mRNA transcript level and change in protein level, it may be beneficial to repeat the microarray experiment to ensure the quality of the mRNA transcript level data. However, given the currently available data, this study provides further support for the observation that changes in the transcriptome are not good predictors of changes in the proteome[41-43].

Comparison of protein quantifications with flux changes

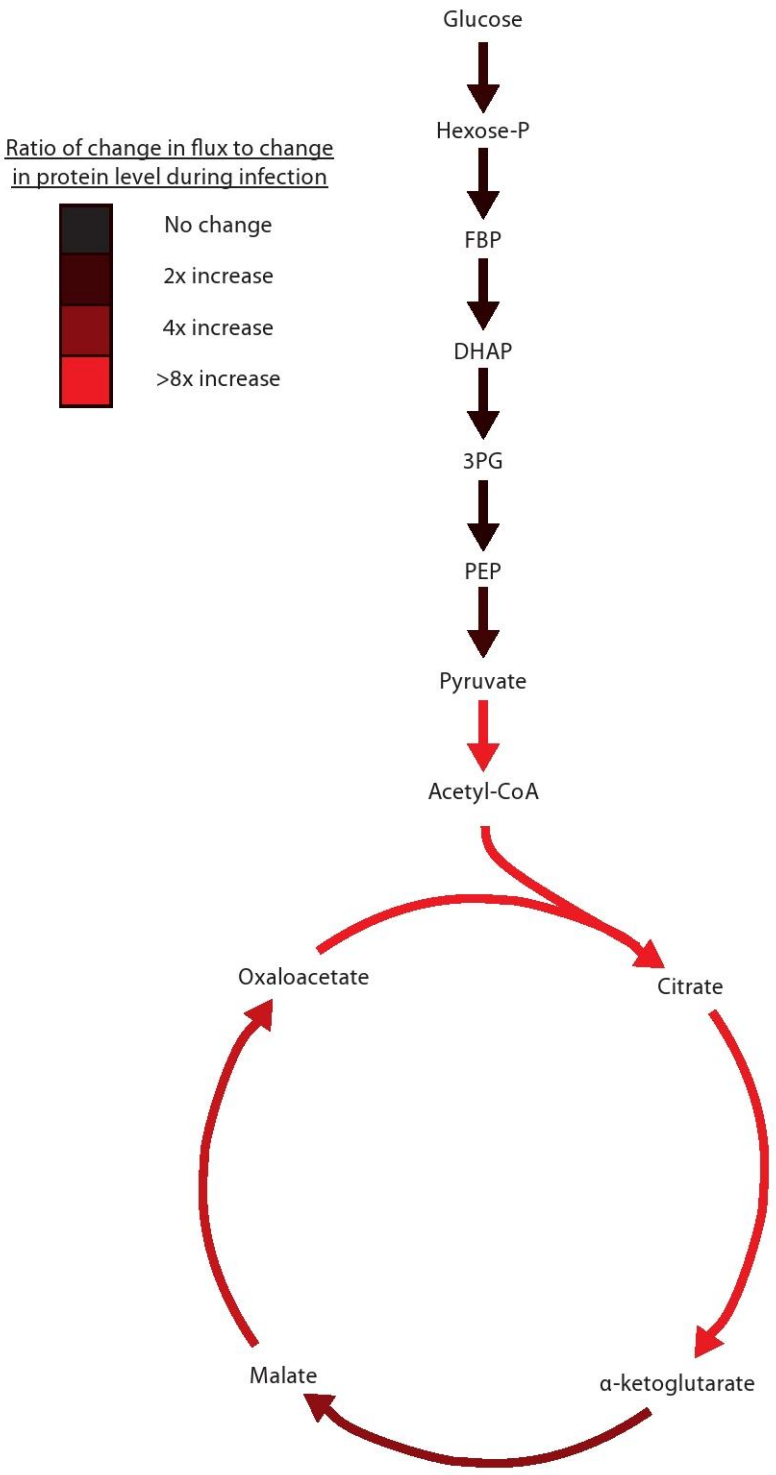


Figure 15 – Ratio of changes in flux to change in enzyme level during infection.

Figure 15 depicts the ratio of the change in flux through key reactions of glycolysis to the change in level of the enzyme (or the average change in abundance of subunits or isozymes of the enzyme) catalyzing the reaction. The ratios of change in flux to change in protein level for the glycolytic enzymes all are close to 1, indicating the changes in flux observed for glycolysis mirror the changes in protein concentration. These ratios suggest the virus increases host cell glycolytic flux primarily by increasing the abundance of various glycolytic enzymes within the cell. According to the Michaelis-Menten kinetic model, reaction velocity is dependent on the concentration of substrate and directly dependent on the concentration of enzyme. Since V_{max} is directly dependent on enzyme concentration, and initial velocity is directly dependent on V_{max}, reaction velocity is directly dependent on enzyme concentration. This direct dependence implies that if changes in flux mirror changes in enzyme concentration, it seems likely that there is a causal relationship between enzyme level and flux. However, this correlation does not prove causation; further work would be needed to fully elucidate the virus's mechanisms of metabolism modification.

The flux changes observed for the reactions of the TCA cycle do not fall in line with the changes in protein level for the enzymes catalyzing these reactions. There are several possible explanations for the observation that activity of the enzymes does not correlate with changes in enzyme abundance. Perhaps the enzymes of the TCA cycle are not operating near their V_{max}. If this was the case, simply adding more substrate will increase flux. Since there is greatly increased flux through the pyruvate dehydrogenase complex, it could be that the virus exerts an extra level of control on only that one key step. The greatly increased mitochondrial pool of acetyl-CoA as a substrate could push the enzymes of the TCA cycle to greater flux than would be predicted simply by examining changes in enzyme level.

It seems likely that the activity of the pyruvate dehydrogenase complex is affected by more than simply enzyme abundance. As noted above, the subunits of the pyruvate dehydrogenase complex were increased between 1.5- and 3-fold under infection, but had an 84-fold increase in metabolic flux[21]. This large discrepancy between the change in enzyme level and flux increase suggests that there is an additional level of regulation by the virus. As mentioned earlier, the pyruvate dehydrogenase complex contains a bound kinase and phosphatase that alter the activity of the complex, inhibiting it through phosphorylation. Calcium ions increase the activity of pyruvate dehydrogenase phosphatase, which dephosphorylates the E1 subunit of the pyruvate dehydrogenase complex, activating it[29]. Since HCMV infection causes calcium release from the endoplasmic reticulum[30], these calcium ions could easily travel to the mitochondria and activate pyruvate dehydrogenase phosphatase. This potential regulation of a key enzyme by the virus seems biologically plausible, but more experiments would be needed to determine the extent of this potential regulation.

Future work could explore the effect of calcium release on pyruvate dehydrogenase activity, treating infected cells with compounds that chelate calcium ions. This chelating agent would sequester calcium and prevent it from activating the phosphatase. This experiment would help determine if flux changes through pyruvate dehydrogenase under infection with HCMV was due in part or wholly to calcium-mediated enzyme regulation.

Additionally, HCMV could cause these massive changes in flux seen in the TCA cycle (Figure 15) by modifying enzymes of the cycle directly, instead of by increasing their substrate concentration. Calcium release from the ER could activate other enzymes of the TCA cycle; calcium ions are known to activate both isocitrate dehydrogenase and alpha-ketoglutarate dehydrogenase[29, 52]. The potential modulation of host cell enzyme activities by viruses through

calcium release from the ER is certainly an avenue for possible research identified because of data obtained through this proteomic screen.

Limitations of quantification with MaxQuant ("All-or-nothing" peptides)

Initially, samples were processed by MaxQuant (the software used for peak detection and quantification) using the correct mass shift for lysine (Δ6 kDa) but an incorrect mass shift for arginine (Δ6). As a result, the algorithm used by MaxQuant to find SILAC pairs failed to recognize peptides containing arginine, leading to the misquantification of many proteins. While a mistake like this would not normally warrant mention in this paper, it ultimately lead to an important discovery about the capability of MaxQuant to quantify proteins present in one sample but not in another. The resultant plot of heavy and light isotope intensities in Figure 16 is based off of this incorrect search parameter.

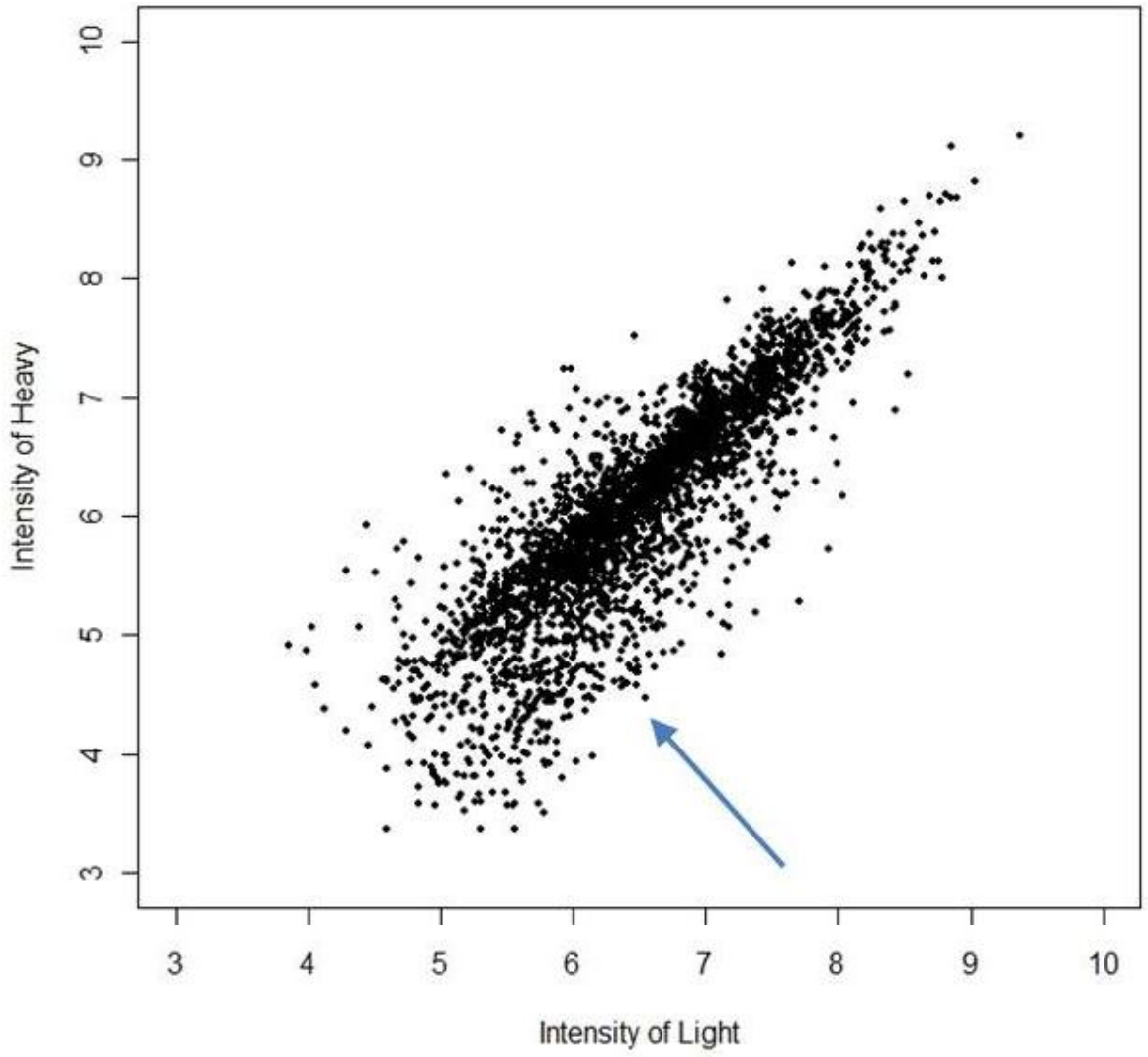


Figure 16 – Intensities of the heavy and light signals for proteins using an incorrect mass shift for arginine (Δ6 instead of Δ10)

When the incorrect mass shift was used for peptides containing arginine, the calculated ratio of heavy to light (infected to mock-infected) would likely be too low, as no peptide would be found at Δ6 Da, but instead would be at Δ10 Da. This would result in many more missed SILAC pairs and errors in quantification. The area containing these proteins with altered quantification is indicated with the arrow; these proteins have an uncharacteristically high intensity in the light sample and low in the heavy sample.

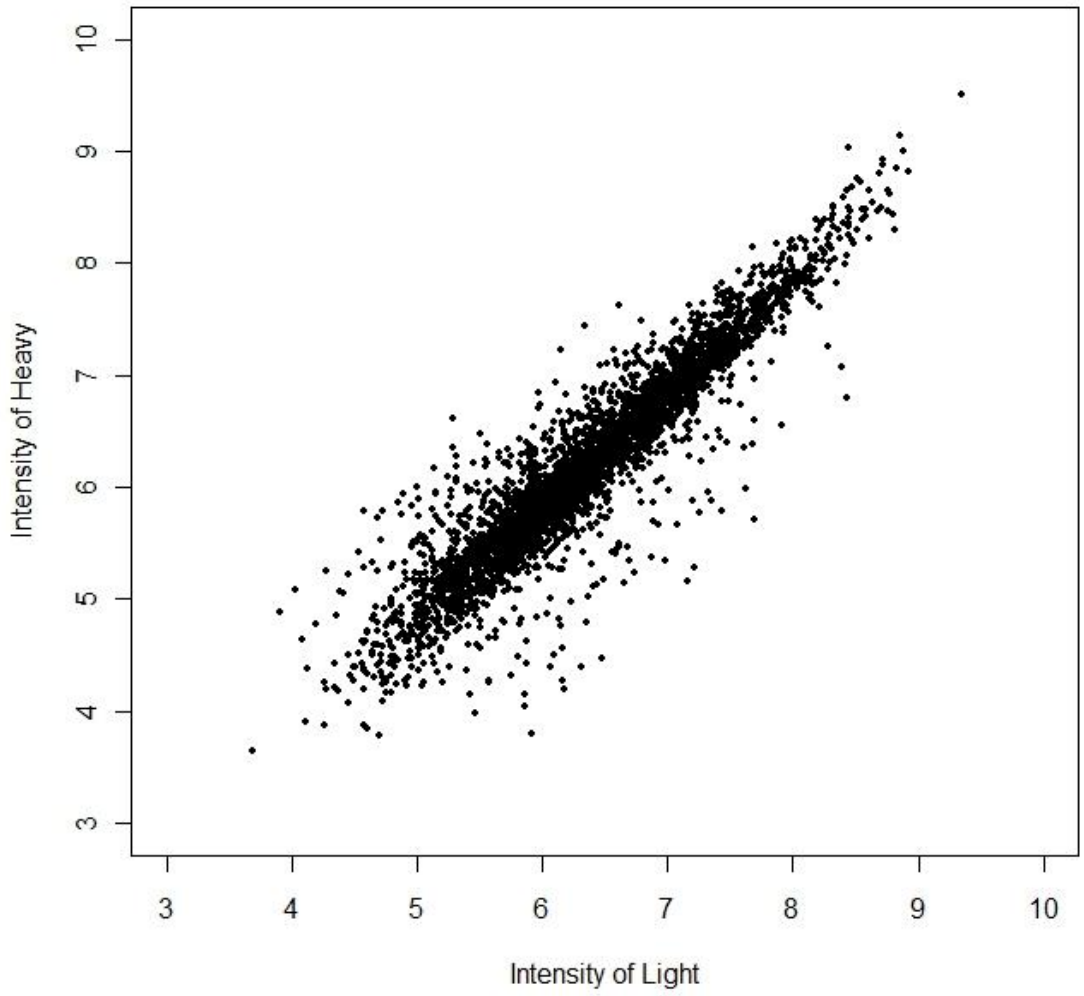


Figure 17 – Intensities of the heavy and light signals for proteins using the correct mass shift for arginine (Δ10)

The plot of intensities in Figure 17 is derived using peptides analyzed with the correct mass shift for arginine. Figure 17 is missing the swarm of data points with much greater intensities in the light sample indicated by the arrow in Figure 16. Many more SILAC pairs were identified using the correct mass shift input.

One would expect any peptides quantified using an incorrect mass to have zero intensity of heavy labeled peptides. This would mean that any protein quantified based off of peptides containing only arginine should be unquantifiable. Since many proteins were quantified based off the quantification of only 1 peptide, it seems likely that a fraction of these were quantified using peptides containing arginine. However, no proteins are found on the x-axis, which would indicate the software failed to identify heavy isotope-containing peptides. To further explore this peculiarity, we examined the basis of quantification for viral proteins.

Interestingly, MaxQuant was able to quantify viral proteins and produce a ratio of their abundances in infected and non-infected cells. Since there should theoretically be no virus proteins in the mock-infected cells, these ratios would seemingly be impossible to calculate. SILAC-based mass spectrometry is only able to obtain relative quantifications of proteins, as without an external standard, absolute quantification is impossible due to unpredictable ionization efficiencies. Therefore, quantification of the virus proteins should theoretically not occur, since there should be no observable “heavy” virus proteins with which to compare the light proteins.

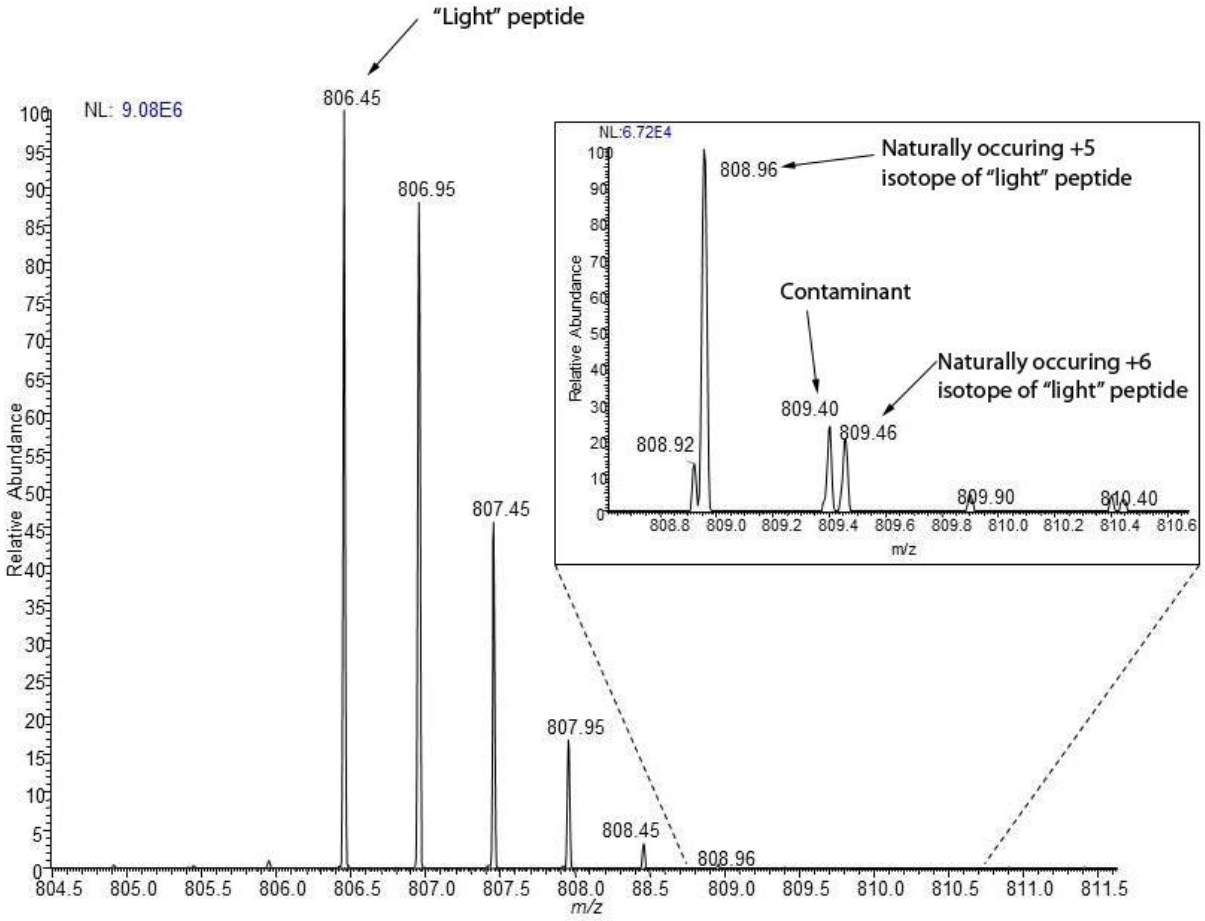


Figure 18 – Quantification of a peptide from pp65 (a viral structural protein) through SILAC

To understand how the MaxQuant software was able to detect and assign a relative abundance to viral proteins, the spectra of individual virus peptides were analyzed. The M+1 peak for the spectra in Figure 18 is seen at 806.45 m/z. The next peak is at 806.95 m/z, indicating a mass shift of 0.5. Since the peptide responsible for this M+2 peak is actually only one neutron heavier, the mass shift of 0.5 indicates the peptide is doubly charged and is in the +2 state. The peptide in question has the sequence NLVPMVATVQGQNLK, with only one lysine and no arginine. This means that the heavy peak will have a predicted mass of around 809.45. The inset spectrum in Figure 18 highlights the region of the spectra where the Δ6 peptide should be found.

The M+6 peak is visible at 809.46 m/z, but there is also a contaminating ion present at 809.40 m/z which also produces a peak pattern consistent with being a +2 ion (see the peaks at 808.92, 809.90, and 810.40 m/z). Comparing the peak heights for the parent ion and the contaminant ion peak produces an approximate H/L ratio of 0.003 to 1. This quantification is then combined with the quantification for the other peptides from the same protein. After all of the runs are complete, the median peptide quantification is read out as the quantification for the protein; the variability in virus protein quantification comes from the strength of signal for various contaminant ions coeluting with the peptides in question.

This analysis of the mechanism through which MaxQuant quantifies a viral protein highlights one of the limitations of the software, explaining how the software was able to calculate quantifications for peptides containing arginine when an incorrect mass shift was used. Since SILAC is a method for determining relative protein concentrations, in order to quantify an "all-or-nothing" protein, which is present in one sample but not the other, the software must find a peak that could possibly represent a heavy or light version of a protein not actually present. Sometimes, a contaminant ion fills that role. Other times, a naturally occurring +6 or +10 isotope could be mislabeled as part of a SILAC pair.

Furthermore, this analysis suggests that some viral proteins may be present in infected cells but not identified if a naturally occurring isotope or contaminant ion failed to act as a its "pair" for analysis.

Proteins in multisubunit complexes

This experiment revealed consistent upregulation of the proteins subunits comprising ATP synthase. It was previously noted that ATP production in HCMV-infected fibroblasts was greater than in mock-infected fibroblasts; this could indicate many things. Since the virus

increases flux through key energy-producing metabolic pathways, such as glycolysis and the TCA cycle, this increase in ATP could simply be due to the production of more NADH and increased flux through oxidative phosphorylation and the electron transport chain. However, the marked increase in the abundance of the subunits of ATP synthase suggests the virus also affects flux through this protein by upregulating the synthesis of each of its subunits to a similar degree.

Additionally, subunits of other multisubunit complexes, like the subunits of the pyruvate dehydrogenase complex or isocitrate dehydrogenase complex, show similar changes in concentration under infection. Interestingly, the subunits comprising two different isozymes of the isocitrate dehydrogenase complex showed different changes in abundance under infection, with the NAD^+-dependent isozymes increasing 2.87- to 3.28-fold but the $NADP^+$-dependent version only increasing 1.12- to 1.25-fold (Table 4), hinting at a possible differential regulation of isozyme levels by the virus.

Gene Ontology and Enrichment Analysis

The term most enriched in the downregulated proteins was "actin binding", showing 7.18-fold enrichment in the downregulated protein grouping relative to the proteome, but only 1.72-fold enrichment in the unchanging proteins relative to the proteome. This indicates a 4.16-fold enrichment of "actin binding" proteins in the downregulated fraction compared to the identified unchanging proteins. Since the virus is known to alter the cytoskeleton, it seems that it may accomplish this by decreasing the abundance of proteins that impact actin binding.

Additionally, proteins involved in amino acid biosynthesis are present to a greater extent in the downregulated fraction. While amino acid metabolism during HCMV infection has not been the topic of much study, it was noted that amino acid pool sizes did not change as a result

of infection (except for alanine, which increased under infection)[25]. It seems the high prevalence of this class of enzymes in the downregulated fraction may indicate more complex modification of amino acid production than previously assumed. A more targeted approach would be needed to say anything more substantial about the virus's impact on this amino acid metabolism.

In upregulated proteins, the most enriched GO terms related to DNA helicases. Helicases unzip the DNA double helix to allow for DNA replication. Since DNA replication is essential for virus replication, it seems logical that the virus would upregulate DNA synthesis and increase the level of various proteins required for DNA synthesis. However, HCMV encodes its own helicase, which has been studied in great detail[53-55]. A literature search revealed no documented interactions between the HCMV helicase and human helicase proteins. Since the helicase proteins are so enriched in the upregulated protein fraction, it seems that there may be some potential interaction between host helicase proteins and viral DNA replication. Further experiments would be needed to fully determine the extent of any possible regulation of human helicase proteins by HCMV. For example, one could add siRNAs specific to the various human helicase proteins to infected cells and monitor virus titers. If human helicase proteins are necessary for virus replication, an siRNA specific to a helicase should knock down virus titer.

Additionally, proteins involved with mitochondrial function are enriched in the upregulated fraction. Since the virus is known to modify mitochondrial enzymes to increase lipid and energy production, this upregulation is consistent with the known biology of the virus, providing some validation for the Gene Ontology analysis.

As evidenced by the preceding discussion, Gene Ontology and enrichment analysis can be a useful tool for identifying major patterns of change between different groups of proteins. However, the approach is not comprehensive and certainly is not conclusive proof of the virus's

effect on host cell proteins. There are likely many aspects of virus infection that were not detected through this analysis, and patterns detected may not actually be real effects. In this context, Gene Ontology should be viewed as a way to confirm the results of the study by checking if they align with previously observed phenomenon as well as a way to identify promising new directions for future research.

Comparing the results of this study with those for other viruses

An interesting parallel can be drawn comparing the results of this study with the results of a similar study examining the effect of Hepatitis C virus infection on liver hepatocytes[37]. This study of hepatitis also revealed increased levels of the enzymes in the glycolytic pathway and TCA cycle, mirroring the effect of HCMV on fibroblasts. However, HCV and HCMV are very different viruses, despite having similar acronyms. HCV is an RNA virus with a small genome consisting of only 10 genes. On the other hand, HCMV is a large, double stranded DNA virus with a genome encoding almost 200 proteins.

Both HCV and HCMV are enveloped viruses, suggesting the increased production of lipids seen under infection by both viruses is needed to produce viral lipid envelopes. This overall increase in flux through citrate synthase suggests that increased lipid production is common to infection by the two viruses, and may possibly be a feature common to enveloped viruses.

However, metabolomics profiling of herpes simplex virus (HSV) tells another story. Though HSV is also an enveloped virus and is closely related to HCMV, it has a very different effect on host cell metabolism than either HCV or HCMV. Though changes in levels of metabolic enzymes of HSV-infected cells have not yet been quantified, the drastically different

changes in flux suggest that there will also be significant differences in the patterns of enzyme synthesis regulation by the virus.

The study of virus modulation of host cell metabolic processes seems to be a promising new avenue for antiviral drug development. Additionally, common patterns of metabolic regulation in very different viruses suggest some common mechanism of virus control of host cell metabolism, possibly indicating the same targeted drug may be able to treat two unrelated viruses.

Future Work

This study demonstrates the power of comparative transcriptome, proteome, and fluxome studies in exploring the effect of HCMV infection on cellular metabolism. These techniques used in tandem allow for a rapid but extensive characterization of virus infection, a complex biological process, and generate many ideas and hypotheses for future potential research. Such work will hopefully elucidate more fully the extent of HCMV's modifications of the host cell proteome. Work by Diamond et al. suggests that HCV regulates levels of metabolic enzymes in a temporal manner, with distinct phases of metabolic modification during infection. To determine if a similar pattern is observed for HCMV, future proteomic analysis should examine changes at different timepoints post-infection (4, 24, 72, and 96 hours post-infection). Further comparative studies of the fluxome of infected cells at different time points post infection in combination with proposed temporal proteomic analyses could demonstrate more fully HCMV's control of metabolism.

Additionally, in this study, an untargeted approach was used to choose the peptides that were quantified. For the first study of the proteome of HCMV-infected fibroblasts, this seems to

be the ideal strategy. However, now that this first pass has been completed, future studies can target specific parts of the proteome. Perhaps future mass spectrometry analysis could be tailored to identify specific metabolic enzymes of interest. Selected reaction monitoring is a targeted approach in which proteins of interest are highlighted. Instead of choosing peptides at random to analyze, this approach takes as input the known elution times and masses of specific peptides and quantifies these peptides only. Using selected reaction monitoring would allow for optimal quantification of key metabolic enzymes and complete coverage of pathways of interest.

Changes in the prevalence of post-translational modifications represent another potential area for future study. Though post-translational modifications such as phosphorylation can have a huge effect on protein function, our MS-based analysis detected very few phosphorylated peptides and proteins. Phosphate groups are especially labile and dissociate from peptides very easily as a result of the high-energy collisions necessary for fragmentation in MS/MS analysis. To detect phosphorylated versions of peptides, it is therefore necessary to enrich peptides for phosphorylation, eliminating peptides that are not phosphorylated. Future analysis of phosphoenriched samples can hopefully shed light on the alteration of protein phosphorylation states by HCMV and obtain a clearer picture of how HCMV induces such large changes in metabolic flux of certain reactions.

Additionally, an siRNA screen to knock down viral proteins could elucidate the mechanism through which the virus increases the levels of various metabolic enzymes. If the synthesis of specific host cell proteins were modified by specific virus proteins, adding siRNAs that decrease the synthesis of these viral proteins would produce a measurable difference in protein levels in infected cells, with protein levels returning to near their levels in mock-infected cells.

Finally, as was previously discussed, future work should explore HCMV's mechanisms of increasing flux through the TCA cycle without increasing enzyme levels at the same magnitude. Though looking at post-translational modifications may shed some light on these mechanisms, it seems there may be a simpler way. As was previously discussed, calcium ions are released as a result of HCMV infection. Since calcium ions increase the activity of many of these enzymes, infected cells could be treated with calcium ion release inhibitors or compounds that chelate calcium. Treatment with these compounds could reduce flux through the TCA cycle, inhibit virus replication, and decrease virus titers.

Conclusion

This study demonstrates that mass spectrometry-based proteomics can be a powerful tool for investigating the complex interactions of a virus and a host cell. Combining proteomic analysis with different modalities of chemistry and biology, such as genomics and lipidomics, allows for a more complete understanding of the effect of HCMV on host cell biological processes, both illuminating the underpinnings of virus infection and highlighting areas of metabolism where more complex modification of host processes is occurring. Additionally, taken in combination with data generated through systems biology analyses of other viruses, this study suggests virus modification of host enzymes is similar across groups of very different viruses. Future work could help identify novel targets for antiviral therapies, specific to not only HCMV but possibly of use in treating other viruses.

Appendix

Relationship between posterior error probability and signal intensity

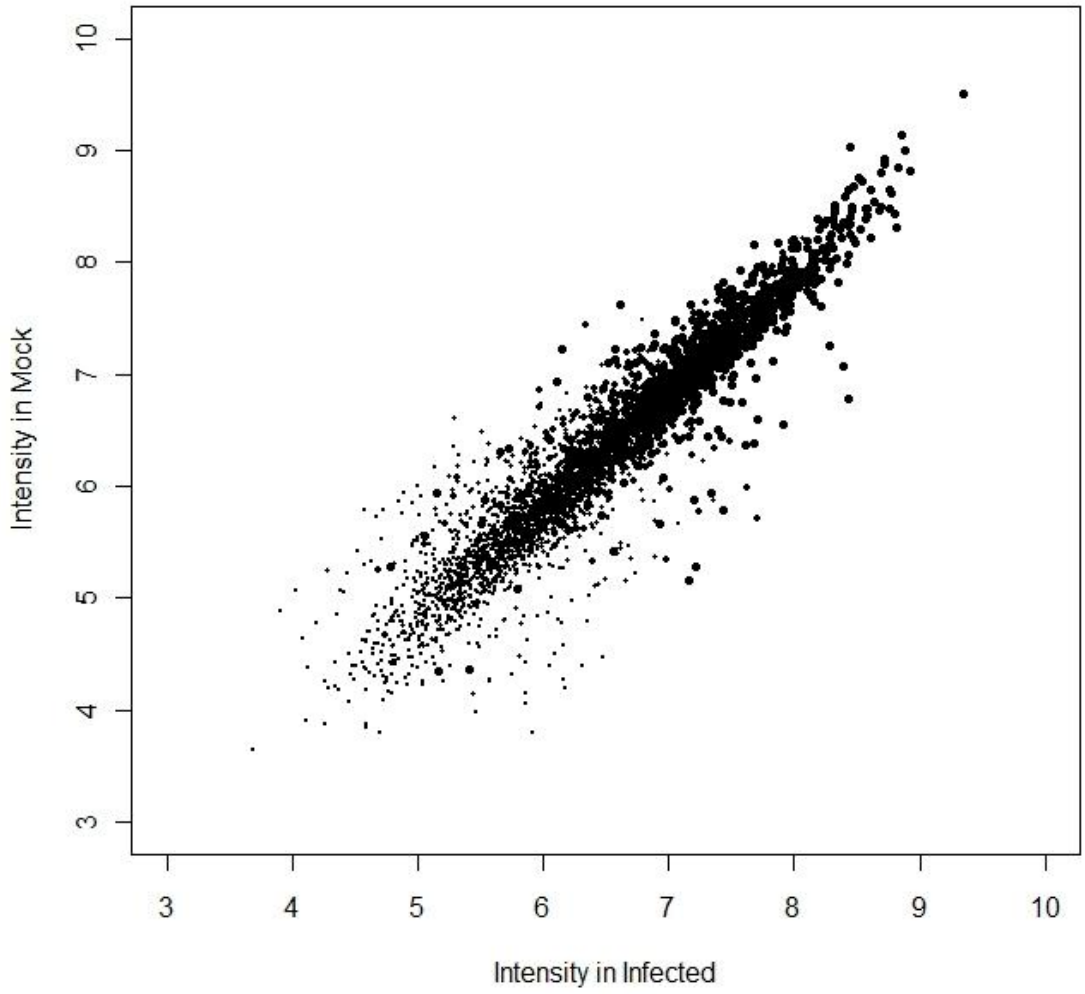


Figure 19 – Relationship between Posterior Error Probability (PEP) and Intensities of Heavy and Light Signals

Figure 19 demonstrates the relationship between intensity of signal and the posterior error probability (PEP), or the likelihood that a protein assignment was incorrect. In Figure 19, the size of the point corresponding to each protein is reflective of the value of PEP. The largest points are those with the smallest PEPs, so we are most certain these identifications are correct. The smaller

points have larger PEPs, indicating less confidence in the identification of the protein. As is expected, there is generally greater confidence in the identification of proteins detected with greater intensity.

Metabolic proteins identified by the SILAC screen

Protein Names	Protein Descriptions	Peptides	PEP	Ratio L/H
ACACA_HUMAN	Acetyl-CoA carboxylase 1	10	2.45×10^{-15}	1.95
ACAD9_HUMAN	Acyl-CoA dehydrogenase family member 9	9	3.75×10^{-34}	1.79
ACADV_HUMAN	Very long-chain specific acyl-CoA dehydrogenase	22	1.09×10^{-74}	0.97
ACD10_HUMAN	Acyl-CoA dehydrogenase family member 10	2	2.16×10^{-2}	1.56
ACLY_HUMAN	ATP-citrate synthase	80	9.35×10^{-129}	1.55
ACPM_HUMAN	Acyl carrier protein	6	1.31×10^{-6}	3.07
ACSF2_HUMAN	Acyl-CoA synthetase family member 2	3	7.22×10^{-3}	1.97
ACSF3_HUMAN	Acyl-CoA synthetase family member 3	8	1.54×10^{-16}	2.06
ACSL3_HUMAN	Long-chain-fatty-acid--CoA ligase 3	16	5.16×10^{-29}	1.98
ACSL4_HUMAN	Long-chain-fatty-acid--CoA ligase 4	3	5.85×10^{-8}	1.70
ALDOA_HUMAN	Fructose-bisphosphate aldolase A	63	2.11×10^{-165}	1.57
ALDOC_HUMAN	Fructose-bisphosphate aldolase C	14	2.23×10^{-56}	1.64
ASSY_HUMAN	Argininosuccinate synthase	25	1.23×10^{-75}	0.91
AT5F1_HUMAN	ATP synthase subunit b	20	4.92×10^{-23}	2.34
ATP5H_HUMAN	ATP synthase subunit d	27	1.92×10^{-23}	2.36
ATP5I_HUMAN	ATP synthase subunit e	3	3.22×10^{-5}	2.62
ATP5J_HUMAN	ATP synthase-coupling factor 6	5	4.05×10^{-27}	2.37
ATP5L_HUMAN	ATP synthase subunit g	9	2.29×10^{-14}	2.84
ATP5S_HUMAN	ATP synthase subunit s	1	3.96×10^{-4}	1.95
ATP6_HUMAN	ATP synthase subunit a	2	3.42×10^{-5}	1.80
ATP8_HUMAN	ATP synthase protein 8	3	1.34×10^{-4}	2.52
ATPA_HUMAN	ATP synthase subunit alpha	111	1.18×10^{-188}	2.13
ATPB_HUMAN	ATP synthase subunit beta	165	0	2.20
ATPD_HUMAN	ATP synthase subunit delta	5	3.58×10^{-10}	2.24
ATPG_HUMAN	ATP synthase subunit gamma	21	3.48×10^{-36}	2.07
ATPO_HUMAN	ATP synthase subunit O	28	1.27×10^{-25}	2.28

BCAT1_HUMAN	Branched-chain-amino-acid aminotransferase, cytosolic	5	5.30×10^{-13}	0.51
CISY_HUMAN	Citrate synthase	23	4.41×10^{-46}	1.96
DHE3_HUMAN	Glutamate dehydrogenase 1	37	6.51×10^{-47}	1.31
DLDH_HUMAN	Dihydrolipoyl dehydrogenase	14	4.06×10^{-30}	2.01
ELOV1_HUMAN	Elongation of very long chain fatty acids protein 1	1	4.21×10^{-7}	3.42
ELOV5_HUMAN	Elongation of very long chain fatty acids protein 5	1	6.72×10^{-3}	1.74
ENOA_HUMAN	Alpha-enolase	172	1.75×10^{-297}	1.50
ENOB_HUMAN	Beta-enolase	2	1.26×10^{-18}	1.30
ENOF1_HUMAN	Mitochondrial enolase superfamily member 1	1	5.09×10^{-3}	1.94
ENOG_HUMAN	Gamma-enolase	30	2.34×10^{-85}	8.98
ENOPH_HUMAN	Enolase-phosphatase E1	3	6.81×10^{-16}	2.48
FAS_HUMAN	Fatty acid synthase	141	3.41×10^{-268}	1.60
G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	193	1.09×10^{-274}	1.32
G6PD_HUMAN	Glucose-6-phosphate 1-dehydrogenase	50	3.12×10^{-69}	1.01
G6PI_HUMAN	Glucose-6-phosphate isomerase	50	7.34×10^{-141}	2.18
GPDM_HUMAN	Glycerol-3-phosphate dehydrogenase	12	8.16×10^{-34}	2.08
GSK3B_HUMAN	Glycogen synthase kinase-3 beta	8	3.53×10^{-6}	2.73
GYS1_HUMAN	Glycogen [starch] synthase, muscle	5	6.60×10^{-19}	2.41
HXK1_HUMAN	Hexokinase-1	41	1.89×10^{-78}	1.54
HXK2_HUMAN	Hexokinase-2	3	1.32×10^{-10}	1.32
IDH3A_HUMAN	Isocitrate dehydrogenase [NAD] subunit alpha	9	1.24×10^{-14}	3.17
IDH3B_HUMAN	Isocitrate dehydrogenase [NAD] subunit beta	12	1.21×10^{-21}	3.32
IDH3G_HUMAN	Isocitrate dehydrogenase [NAD] subunit gamma	2	1.52×10^{-6}	2.87
IDHC_HUMAN	Isocitrate dehydrogenase [NADP] cytoplasmic	32	7.05×10^{-64}	1.12
IDHP_HUMAN	Isocitrate dehydrogenase [NADP]	26	7.96×10^{-41}	1.25
K6PF_HUMAN	6-phosphofructokinase, muscle type	10	1.80×10^{-30}	2.52
K6PL_HUMAN	6-phosphofructokinase, liver type	18	5.62×10^{-62}	1.09
K6PP_HUMAN	6-phosphofructokinase type C	22	3.88×10^{-57}	1.97
KPYM_HUMAN	Pyruvate kinase isozymes	221	0	1.04

	M1/M2			
LDHA_HUMAN	L-lactate dehydrogenase A chain	106	4.27×10^{-158}	1.16
LDHB_HUMAN	L-lactate dehydrogenase B chain	83	3.78×10^{-77}	2.24
MAOM_HUMAN	NAD-dependent malic enzyme	12	6.71×10^{-47}	3.41
MAOX_HUMAN	NADP-dependent malic enzyme	14	5.20×10^{-27}	1.93
MDHC_HUMAN	Malate dehydrogenase, cytoplasmic	37	1.07×10^{-85}	2.20
MDHM_HUMAN	Malate dehydrogenase, mitochondrial	36	1.69×10^{-76}	2.19
MPI_HUMAN	Mannose-6-phosphate isomerase	4	4.12×10^{-7}	1.64
NDUA2_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2	2	3.48×10^{-7}	1.68
NDUA4_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4	3	1.45×10^{-3}	2.02
NDUA5_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	5	7.39×10^{-17}	1.93
NDUA6_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6	5	1.14×10^{-2}	2.48
NDUA8_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8	2	5.70×10^{-8}	2.12
NDUA9_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9	10	3.97×10^{-34}	1.86
NDUAA_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10	7	3.76×10^{-20}	1.79
NDUAC_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	7	9.60×10^{-6}	1.44
NDUAD_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13	11	7.80×10^{-8}	1.86
NDUB1_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 1	2	1.09×10^{-7}	1.78
NDUB3_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex	1	1.24×10^{-2}	1.93

	subunit 3			
NDUB4_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4	5	9.28×10^{-7}	1.85
NDUB5_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 5	2	3.90×10^{-3}	2.10
NDUB6_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 6	4	2.61×10^{-6}	2.36
NDUB8_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8	2	2.75×10^{-9}	1.65
NDUB9_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9	3	2.03×10^{-14}	1.69
NDUBA_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10	8	1.17×10^{-15}	1.87
NDUBB_HUMAN	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11	2	1.17×10^{-5}	2.90
NDUF4_HUMAN	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 4	2	1.36×10^{-2}	2.47
NDUS2_HUMAN	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2	13	2.01×10^{-21}	1.89
NDUS3_HUMAN	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3	11	1.11×10^{-42}	1.74
NDUS4_HUMAN	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4	2	7.04×10^{-5}	1.58
NDUS7_HUMAN	NADH dehydrogenase [ubiquinone] iron-sulfur protein 7	1	3.58×10^{-3}	1.60
NDUS8_HUMAN	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8	4	1.70×10^{-4}	1.48
NDUV1_HUMAN	NADH dehydrogenase [ubiquinone] flavoprotein 1	6	1.25×10^{-19}	1.60
NDUV2_HUMAN	NADH dehydrogenase [ubiquinone] flavoprotein 2	4	2.84×10^{-14}	1.79
NQO1_HUMAN	NAD(P)H dehydrogenase [quinone] 1	14	1.64×10^{-25}	0.81

ODBA_HUMAN	2-oxoisovalerate dehydrogenase subunit alpha	1	6.23×10^{-3}	2.01
ODO1_HUMAN	2-oxoglutarate dehydrogenase	27	6.20×10^{-82}	2.82
ODO2_HUMAN	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	14	1.01×10^{-44}	1.62
ODP2_HUMAN	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	18	6.89×10^{-31}	2.14
ODPA_HUMAN	Pyruvate dehydrogenase E1 component subunit alpha, somatic form	6	6.00×10^{-17}	2.32
ODPB_HUMAN	Pyruvate dehydrogenase E1 component subunit beta	17	4.57×10^{-37}	1.88
ODPX_HUMAN	Pyruvate dehydrogenase protein X component	2	6.38×10^{-4}	2.80
PCKGM_HUMAN	Phosphoenolpyruvate carboxykinase [GTP]	9	1.87×10^{-21}	1.55
PDP1_HUMAN	[Pyruvate dehydrogenase [acetyl-transferring]]-phosphatase 1	5	2.84×10^{-5}	1.59
PDPR_HUMAN	Pyruvate dehydrogenase phosphatase regulatory subunit	5	2.13×10^{-10}	1.97
PGAM1_HUMAN	Phosphoglycerate mutase 1	54	9.93×10^{-81}	1.67
PGK1_HUMAN	Phosphoglycerate kinase 1	94	1.11×10^{-151}	1.89
PGM1_HUMAN	Phosphoglucomutase-1	28	4.36×10^{-53}	1.62
PGM2_HUMAN	Phosphoglucomutase-2	11	3.32×10^{-13}	2.54
PGM2L_HUMAN	Glucose 1,6-bisphosphate synthase	3	6.21×10^{-7}	2.24
PLCD3_HUMAN	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase delta-3	2	2.44×10^{-16}	0.41
PRPS1_HUMAN	Ribose-phosphate pyrophosphokinase 1	7	1.52×10^{-9}	1.31
PYC_HUMAN	Pyruvate carboxylase	11	1.18×10^{-25}	2.55
PYGB_HUMAN	Glycogen phosphorylase, brain form	32	9.37×10^{-78}	1.20
PYGL_HUMAN	Glycogen phosphorylase, liver form	25	3.12×10^{-57}	1.50
PYGM_HUMAN	Glycogen phosphorylase, muscle form	1	9.14×10^{-9}	1.23
SERA_HUMAN	D-3-phosphoglycerate dehydrogenase	38	1.86×10^{-38}	0.87

References

(1) Crough, T.; Khanna, R. Immunobiology of human cytomegalovirus: from bench to bedside. *Clin. Microbiol. Rev.* **2009**, *22*, 76-98, Table of Contents.

(2) Kenneson, A.; Cannon, M. J. Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev. Med. Virol.* **2007**, *17*, 253-276.

(3) Ho, M. The history of cytomegalovirus and its diseases. *Med. Microbiol. Immunol.* **2008**, *197*, 65-73.

(4) Kutza, A. S.; Muhl, E.; Hackstein, H.; Kirchner, H.; Bein, G. High incidence of active cytomegalovirus infection among septic patients. *Clin. Infect. Dis.* **1998**, *26*, 1076-1082.

(5) Mutimer, D.; Mirza, D.; Shaw, J.; O'Donnell, K.; Elias, E. Enhanced (cytomegalovirus) viral replication associated with septic bacterial complications in liver transplant recipients. *Transplantation* **1997**, *63*, 1411-1415.

(6) Prosch, S.; Wendt, C. E.; Reinke, P.; Priemer, C.; Oppert, M.; Kruger, D. H.; Volk, H. D.; Docke, W. D. A novel link between stress and human cytomegalovirus (HCMV) infection: sympathetic hyperactivity stimulates HCMV activation. *Virology* **2000**, *272*, 357-365.

(7) Zhang, M.; Yang, Y.; Yang, X.; Cai, J. Human cytomegalovirus infection is a novel etiology for essential hypertension. *Med. Hypotheses* **2011**.

(8) Fowler, K. B.; Boppana, S. B. Congenital cytomegalovirus (CMV) infection and hearing deficit. *J. Clin. Virol.* **2006**, *35*, 226-231.

(9) Britt, W. J.; Boppana, S. Human cytomegalovirus virion proteins. *Hum. Immunol.* **2004**, *65*, 395-402.

(10) Wang, D.; Yu, Q. C.; Schroer, J.; Murphy, E.; Shenk, T. Human cytomegalovirus uses two distinct pathways to enter retinal pigmented epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20037-20042.

(11) Stinski, M. F.; Meier, J. L. In *Immediate-early viral gene regulation and function;* Arvin, A., Campadelli-Fiume, G., Mocarski, E., Moore, P. S., Roizman, B., Whitley, R. and Yamanishi, K., Eds.; Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis; Cambridge, 2007; .

(12) Gibson, W. Structure and formation of the cytomegalovirus virion. *Curr. Top. Microbiol. Immunol.* **2008**, *325*, 187-204.

(13) Noble, S.; Faulds, D. Ganciclovir. An update of its use in the prevention of cytomegalovirus infection and disease in transplant recipients. *Drugs* **1998**, *56*, 115-146.

(14) Sugawara, M.; Huang, W.; Fei, Y. J.; Leibach, F. H.; Ganapathy, V.; Ganapathy, M. E. Transport of valganciclovir, a ganciclovir prodrug, via peptide transporters PEPT1 and PEPT2. *J. Pharm. Sci.* **2000**, *89*, 781-789.

(15) Sharland, M.; Luck, S.; Griffiths, P.; Cotton, M. Antiviral therapy of CMV disease in children. *Adv. Exp. Med. Biol.* **2011**, *697*, 243-260.

(16) Meyer, P. R.; Rutvisuttinunt, W.; Matsuura, S. E.; So, A. G.; Scott, W. A. Stable complexes formed by HIV-1 reverse transcriptase at distinct positions on the primer-template controlled by binding deoxynucleoside triphosphates or foscarnet. *J. Mol. Biol.* **2007**, *369*, 41-54.

(17) Becker, M. N.; Obraztsova, M.; Kern, E. R.; Quenelle, D. C.; Keith, K. A.; Prichard, M. N.; Luo, M.; Moyer, R. W. Isolation and characterization of cidofovir resistant vaccinia viruses. *Virol. J.* **2008**, *5*, 58.

(18) Roush, W. Antisense aims for a renaissance. *Science* **1997**, *276*, 1192-1193.

(19) Maynard, N. D.; Gutschow, M. V.; Birch, E. W.; Covert, M. W. The virus as metabolic engineer. *Biotechnol. J.* **2010**, *5*, 686-694.

(20) Jakobsson, A.; Westerberg, R.; Jacobsson, A. Fatty acid elongases in mammals: their regulation and roles in metabolism. *Prog. Lipid Res.* **2006**, *45*, 237-249.

(21) Munger, J.; Bennett, B. D.; Parikh, A.; Feng, X. J.; McArdle, J.; Rabitz, H. A.; Shenk, T.; Rabinowitz, J. D. Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nat. Biotechnol.* **2008**, *26*, 1179-1186.

(22) Min Jou, W.; Haegeman, G.; Ysebaert, M.; Fiers, W. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature* **1972**, *237*, 82-88.

(23) Horning, E. C.; Horning, M. G. Metabolic profiles: gas-phase methods for analysis of metabolites. *Clin. Chem.* **1971**, *17*, 802-809.

(24) Wenk, M. R. The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **2005**, *4*, 594-610.

(25) Munger, J.; Bajad, S. U.; Coller, H. A.; Shenk, T.; Rabinowitz, J. D. Dynamics of the cellular metabolome during human cytomegalovirus infection. *PLoS Pathog.* **2006**, *2*, e132.

(26) Chambers, J. W.; Maguire, T. G.; Alwine, J. C. Glutamine metabolism is essential for human cytomegalovirus infection. *J. Virol.* **2010**, *84*, 1867-1873.

(27) Brown, J. P.; Perham, R. N. Selective inactivation of the transacylase components of the 2-oxo acid dehydrogenase multienzyme complexes of Escherichia coli. *Biochem. J.* **1976**, *155*, 419-427.

(28) Linn, T. C.; Pettit, F. H.; Reed, L. J. Alpha-keto acid dehydrogenase complexes. X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **1969**, *62*, 234-241.

(29) Denton, R. M.; Randle, P. J.; Martin, B. R. Stimulation by calcium ions of pyruvate dehydrogenase phosphate phosphatase. *Biochem. J.* **1972**, *128*, 161-163.

(30) Sharon-Friling, R.; Goodhouse, J.; Colberg-Poley, A. M.; Shenk, T. Human cytomegalovirus pUL37x1 induces the release of endoplasmic reticulum calcium stores. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 19117-19122.

(31) Spencer, C. M.; Schafer, X. L.; Moorman, N. J.; Munger, J. Human Cytomegalovirus Induces the Activity and Expression of Acetyl-CoA Carboxylase, a Fatty Acid Biosynthetic Enzyme whose Inhibition Attenuates Viral Replication. *J. Virol.* **2011**.

(32) Botelho, D.; Wall, M. J.; Vieira, D. B.; Fitzsimmons, S.; Liu, F.; Doucette, A. Top-down and bottom-up proteomics of SDS-containing solutions following mass-based separation. *J. Proteome Res.* **2010**, *9*, 2863-2870.

(33) Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198-207.

(34) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* **2005**, *40*, 430-443.

(35) Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.; Wouters, E. R.; Senko, M.; Makarov, A.; Mann, M.; Horning, S. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol. Cell. Proteomics* **2009**, *8*, 2759-2769.

(36) Antrobus, R.; Grant, K.; Gangadharan, B.; Chittenden, D.; Everett, R. D.; Zitzmann, N.; Boutell, C. Proteomic analysis of cells in the early stages of herpes simplex virus type-1 infection reveals widespread changes in the host cell proteome. *Proteomics* **2009**, *9*, 3913-3927.

(37) Diamond, D. L.; Syder, A. J.; Jacobs, J. M.; Sorensen, C. M.; Walters, K. A.; Proll, S. C.; McDermott, J. E.; Gritsenko, M. A.; Zhang, Q.; Zhao, R.; Metz, T. O.; Camp, D. G.,2nd; Waters, K. M.; Smith, R. D.; Rice, C. M.; Katze, M. G. Temporal proteome and lipidome profiles reveal hepatitis C virus-associated reprogramming of hepatocellular metabolism and bioenergetics. *PLoS Pathog.* **2010**, *6*, e1000719.

(38) Vogels, M. W.; van Balkom, B. W.; Kaloyanova, D. V.; Batenburg, J. J.; Heck, A. J.; Helms, J. B.; Rottier, P. J.; de Haan, C. A. Identification of host factors involved in coronavirus replication by quantitative proteomics analysis. *Proteomics* **2011**, *11*, 64-80.

(39) Nelson, M. M.; Jones, A. R.; Carmen, J. C.; Sinai, A. P.; Burchmore, R.; Wastling, J. M. Modulation of the host cell proteome by the intracellular apicomplexan parasite Toxoplasma gondii. *Infect. Immun.* **2008**, *76*, 828-844.

(40) Costenoble, R.; Picotti, P.; Reiter, L.; Stallmach, R.; Heinemann, M.; Sauer, U.; Aebersold, R. Comprehensive quantitative analysis of central carbon and amino-acid metabolism in Saccharomyces cerevisiae under multiple conditions by targeted proteomics. *Mol. Syst. Biol.* **2011**, *7*, 464.

(41) de Sousa Abreu, R.; Penalva, L. O.; Marcotte, E. M.; Vogel, C. Global signatures of protein and mRNA expression levels. *Mol. Biosyst* **2009**, *5*, 1512-1526.

(42) Greenbaum, D.; Colangelo, C.; Williams, K.; Gerstein, M. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* **2003**, *4*, 117.

(43) Griffin, T. J.; Gygi, S. P.; Ideker, T.; Rist, B.; Eng, J.; Hood, L.; Aebersold, R. Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. *Mol. Cell. Proteomics* **2002**, *1*, 323-333.

(44) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1*, 376-386.

(45) Cox, J.; Matic, I.; Hilger, M.; Nagaraj, N.; Selbach, M.; Olsen, J. V.; Mann, M. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **2009**, *4*, 698-705.

(46) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359-362.

(47) Rappsilber, J.; Ishihama, Y.; Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **2003**, *75*, 663-670.

(48) 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*, 1367-1372.

(49) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**.

(50) Dennis, G.,Jr; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **2003**, *4*, P3.

(51) Haase, H.; Alvarez, J.; Petzhold, D.; Doller, A.; Behlke, J.; Erdmann, J.; Hetzer, R.; Regitz-Zagrosek, V.; Vassort, G.; Morano, I. Ahnak is critical for cardiac Ca(V)1.2 calcium channel function and its beta-adrenergic regulation. *FASEB J.* **2005**, *19*, 1969-1977.

(52) Denton, R. M. Regulation of mitochondrial dehydrogenases by calcium ions. *Biochim. Biophys. Acta* **2009**, *1787*, 1309-1316.

(53) Anders, D. G.; McCue, L. A. The human cytomegalovirus genes and proteins required for DNA synthesis. *Intervirology* **1996**, *39*, 378-388.

(54) McMahon, T. P.; Anders, D. G. Interactions between human cytomegalovirus helicase-primase proteins. *Virus Res.* **2002**, *86*, 39-52.

(55) Pari, G. S. Nuts and bolts of human cytomegalovirus lytic DNA replication. *Curr. Top. Microbiol. Immunol.* **2008**, *325*, 153-166.