## CHARACTERIZATION OF AMINO ACID METABOLISM IN ESCHERICHIA COLI

by

**Brian Grant Phillips** 

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Degree in Major with Distinction

Spring 2017

© 2017 All Rights Reserved

#### **CHARACTERIZATION OF**

#### AMINO ACID METABOLISM

#### IN ESCHERICHIA COLI

by

**Brian Grant Phillips** 

Approved:

Maciek Antoniewicz, PhD Professor in charge of thesis on behalf of the Advisory Committee

Approved:

Wilfred Chen, PhD Committee member from the Department of Chemical and Biomolecular Engineering

Approved:

Millicent Sullivan, PhD Committee member from the Board of Senior Thesis Readers

Approved:

Michael Arnold, Ph.D. Director, University Honors Program

#### ACKNOWLEDGMENTS

I would like to extend my deepest gratitude towards Dr. Maciek Antoniewicz and members of the Antoniewicz group. I came into this knowing very little concerning the study of metabolic engineering but through their patience and Dr. Antoniewicz's excellence both as a teacher and a scholar I have learned far more than I could have ever hoped. Finally, I would like to thank my family and my friends for giving me invaluable support throughout this endeavor. "Because your own strength is unequal to the task, do not assume that it is beyond the powers of man; but if anything is within the powers and province of man, believe that it is within your own compass also."

## TABLE OF CONTENTS

LIST OF TA	ABLES			
	GUKES T	••••••	•••••	
ABSIRAC	1	•••••	•••••	
11	GOAI	S AND M	ίοτινατι	ON 1
1.1	OVER		FRESEAR	CH ACTIVITIES 2
2.1	DREV			MINO ACID BIOMASS PROFILES
2.1	DISCI		) F PREVIO	UIS STUDY OF AMINO ACID
2.2	BIOM	ASS PRO	FII FS	5 STODT OF ANIAG ACID
23	PRFI	IMINARY	Y STUDY (	OF AMINO ACID BIOMASS PROFILES 5
2.5	DISCI		E PRELIN	AINARY VERSUS PREVIOUS AMINO
2.4		BIOMAS	S DRUEIII	FS 6
3 1	SECO			ES AMINO ACID BIOMASS DDOEILES 7
5.1	SECO			AWINO ACID DIOWASS I KOFILES
	311	METHO	DS	7
	312	SECONI	DARY GR	OWTH RESULT 8
	313	PRODUC	T FORM	ATION RESULT FOR ALANINE 11
	5.1.5	TRODUC		
4.1	BACK	GROUNI	$O OF^{13}C-N$	/ ETABOLIC FLUX ANALYSIS
4.2	PRED	ICTED LA	ABELING	PATTERN FOR FULLY LABELED
	ALAN	JINE ASF	PARTATE	AND GLUTAMINE 13
43	HYPC	THESIS (	OF CENTR	AI METABOLISM FOR ALANINE
1.5	ASPA	RTATE &	& GLUTAI	MINE 15
51	METH	IOD OF P	REPARAT	TION AND MONITORING OF BIOMASS
5.1	GROV	NTH FOR	$^{13}C-MET$	ABOLIC FLUX ANALYSIS
52	RESU	TTS AND		SION 19
5.2	RESU		DISCOS	17
	5.2.1	METHO	D OF ANA	LYSIS FOR BIOMASS FROM
	0.2.1	PRELIM	INARY <sup>13</sup>	C-MFA EXPERIMENT 20
	522	GC-MS I	INTEGRA'	TION RESULT OF PRELIMINARY $^{13}C_{-}$
	0.2.2	MFA EX	PERIMEN	T (PREDICTED METABOLITE
		LABELI	NG)	21
	523	PRFI IM	INARY M	FTABOLIC FLUX PROFILES FOR
	5.2.5			ASP AND GLN 24
		SINULA		
		5231	METHOD	24
		5232	RESULTS	25
		5.2.5.2		
			5.2.3.2.1	<sup>13</sup> C-MFA OF SECONDARY ALA
				EXPERIMENT
	5.2.4	CONCLU	USION	

REFE	ERENCES	
А	APPENDIX TITLE-One	
	Appendix 1.1	
	Glycolysis	
	Pentose phosphate pathway	
	Entner–Doudoroff pathway	
	TCA Cycle	
	Glyoxylate Shunt	
	Amphibolic reactions	
	Acetic acid formation	
	Amino acid biosynthesis	
	One-carbon metabolism	
	Oxidative phosphorylation	
	Transhydrogenation	
	ATP hydrolysis	
	Transport	
	Biomass formation	
	CO2 exchange	
	Appendix 1.2	

## LIST OF TABLES

Table 1: $\mu^{max}  and  (-q_s$ ) values for Ala, As	p, and Gln10
Table 2: <sup>13</sup> C labeling of GC-MS amino ac	id fragments22

### LIST OF FIGURES

Figure 1: OD <sub>600</sub> of biomass grown under different amino acid conditions5
Figure 2: OD <sub>600</sub> of biomass grown in unlabeled, 20mM amino acid growth conditions
Figure 3: OD <sub>600</sub> measurements for Ala, Asp, and Gln cultures
Figure 4: $ln(C_x)$ as a function of time for Ala, Asp, and Gln cultures
Figure 5: Biomass concentration (C <sub>x</sub> ) and substrate concentration (C <sub>s</sub> ) for Ala, Asp, and Gln cultures
Figure 6: Product concentration ( $C_{ac}$ ) and substrate concentration ( $C_{ala}$ ) for Ala 12
Figure 7: Model of central carbon metabolism for <i>E. coli</i>
Figure 8: Role of amino acids in central carbon metabolism of <i>E. coli</i>
Figure 9: Growth profile of respective <sup>13</sup> C conditions for preliminary tracer experiment
Figure 10: Metabolic Flux Profile for First <sup>13</sup> C-MFA of <i>E. coli</i> grown on labeled Ala25
Figure 11: Metabolic Flux Profile for <sup>13</sup> C-MFA of <i>E. coli</i> grown on labeled Asp 26
Figure 12: Metabolic Flux Profile for <sup>13</sup> C-MFA of <i>E. coli</i> grown on labeled Gln 28
Figure 13: Metabolic Flux Profile for Second <sup>13</sup> C-MFA of <i>E. coli</i> grown on labeled Ala

#### ABSTRACT

The metabolism of alanine (Ala), aspartate (Asp), and glutamine (Gln) in *Escherichia coli* was studied using <sup>13</sup>C-metabolic flux analysis (<sup>13</sup>C-MFA) to assist future efforts to engineer methylotrophy (i.e. methanol utilization) for this organism. Growth analysis were performed using Ala, Asp, and Gln as substrates which revealed that *E. coli* has biomass specific substrate uptake rates [mmol/g DW biomass/hr] of 11.6, 5.93, and 1.83, respectively, for the three respective substrates. The biomass specific growth rates [1/hr] for Ala, Asp, and Gln culture were 0.222, 0.138, and 0.033, respectively. Parallel <sup>13</sup>C tracer experiments were also performed using <sup>13</sup>C-Ala, <sup>13</sup>C-Asp, and <sup>13</sup>C-Gln tracers and the labeling data was analyzed using <sup>13</sup>C-MFA. The resulting flux results revealed that amino acid metabolism of Ala, Asp, and Gln in *E. coli* is characterized by high fluxes through the TCA cycle in central carbon metabolism with much lower fluxes through other metabolic pathways.

#### Chapter 1

#### INTRODUCTION

#### 1.1 GOALS AND MOTIVATION

In the 21<sup>st</sup> century, the biochemical conversion of abundant feedstocks to highvalue chemicals is increasingly desired. Of these feedstocks, methanol sourced from methane is an attractive substrate for fermentation due to its high degree of reduction (i.e. energy content) Methylotrophs, or organisms which preliminarily uptake methanol as a substrate for growth, are currently not employed in industry due to the lack of requisite genetic tools for such an endeavor.<sup>1</sup> Therefore, there is significant interest in employing more widely used organisms such as *Escherichia coli* for this task, which is the organism we focused on in this work.

Previous work performed by Whitaker et al.<sup>2</sup> demonstrated that *E. coli* can be successfully engineered to take up methanol as a co-substrate for cell growth. More recently, work within the Antoniewicz group demonstrated that *E. coli* engineered for methylotrophy, when grown in a culture containing a co-substrate of yeast extract,

<sup>&</sup>lt;sup>1</sup> Gonzalez, J., Bennet R., Papoutsakis, E., Antoniewicz, M., 2017

<sup>&</sup>lt;sup>2</sup> Whitaker, W.B., Jones, J.A., Bennett, K., Gonzalez, J., Vernacchio, V.R., Collins, S.M., Palmer, M. a., Schmidt, S., Antoniewicz, M.R., Koffas, M. a., Papoutsakis, E.T., **2016.** Engineering the Biological Conversion of Methanol to Specialty Chemicals in Escherichia coli. Metab. Eng. 1–11.

exhibits significantly improved biomass growth. Exactly why this effect is observed is still unclear but amino acids are the primary component of yeast extract and thus their effect on cellular metabolism should be analyzed to rationalize the effect of yeast extract on biomass growth. Preliminary studies were conducted in the Antoniewicz lab to investigate the effect of the presence of particular amino acids within the extract as it pertains to cellular metabolism.<sup>1</sup>

To supplement these research efforts, this work investigated a class of amino acids which exhibited particularly improved biomass growth in co-culture with methanol. The three amino acids analyzed in this work are alanine (Ala), aspartate (Asp), and glutamine (Gln). By investigating the metabolism of these amino acids this work seeks to provide a solid foundation which can further improve efforts to improve biomass growth on methanol as a co-substrate.

#### **1.2 OVERVIEW OF RESEARCH ACTIVITIES**

For the purpose of determining metabolic fluxes in the presence of Ala, Asp, and Gln as substrates, the following research was conducted. First, a preliminary study on the effect of amino acid presence on biomass growth was replicated with respect to prior efforts<sup>1</sup> to ensure that the effect of each amino acid on biomass cell growth was reproducible. This experiment also explored if there was a significant difference in growth rate for each amino acid if the concentration of the substrate was increased from 5mM to 20mM. Subsequently, a more detailed study of cell growth was conducted to obtain biomass specific uptake rates and growth rates for each amino acid condition. This information was used to calculate the biomass yield on each substrate. Measured metabolic rates were also used in subsequent <sup>13</sup>C-flux analysis study.

2

Finally, parallel <sup>13</sup>C tracer experiments were conducted to obtain information necessary for the estimation of metabolic flux profiles of *E. coli* grown on each amino acid individually. This analysis was performed using the Metran software (developed previously by Prof. Antoniewicz) and was conducted with respect to the currently accepted model of the metabolic network of *E. coli*.

#### Chapter 2

#### 2.1 PREVIOUS STUDY OF AMINO ACID BIOMASS PROFILES

Previous work in the Antoniewicz group<sup>1</sup> determined the effect of various amino acids on cell growth of *E. coli*. In particular, *E. coli* was first pre-cultured using a typical growth substrate (i.e. a sugar) and then grown in the presence of amino acid alone. Specifically, the inoculum was grown in M9 minimal medium first in glucose until stationary phase; this was determined by following biomass concentrations as a function of time. Growth conditions were at 37° C in 125 mL Erlenmeyer flasks. Subsequently, 1 mL inoculum from the glucose culture were grown in M9 minimal media containing a particular amino acid for cell growth.

For amino acid condition cells were cultured again at 37° C in 125 mL Erlenmeyer flasks at 5 mM working concentrations for each particular amino acid. To investigate cell growth as a function of time, 200  $\mu$ L samples were withdrawn at regular time intervals and biomass concentration was determined by measuring optical density at 600 nm wavelength (OD<sub>600</sub>) using a spectrophotometer (Eppendorf BioPhotometer). This resulted in the growth curves shown in Figure 1.



Figure 1: OD<sub>600</sub> of biomass grown under different amino acid conditions.

#### 2.2 DISCUSSION OF PREVIOUS STUDY OF AMINO ACID BIOMASS PROFILES

Figure 1 illustrates that *E. coli* was able to grow in the presence of alanine (Ala), aspartate (Asp), and glutamine (Gln). Interestingly, although known degradation pathways exist for serine (Ser) and thrionine (Thr) for *E. coli* cultures did not seem to grow in the presence of these amino acids as substrates. Why this was observed remains unclear; this experiment was only conducted once and therefore we should not rule out the influence of repeatability on this observation. There could also be a rational reason, however, for this occurrence which might of be of interest for future investigations. Whatever the conclusion may be, Figure 1 illustrates that *E. coli* appears to achieve stationary phase first in Ala culture (i.e. fastest growth on this substrate), while Asp and Gln cultures reached stationary phase subsequently afterwards.

#### 2.3 PRELIMINARY STUDY OF AMINO ACID BIOMASS PROFILES

Given the previous study of amino acid biomass profiles, to investigate the effect of amino acid influence on biomass cell growth the experiment described in Section 2.1 was repeated in but this time using 20mM working concentrations of Ala, Asp, and Gln. By increasing the concentration of each amino acid fourfold, it was hypothesized that even greater differences between biomass growth profiles would be observed for each amino acid condition. The results of these growth experiments are shown in Figure 2.



Figure 2: OD<sub>600</sub> of biomass grown in unlabeled, 20mM amino acid growth conditions.

#### 2.4 DISCUSSION OF PRELIMINARY VERSUS PREVIOUS AMINO ACID BIOMASS PROFILES

From Figure 2 it is apparent that increasing the concentration of each amino acid fourfold successfully increased the relative differences in biomass cell growth for each substrate condition. Furthermore, in comparison to the previous study of biomass cell growth, the observed relative time to stationary phase was again noticed. *E. coli* first achieved stationary phase for Ala culture, and achieved stationary phase at later times in time for the Asp and Gln cultures. Interestingly, despite taking more time to reach stationary phase, the results from Asp culture would suggest that more biomass was produced under this condition versus the other conditions. To more precisely qualify these observations, however, more precise, quantitative information would be needed. Thus, a more detailed subsequent investigation of biomass cell growth was conducted.

#### Chapter 3

#### 3.1 SECONDARY STUDY OF AMINO ACID BIOMASS PROFILES

#### 3.1.1 METHODS

For this second set of growth experiments, cultures were first grown at 37°C in 125mL Erlenmeyer flasks in M9 minimal media with 20mM glucose serving as the carbon source for growth. When the pre-culture reached stationary phase, evidenced by constant  $OD_{600}$  measurements, 100 µL inoculum was withdrawn and transferred to new cultures containing the respective amino acids as the only substrate. Thus, the amino acids studied (again) were, alanine, aspartate, and glutamine. For each culture the amino acid was present at 20mM working concentration in M9 minimal medium. To monitor the biomass concentration as function of time within each condition the same procedure was followed as outlined in section 2.1

For this experiment, in addition to measuring increase in biomass concentration, the substrate concentration as a function of time was also monitored. For each culture, 150  $\mu$ L samples were withdrawn from culture and subsequently mixed with a <sup>13</sup>C-labeled standard. Mixtures were then subjected to GC-MS analysis to determine the ratio of unlabeled amino acid (M0) to fully labeled amino acid (Mn) from which the concentration of each amino acid could be calculated by the following equation:

$$C_{sample,i} = \frac{MO}{M3} \cdot \frac{V_{std}}{V_{sample}} \cdot C_{std} \qquad (1)$$

Finally, for all cultures the samples were also analyzed on HPLC to determine acetate concentration as function of time (HPLC analysis was performed by Jackie

Gonzalez.). More specifically, to determine the concentration of acetate (Ac) in the samples, 100  $\mu$ L samples were withdrawn at various time points throughout culture and subsequently centrifuged at 14,000 rpm for 5 minutes. Supernatant from these extracts was subsequently analyzed for concentration via HPLC calibrated to a known internal standard of acetate. By determining both product concentration and substrate concentration as a function of time, in addition to biomass concentration as a function of time, in addition to biomass yields and product yields.

#### 3.1.2 SECONDARY GROWTH RESULT



Results of cell growth analysis of E. coli on each amino acid is shown in

Figure 3: OD<sub>600</sub> measurements for Ala, Asp, and Gln cultures.



Figure 4:  $ln(C_x)$  as a function of time for Ala, Asp, and Gln cultures.

Determination of growth rate for each substrate is shown in Figure 4. Here,  $C_x$  denotes the concentration of biomass in units of mCmol/L.  $C_x$  was obtained by assuming that 1 unit of OD<sub>600</sub> corresponds to 0.32 g of dry weight of biomass per liter, and assuming 25g/Cmol for dry weight biomass. Cell growth rates were determined for each substrate from three time point measurements. Figure 4 illustrates that Ala produced the highest growth rate, followed by Asp, and Gln, i.e. growth rates of 0.222, 0.138, and 0.033 (1/h), respectively.

Next, the biomass yield for each amino acid substrate was determined by plotting the substrate concentration as a function of biomass concentration, as shown in Figure 5.



Figure 5: Biomass concentration ( $C_x$ ) and substrate concentration ( $C_s$ ) for Ala, Asp, and Gln cultures.

By multiplying the slope of each line in Figure 5 by the previously determined growth rate for each condition, and assuming that dry biomass has a molecular weight of 25 g/ Cmol, the biomass specific substrate uptake rate ( $-q_s$ ) for each amino acid was obtained (Table 1).

	Ala	Asp	Gln
$\mu^{max}$ [1/h]	0.222	0.138	0.033
-qs [mmol/g DW biomass/hr]	11.6	5.93	1.83

Table 1:  $\mu^{max}$  and (–q<sub>s</sub>) values for Ala, Asp, and Gln

Interestingly, our results illustrate that the measured  $Y_{sx}$  values are in disagreement with theoretical  $Y_{sx}^{max}$  values for these amino acid substrates. That is, if we calculate the degree of reduction for Ala (C3H7NO2), Asp (C4H7NO4), and Gln (C5H10N2O3), we obtain degrees of reduction of 4, 3, and 3.6, respectively. Generally, higher biomass yield is obtained for substrates with higher degrees of reduction (i.e. more energy content). Thus, we would expect that the  $Y_{sx}^{max}$ [mCmol/mCmol] values to be in descending order as Ala, Gln, and then Asp.

Because  $Y_{sx}$  is equal to  $\mu^{max}$  divided by  $-q_s$  for each amino acid, if the numbers are again adjusted for proper units,  $Y_{sx}$  [mCmol/mCmol] is calculated as 0.020, 0.016, and 0.005 for Ala, Asp, and Gln, respectively. Thus, the experimental results seem to contradict the prediction of the magnitude of  $Y_{sx}$ . Why this contradiction exists is unclear and might be of interest in future work.

#### 3.1.3 PRODUCT FORMATION RESULT FOR ALANINE

Out of the three amino acid substrates studied here, acetate was produced only when Ala was used as the substrate, i.e. no acetate was detected in Asp and Gln cultures. To determine the acetate yield, acetate concentration was plotted as a function of alanine concentration (Figure 6.)



Figure 6: Product concentration (Cac) and substrate concentration (Cala) for Ala

In constructing the figure above, two out of three data points were used because it was determined that *E. coli* most likely started using Ac as a carbon source for cell growth during the second half of the growth phase. Nonetheless, the slope of the line illustrated in Figure 6 directly informs us about the yield,  $Y_{sp}$  of acetate on Ala which is 0.394 [mol/mol]. This measurement was also employed in <sup>13</sup>C metabolic flux analysis, described next.

#### **Chapter 4**

#### 4.1 BACKGROUND OF <sup>13</sup>C-METABOLIC FLUX ANALYSIS

<sup>13</sup>C-metabolic flux analysis (<sup>13</sup>C-MFA) is a technique for determining the metabolic flux profiles associated with a particular organism as it grows on a given substrate of study. It relies upon <sup>13</sup>C labeled substrates; when <sup>13</sup>C is assimilated into an organism's metabolites it provides information regarding the relative fluxes through each pathway of an organism's central metabolism, *ex post facto*. Exactly how <sup>13</sup>C is assimilated into a respective organisms metabolites is determined typically through GC-MS or NMR experiments. These experiments provide information regarding the isotopomer distribution of <sup>13</sup>C atoms which have been assimilated into various metabolites such as amino acids. By knowing distribution of isotopomers, and also by knowing the stoichiometry and carbon rearrangements associated with biochemical reactions within a metabolic network we can generate a simulated flux profile which describes metabolic flux associated with a given growth condition for an organism of study. For the purpose of simulation, the most efficient method to date is the elementary metabolite unit (EMU) framework which has been well described and employed by Antoniewicz et al.<sup>3</sup>

#### 4.2 PREDICTED LABELING PATTERN FOR FULLY LABELED ALANINE, ASPARTATE, AND GLUTAMINE

Before <sup>13</sup>C-MFA was conducted, models illustrating the incorporation of <sup>13</sup>C into the metabolites associated with *E. coli* central metabolism were developed for

<sup>&</sup>lt;sup>3</sup> M.R. Antoniewicz, J.K. Kelleher, G. Stephanopoulos. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. Metab. Eng., 9 (2007), pp. 68–86

each amino acid, assuming that they were labeled at different carbon positions to begin with. These models were used to predict the behavior of metabolic flux profiles which would be simulated for a given <sup>13</sup>C labeled amino acid.

For this work, reactions listed within Appendix 1.1 were used to generate models for alanine (Ala), aspartate (Asp), and glutamine (Gln) which depict the fate of <sup>13</sup>C originating from a respective carbon position associated with the particular amino acid. Within Appendix 1.1 it is important to note that lettered carbons rearrange throughout the course of different reactions. For instance, in the case of glutamate generation from AKG (v<sub>36</sub>) the fact that (abcde) appears for both Glu and AKG reflects that <sup>13</sup>C sourced from AKG would end up in identical carbon positions in Glu. This is not always the case, however, and thus for accurate simulation it is important that we keep track of carbon position rearrangements. In that context, these efforts produced the models depicted in Appendix 1.2.

Comparing between Ala, Asp, and Gln in Appendix 1.2 it becomes clear that a metabolic network model simulated for Gln should be expected to be more difficult to observe in practice. As illustrated in Appendix 1.2, there are more possible labeling states for metabolites within the model as opposed to Ala and Asp. The net effect of this probably explains why it was difficult to obtain a satisfactory result around the part of metabolism involving PEP and Pyr for glutamine as the substrate (see next sections.)

14

#### 4.3 HYPOTHESIS OF CENTRAL METABOLISM FOR ALANINE, ASPARTATE, & GLUTAMINE

With respect to *E. coli* metabolism the following metabolic network model was employed for simulations. It was developed previously by the Antoniewicz group based on *E. coli* model in the KEGG database<sup>4</sup> and from previous experimental efforts.<sup>5</sup>



<sup>&</sup>lt;sup>4</sup> The KEGG (Kyoto Encyclopedia of Genes and Genomes) database integrates genomic, chemical, and systemic functional information for organisms of study. *E. coli* being a very commonly employed organism in academia has a richly understood metabolic network model as a result of decades of effort.

<sup>&</sup>lt;sup>5</sup> M.R. Antoniewicz, D.F. Kraynie, L.A. Laffend, J. Gonzalez-Lergier, J.K. Kelleher, G. Stephanopoulos, Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol, Metab. Eng., 9 (2007), pp. 277–292

Figure 7: Model of central carbon metabolism for E. coli

At a glance, the main components of *E. coli* central metabolism center on three main pathways. The central part of metabolism stemming downward from the breakdown of glucose is the glycolysis pathway. Immediately to the right of the glycolysis pathway is the oxidative pentose phosphate pathway. And finally beneath the endpoint of the glycolytic pathway, there is the TCA cycle. Biochemical reactions implicit to this model were adapted from Leighty and Antoniewicz<sup>6</sup> (See Appendix 1.1)

With respect to central carbon metabolism for in this work we predicted the degradation and integration of carbon associated with each amino acid in the context of the figure shown below.<sup>7</sup>

<sup>&</sup>lt;sup>6</sup> R.W. Leighty, M.R. Antoniewicz, Parallel labeling experiments with [U-13C] glucose validate E. coli metabolic network model for 13C metabolic flux analysis, Metab. Eng., 14 (2012), pp. 533–541

<sup>&</sup>lt;sup>7</sup> http://www.uky.edu/~dhild/biochem/24/lect24.html



Figure 8: Role of amino acids in central carbon metabolism of E. coli.

First, for the case of Ala it was hypothesized that it could enter the central carbon metabolism via the reversal of the reaction which generates alanine from pyruvate. This reaction is typically depicted as:

$$Pyr + Glu \rightarrow Ala + AKG.$$
(2)

For the case of Asp, the exact reaction with which it is integrated into *E. coli's* central metabolism was less clear because amino acid biosynthesis reactions within the

network model employ Asp several times. (See Appendix 1.1) Despite this consideration, it was hypothesized that Asp when acting as a sole carbon source for growth would be assimilated into the network via the generation of oxaloacetate (OAC) thus entering the TCA cycle. Finally, with respect to Gln it was hypothesized that it would enter central metabolism through the generation of  $\alpha$ -ketoglutarate (AKG) and thus also enter the TCA cycle.

#### Chapter 5

#### 5.1 METHOD OF PREPARATION AND MONITORING OF BIOMASS GROWTH FOR <sup>13</sup>C-METABOLIC FLUX ANALYSIS

For the purpose of the <sup>13</sup>C-metabolic flux analysis (<sup>13</sup>C-MFA) cultures were first grown at 37°C in 125mL Erlenmeyer flasks in M9 minimal medium with 20mM glucose serving as the carbon source. When cultures reached stationary phase, as evidenced by constant OD<sub>600</sub> measurements, 100  $\mu$ L inoculum was withdrawn and transferred to 9 different cultures. These cultures contained the following <sup>13</sup>C-tracers: were as follows: [1-Ala], [2-Ala], [3-Ala] for alanine cultures, [1,4-Asp], [2-Asp], [4-Asp] for aspartate cultures, and [1-Gln], [2-Gln], [5-Gln] for glutamine cultures. The notation indicates which carbon positions were labeled with <sup>13</sup>C atoms, e.g. [1,4-Asp] is labeled at the first and fourth carbon positions. For each culture the isotopic purity of the <sup>13</sup>C labeled substrate was >99% (Sigma-Aldrich), and substrates were prepared at 20mM working concentration in M9 minimal media. To monitor cell growth as a function of time for each condition the same procedure was followed as outlined in section 2.1

#### 5.2 **RESULTS AND DISCUSSION**

The measured growth curves for the labeling experiments are shown below in Figure 9.



Figure 9: Growth profile of respective <sup>13</sup>C conditions for preliminary tracer experiment.

### 5.2.1 METHOD OF ANALYSIS FOR BIOMASS FROM PRELIMINARY <sup>13</sup>C-MFA EXPERIMENT

In order to analyze the <sup>13</sup>C labeling of cellular metabolites for each substrate and tracer, 1mL of culture was extracted as soon as it was deemed to have reached exponential growth phase (OD>0.5) It was important to analyze metabolites associated with cells growing in this growth phase because samples from this stage of growth reflect cells at their most active metabolic state and thus most relevant metabolic conditions for cell growth. Samples were subsequently centrifuged at 14,000 rpm for 5

minutes (Beckmann Coulter) and supernatant from each sample was removed and discarded (i.e. only the cell pellet was used for <sup>13</sup>C-labeling analysis)

For each sample, biomass pellets resulting from the centrifugation treatment were subsequently hydrolyzed overnight with 0.5 mL of 6M HCl at 110°C. Following hydrolysis, samples were evaporated at 65°C for a period of 2 hours to prepare them for TBDMS derivatization and subsequent GC-MS analysis. 50  $\mu$ L of pyridine (Sigma-Aldrich) and 100  $\mu$ L of TBDMS (Sigma-Aldrich) were subsequently added to samples, which were then subjected to GC-MS analysis. For GC-MS analysis an Agilent 7890B GC system equipped with a DB-5MS capillary column was employed (30 m, 0.25 mm i.d., 0.25  $\mu$ m-phase thickness; Agilent J&W Scientific) The GC was connected to an Agilent 5977A Mass Spectrometer operating under ionization by electron impact (EI) at 70 eV. Mass isotopomer distributions were obtained by integration<sup>8</sup> and corrected for natural isotope abundances, as is commonly done.<sup>9</sup>

#### 5.2.2 GC-MS INTEGRATION RESULT OF PRELIMINARY <sup>13</sup>C-MFA EXPERIMENT (PREDICTED METABOLITE LABELING)

For the purpose of obtaining mass isotopomer distributions pertaining to amino acids of GC-MS analyzed biomass, integration results for a given mass-to-charge ratio (m/z) of an amino acid were analyzed to assess the carbon labeling associated with a

<sup>&</sup>lt;sup>8</sup> Antoniewicz, M. R., Kelleher, J. K., Stephanopoulos, G., 2007. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. Anal Chem. 79, 7554-9.

<sup>&</sup>lt;sup>9</sup> Fernandez, C. A., Des Rosiers, C., Previs, S. F., David, F., Brunengraber, H., 1996. Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom. 31, 255-62.

particular tracer. As an example, in the case of Ala-232 and Ala-233, if Ala-233 is significantly more present than Ala-232 after integration of GC-MS data, then we can say that it has effectively been labeled by one <sup>13</sup>C-atom. For other amino acid fragments, this number varied between 0 and 3. Higher numbers observed were often associated with multiple precursor metabolites for a given metabolite fragment. Analyzing GC-MS fragments for each condition produced the data shown in Table 2.

										Precursor Metabolites
-	1-Ala	2-Ala	3-Ala	1,4-Asp	2-Asp	4-Asp	1-Gln	2-Gln	5-Gln	
Ala232	0	1	1	0	1	0	0	0	0	Pyr (2-3)
Ala260	1	1	1	1	1	0	0	0	0	Pyr (1-2-3)
Gly218	0	0	0	0	0	0	0	0	0	
Gly246	1	1	0	1	1	0	0	0	0	
Val260	0	2	2	0	2	0	0	0	0	Pyr (2-3) + Pyr (2-3)
Val288	1	2	2	1	2	0	0	0	0	Pyr (1-2-3) + Pyr (2-3)
Leu274	0	2	3	0	3	0	0	0	0	AcCoA(2) + Pyr (2-3) + Pyr (2-3)
Ile200	0	2	3	1	2	0	0	0	0	OAC (2-3-4)+Pyr (2-3)
Ile274	0	2	3	1	2	1	0	0	0	OAC (2-3-4)+Pyr (2-3)
Met218	0	1	3	1	1	1	0	0	0	OAC (2-3-4) + C-1
Met320	0	2	3	2	2	1	0	1	1	OAC (1-2-3-4) + C-1
Ser362	0	1	1	0	1	0	0	0	0	
Ser390	1	1	1	1	1	0	0	0	0	
Thr376	0	1	2	1	1	1	0	0	0	OAC (2-3-4)
Thr404	0	1	3	2	1	1	0	1	1	OAC (1-2-3-4)
Phe302	1	1	1	1	0	0	0	0	0	PEP (1-2)
Phe308	2	3	3	2	3	1	1	1	1	PEP(2-3) + PEP (2-3) + E4P (1-2-3-4)
Phe336	3	3	3	2	3	1	0	1	1	PEP (1-2-3) + PEP(2-3) + E4P (1-2-3-4)
Asp302	1	1	1	1	1	0	0	0	0	OAC (1-2)
Asp390	0	1	2	1	1	1	0	0	0	OAC (2-3-4)
Asp418	0	2	3	2	1	1	0	1	1	OAC (1-2-3-4)
Glu330	0	2	3	0	2	0	0	1	1	AKG (2-3-4-5)
Glu432	0	2	3	1	2	1	1	1	1	AKG (1-2-3-4-5)
Tyr302	1	1	0	1	0	0	0	0	0	PEP (1-2)

Table 2: <sup>13</sup>C labeling of GC-MS amino acid fragments.

Precursor metabolites listed adjacent to a particular fragment indicate the

carbon positions within the given metabolite which produced the given labeling

associated with a particular fragment. This type of information is important because it

informs us of pathways within central metabolism that are predicted to be active. It also in our case corroborates predictions made in section 4.3 regarding the point of entry into central metabolism for each amino acid.

From this information the following information could be obtained. First, with respect to labeled Ala from Val288 fragment, it can be predicted that there is significant metabolic flux near the pyruvate node within the model of *E. coli* central metabolism depicted in section 3.1.2. This is in good agreement with our hypothesis concerning the entry of Ala into central metabolism. With respect to Asp, labeling at Phe336 suggests that we would expect significant metabolic flux near the PEP node (i.e. this would also be predicted for Ala as well.) The PEP node is not far removed from the OAC node where Asp is proposed to enter metabolism; thus, this observation also supports our hypothesis. Finally, for Gln, the <sup>13</sup>C labeling suggests that there would be less flux through the pyruvate node but that PEP and AKG nodes would be more active; this also corroborates the prediction made for Gln.

## 5.2.3 PRELIMINARY METABOLIC FLUX PROFILES FOR SIMULATED ALA, ASP, AND GLN

#### 5.2.3.1 METHOD

<sup>13</sup>C-MFA using labeling data from the 9 tracer experiments was performed using the Metran software.<sup>10</sup> This software is developed from the elementary metabolite unit (EMU) framework mentioned in section 4.1 Fluxes for the E. coli metabolic network model were estimated for each condition by minimizing the variance-weighted sum of squared residuals (SSR) between experimentally measured and model predicted mass isotopomer distributions of amino acids. For each amino acid, results from each labeling condition were fitted in parallel and this was achieved through non-linear least squares regression using the Metran software. Flux estimation was repeated at a minimum of 10 times, starting with random initial values for all fluxes to find a global solution for a particular case. Fluxes were scaled to the uptake rate of the particular amino acid being considered; for example in the case of Ala, fluxes were scaled to Ala (where this represents Ala which is utilized from the medium.) For all cases, the scaling flux was set to a value of 100 for the amino acid of interest. Once a satisfactory SSR was obtained for a given simulation, 95% confidence intervals were computed for simulated fluxes by evaluating the sensitivity of minimized SSR to variations in flux in Metran.<sup>11</sup>

<sup>&</sup>lt;sup>10</sup> H. Yoo, M.R. Antoniewicz, G. Stephanopoulos, J.K. Kelleher, Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. J. Biol. Chem., 283 (2008), pp. 20621–20627

<sup>&</sup>lt;sup>11</sup> M.R. Antoniewicz, J.K. Kelleher, G. Stephanopoulos. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab. Eng., 8 (2006), pp. 324–337

#### 5.2.3.2 RESULTS

For the case of Ala, <sup>13</sup>C-MFA produced the metabolic flux profile shown in Figure 10.



Figure 10: Metabolic Flux Profile for First <sup>13</sup>C-MFA of *E. coli* grown on labeled Ala.

This result demonstrates that, *E. coli* metabolism of Ala features minimal flux associated with the oxidative pentose phosphate pathway (i.e. reactions to the right of

the beginning steps of glycolysis.) Importantly, this result illustrates significant metabolic flux at the entry point of the TCA cycle, specifically with respect to the generation of AcCoA and Ac. Thus, the fact that a statistically satisfactory fit was obtained supports our hypothesis that Ala enters metabolism through pyruvate.

For the case of Asp, <sup>13</sup>C-MFA produced the metabolic flux profile shown in Figure 11.



Figure 11: Metabolic Flux Profile for <sup>13</sup>C-MFA of *E. coli* grown on labeled Asp.

From this flux result for Asp, similarities are observed with respect to what was observed for the case of Ala. For instance, the flux model illustrates that minimal metabolic flux is associated with the oxidative pentose phosphate pathway. Also, possible flux for gluconeogenesis is indicated to be small relative to the flux associated with the TCA cycle and the PEP and OAC nodes. There is significant metabolic flux associated with the OAC node; particularly, flux associated with the reaction to generate pyruvate from malate associated with the TCA cycle is significant.



For the case of Gln, <sup>13</sup>C-MFA produced Figure 11.

Figure 12: Metabolic Flux Profile for <sup>13</sup>C-MFA of *E. coli* grown on labeled Gln.

From the results for Gln, we observe again that metabolic flux is predicted to be minimal through the oxidative pentose phosphate pathway and relatively insignificant in the final steps of gluconeogenesis. For the case of Gln, a satisfactory confidence interval for the PEP and Pyr equilibrium reaction was not obtained. Why this might occur is explained in section 4.2. Overall, the flus result illustrates that the metabolic flux profile for Gln is most qualitatively similar to Ala in that the Mal to Pyr pathway is predicted to have negligible flux and that the most active node centers about Pyr.

The results for Ala, Asp, and Gln are especially interesting when scrutinized with respect to the following equation for biomass synthesis<sup>6</sup> which is implicit in the model used for simulation:

0.488 Ala+0.281 Arg+0.229 Asn+0.229 Asp+0.087 Cys+0.250 Glu+0.250Gln+0.582 Gly+0.090 His+0.276 Ile+0.428 Leu+0.326 Lys+0.146 Met+0.176 Phe+0.210 Pro+0.205 Ser+0.241 Thr+0.054 Trp+0.131 Tyr+0.402 Val+0.205 G6P+0.071 F6P+0.754 R5P+0.129 GAP+0.619 3PG+0.051 PEP+0.083 Pyr+2.510 AcCoA+0.087 AKG+0.340 OAC+0.443 MEETHF+33.247 ATP+5.363 NADPH $\rightarrow$ 39.68 Biomass+1.455 NADH (3)

The equation above provides a framework (from energy and carbon balances) with which biomass is formulated from key amino acids and key metabolites. Key metabolites relevant to this model appear after Val in the above equation. It is interesting to note then that from the metabolic flux profiles presented, for all three amino acids metabolic flux is predicted to be much smaller in steps involving the key metabolites of G6P, F6P, R5P, GAP, and 3PG. In other words, these flux results suggest that for biomass synthesis from amino acid conditions there is a relative shift away from the formation of these key metabolites necessary for biomass synthesis.

Intuitively however, this might be expected in the context of the metabolic network model because these metabolites are more typically generated from glycolysis

of carbon sources such as glucose. Thus, if we grow *E. coli* in an environment where the routes of synthesis for these metabolites are not available, then we should expect that flux would diminish through a given key metabolite as it is increasingly far removed from the proposed point of entry for the alternative carbon source (amino acids in this study.) For instance, with the exception of Ala, it is proposed that these amino acids enter central metabolism through the TCA cycle; F6P, however, is usually encountered as a metabolite not far removed from glucose due to its entry point in glycolysis. Thus, the observation that flux is minimal through the F6P node for Asp and Gln in their models is intuitive. While more experimental verification would certainly be necessary to qualify this observation, this consideration illustrates that equation (3) only offers a static picture regarding the synthesis of biomass. Including results from metabolic flux analysis is what allows a better understanding of metabolism to emerge.

#### 5.2.3.2.1 <sup>13</sup>C-MFA OF SECONDARY ALA EXPERIMENT

For the purpose of obtaining the metabolic network model pertinent to Ala, the procedure outlined in section 5.2.3.1 was again followed with the only difference being that this time the scaling flux for simulation was chosen as Ac production. From the acetate yield measurement obtained in Section 3.1.3, the relative secretion rate of acetate number was 40 (i.e. relative to 100 moles of Ala taken up.) Once, this information was entered and the previously outlined procedure was followed, the following flux result shown in Figure 13 was obtained.



Figure 13: Metabolic Flux Profile for Second <sup>13</sup>C-MFA of *E. coli* grown on labeled Ala.

Comparing this result to the previous result for Ala illustrates that qualitatively, the metabolic flux profile is about the same as the one obtained for the parallel tracer experiment without input of acetate secretion flux. For instance, minimal flux is predicted for the oxidative pentose phosphate pathway, while significant flux is predicted through the Pyr and AcCoA nodes. Note that the flux to produce Ac is very close to 40, because of measured flux employed for this analysis.

#### 5.2.4 CONCLUSION

This work aimed to develop a better understanding of E. coli metabolism of Ala, Asp, and Gln. In cell culture experiments it was observed that E. coli had the highest growth rate for Ala, then Asp, and that slowest growth was observed for Gln. This observation might have important consequences with respect to industrial fermentation of E. coli on amino acid substrates and also in future experimental efforts to improve *E. coli* methylotrophy, as indicated by previous efforts.<sup>1</sup> Preliminary metabolic flux profiles rendered for Ala, Asp, and Gln generally illustrated that for E. *coli* grown on these substrates metabolic flux is predicted to be minimal for metabolic pathways further removed from the TCA cycle, such as the oxidative pentose phosphate pathway and for gluconeogenesis. Significant metabolic flux, however, is predicted for central metabolism around the TCA cycle and thus indicates that it should be an area of focus for future metabolic engineering efforts. Sufficient agreement between model predictions rendered via parallel <sup>13</sup>C tracer experiments qualify the fidelity of the <sup>13</sup>C<sup>-</sup>MFA results as provided by Metran. Thus overall metabolic flux models of central metabolism for E. coli were successfully determined for Ala, Asp, and Gln and will assist in future experimental and modeling efforts.

#### REFERENCES

- 1.) Gonzalez, J., Bennet R., Papoutsakis, E., Antoniewicz, M., 2017
- Whitaker, W.B., Jones, J.A., Bennett, K., Gonzalez, J., Vernacchio, V.R., Collins, S.M., Palmer, M. a., Schmidt, S., Antoniewicz, M.R., Koffas, M. a., Papoutsakis, E.T., **2016.** Engineering the Biological Conversion of Methanol to Specialty Chemicals in Escherichia coli. Metab. Eng. 1–11.
- 3.) M.R. Antoniewicz, J.K. Kelleher, G. Stephanopoulos. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. Metab. Eng., 9 (2007), pp. 68–86
- M.R. Antoniewicz, D.F. Kraynie, L.A. Laffend, J. Gonzalez-Lergier, J.K. Kelleher, G. Stephanopoulos, Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of E. coli producing 1,3-propanediol, Metab. Eng., 9 (2007), pp. 277–292
- 5.) R.W. Leighty, M.R. Antoniewicz, Parallel labeling experiments with [U-13C]glucose validate E. coli metabolic network model for 13C metabolic flux analysis, Metab. Eng., 14 (2012), pp. 533–541
- 6.) http://www.uky.edu/~dhild/biochem/24/lect24.html
- Antoniewicz, M. R., Kelleher, J. K., Stephanopoulos, G., 2007. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. Anal Chem. 79, 7554-9.
- Fernandez, C. A., Des Rosiers, C., Previs, S. F., David, F., Brunengraber, H., 1996. Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom. 31, 255-62.
- 9.) H. Yoo, M.R. Antoniewicz, G. Stephanopoulos, J.K. Kelleher, Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. J. Biol. Chem., 283 (2008), pp. 20621–20627
- M.R. Antoniewicz, J.K. Kelleher, G. Stephanopoulos. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab. Eng., 8 (2006), pp. 324–337
- 11.) S.B. Crown, C.P. Long, M.R. Antoniewicz. Optimal tracers for parallel labeling experiments and 13C metabolic flux analysis: a new precision and synergy scoring system. Metab Eng, 38 (2016), pp. 10–18

## Appendix-One

## **APPENDIX TITLE-One**

## Appendix 1.1

## <u>Metabolic Network Model for E. coli (adopted from Leighty and Antoniewicz<sup>6</sup>):</u>

	Glycolysis
<b>V</b> 1	Gluc.ext (abcdef)+ATP $\rightarrow$ G6P (abcdef)
V2	G6P (abcdef)↔F6P (abcdef)
V <sub>3</sub>	F6P (abcdef)+ATP→FBP (abcdef)
<b>V</b> 4	FBP (abcdef)↔DHAP (cba)+GAP (def)
<b>V</b> 5	DHAP (abc)↔GAP (abc)
<b>V</b> <sub>6</sub>	GAP (abc)↔3PG (abc)+ATP+NADH
V7	3PG (abc)↔PEP (abc)
V <sub>8</sub>	PEP (abc)→Pyr (abc)+ATP
	Pentose phosphate pathway
V9	G6P (abcdef)→6PG (abcdef)+NADPH
<b>V</b> 10	6PG (abcdef)→Ru5P (bcdef)+CO₂ (a)+NADPH
<b>V</b> 11	Ru5P (abcde)↔X5P (abcde)
<b>V</b> 12	Ru5P (abcde)↔R5P (abcde)

 $V_{13}$ 

	X5P (abcde)↔TK-C2 (ab)+GAP (cde)
<b>V</b> <sub>14</sub>	F6P (abcdef)↔TK-C2 (ab)+E4P (cdef)
<b>V</b> 15	S7P (abcdefg)↔TK-C2 (ab)+R5P (cdefg)
<b>V</b> <sub>16</sub>	F6P (abcdef)↔TA-C3 (abc)+GAP (def)
<b>V</b> 17	S7P (abcdefg)↔TA-C3 (abc)+E4P (defg)
	Entner–Doudoroff pathway
<b>V</b> <sub>18</sub>	6PG (abcdef)→KDPG (abcdef)
<b>V</b> 19	KDPG (abcdef)→Pyr (abc)+GAP (def)
	TCA Cycle
V <sub>20</sub>	Pyr (abc)→AcCoA (bc)+CO₂ (a)+NADH
V21	OAC (abcd)+AcCoA (ef)→Cit (dcbfea)
<b>V</b> <sub>22</sub>	Cit (abcdef)⇔lCit (abcdef)
V <sub>23</sub>	ICit (abcdef)↔AKG (abcde)+CO₂ (f)+NADPH
V24	AKG (abcde)→SucCoA (bcde)+CO₂ (a)+NADH
• 20	SucCoA (abcd)⇔Suc (1/2 abcd+1/2 dcba)+ATP
V <sub>26</sub>	Suc (1/2 abcd+1/2 dcba)↔Fum (1/2 abcd+1/2 dcba)+FADH2
V <sub>27</sub>	Fum (1/2 abcd+1/2 dcba)↔Mal (abcd)
<b>V</b> <sub>28</sub>	Mal (abcd)↔OAC (abcd)+NADH

#### Glyoxylate Shunt

 $V_{29}$ 

ICit (abcdef)→Glyox (ab)+Suc (1/2 edcf+1/2 fcde)
--

**V**<sub>30</sub>

Glyox (ab)+AcCoA (cd) $\rightarrow$ Mal (abdc)

Amphibolic reactions

**V**31

Mal (abcd)→Pyr (abc)+CO<sub>2</sub> (d)+NADPH

V<sub>32</sub>

Mal (abcd)→Pyr (abc)+CO<sub>2</sub> (d)+NADH

PEP (abc)+CO<sub>2</sub> (d) $\rightarrow$ OAC (abcd)

 $V_{33}$  $V_{34}$ 

- OAC (abcd)+ATP $\rightarrow$ PEP (abc)+CO<sub>2</sub> (d)

Acetic acid formation

 $V_{35}$ 

AcCoA	(ab)↔Ac	(ab)+ATP
-------	---------	----------

Amino acid biosynthesis

<b>V</b> <sub>36</sub>	AKG (abcde)+NADPH+NH3→Glu (abcde)
V <sub>27</sub>	
• 51	Glu (abcde)+ATP+NH3→Gln (abcde)
V <sub>38</sub>	Glu (abcde)+ATP+2 NADPH→Pro (abcde)
V <sub>39</sub>	Glu (abcde)+CO₂ (f)+Gln (ghijk)+Asp (lmno)+AcCoA (pq)+5 ATP+NADPH→Arg (abcdef)+AKG (ghijk)+Fum (lmno)+Ac (pq)
<b>V</b> <sub>40</sub>	OAC (abcd)+Glu (efghi)→Asp (abcd)+AKG (efghi)
<b>V</b> <sub>41</sub>	Asp (abcd)+2 ATP+NH3→Asn (abcd)
V <sub>42</sub>	Pyr (abc)+Glu (defgh)→Ala (abc)+AKG (defgh)

V43

3PG (abc)+Glu (defgh)→Ser (abc)+AKG (defgh)+NADH

**V**44

Ser (abc) $\leftrightarrow$ Gly (ab)+MEETHF (c)

V<sub>45</sub>-V<sub>58</sub>:

Gly (ab)↔CO<sub>2</sub> (a)+MEETHF (b)+NADH+NH3

Thr (abcd)→Gly (ab)+AcCoA (cd)+NADH

Ser (abc)+AcCoA (de)+3 ATP+4 NADPH+SO4→Cys (abc)+Ac (de)

Asp (abcd)+Pyr (efg)+Glu (hijkl)+SucCoA (mnop)+ATP+2 NADPH→LL-DAP (abcdgfe)+AKG (hijkl)+Suc (1/2 mnop+1/2 ponm)

LL-DAP (abcdefg) $\rightarrow$ Lys (abcdef)+CO<sub>2</sub> (g)

Asp (abcd)+2 ATP+2 NADPH $\rightarrow$ Thr (abcd)

Asp (abcd)+METHF (e)+Cys (fgh)+SucCoA (ijkl)+ATP+2 NADPH →Met (abcde)+Pyr (fgh)+Suc (1/2 ijkl+1/2 lkji)+NH3

Pyr (abc)+Pyr (def)+Glu (ghijk)+NADPH→Val (abcef)+CO₂ (d)+AKG (ghijk)

AcCoA (ab)+Pyr (cde)+Pyr (fgh)+Glu (ijklm)+NADPH  $\rightarrow$ Leu (abdghe)+CO<sub>2</sub>(c)+CO<sub>2</sub> (f)+AKG (ijklm)+NADH

Thr (abcd)+Pyr (efg)+Glu (hijkl)+NADPH $\rightarrow$ IIe (abfcdg)+CO<sub>2</sub> (e)+AKG (hijkl)+NH3

PEP (abc)+PEP (def)+E4P (ghij)+Glu (klmno)+ATP+NADPH→Phe (abcefghij)+CO₂ (d)+AKG (klmno)

PEP (abc)+PEP (def)+E4P (ghij)+Glu (klmno)+ATP+NADPH →Tyr (abcefghij)+CO₂ (d)+AKG (klmno)+NADH

Ser (abc)+R5P (defgh)+PEP (ijk)+E4P (lmno)+PEP (pqr)+Gln (stuvw)+3 ATP+NADPH→Trp (abcedklmnoj)+CO₂ (i)+GAP (fgh)+Pyr (pqr)+Glu (stuvw)

R5P (abcde)+FTHF (f)+Gln (ghijk)+Asp (lmno)+5 ATP  $\rightarrow$ His (edcbaf)+AKG (ghijk)+Fum (lmno)+2 NADH

One-carbon metabolism

 $V_{59}$ 

```
MEETHF (a)+NADH\rightarrowMETHF (a)
```

V<sub>60</sub>

MEETHF (a)→FTHF (a)+NADPH

Oxidative phosphorylation

**V**61

NADH+1/2 O<sub>2</sub> $\rightarrow$ 3 ATP V<sub>62</sub> FADH2+1/2 O<sub>2</sub> $\rightarrow$ 2 ATP

**Transhydrogenation** 

V<sub>63</sub>

 $\mathsf{NADH}{\leftrightarrow}\mathsf{NADPH}$ 

ATP hydrolysis

**V**64

ATP→ATP:ext

**Transport** 

v65

Ac (ab)→Ac.ext (ab)

**V**66

 $CO_2$  (a) $\rightarrow$ CO<sub>2</sub>.ext (a)

 $O_2.ext \rightarrow O_2$ 

 $V_{68}$ 

V<sub>67</sub>

NH3.ext→NH3

**V**69

 $SO4.ext \rightarrow SO4$ 

**Biomass formation** 

**V**<sub>70</sub>

0.488 Ala+0.281 Arg+0.229 Asn+0.229 Asp+0.087 Cys+0.250 Glu+0.250 Gln+0.582 Gly+0.090 His+0.276 Ile+0.428 Leu+0.326 Lys+0.146 Met+0.176 Phe+0.210 Pro+0.205 Ser+0.241 Thr+0.054 Trp+0.131 Tyr+0.402 Val+0.205 G6P+0.071 F6P+0.754 R5P+0.129 GAP+0.619 3PG+0.051 PEP+0.083 Pyr+2.510 AcCoA+0.087 AKG+0.340 OAC+0.443 MEETHF+33.247 ATP+5.363 NADPH→39.68 Biomass+1.455 NADH

#### CO2 exchange

 $V_{71}$ 

 $CO_2$ .unlabeled (a)+ $CO_2$  (b) $\rightarrow CO_2$  (a)+ $CO_2$ .out (b)

## Appendix 1.2

# Predicted Mass Isotopomer Distributions for [U-Ala], [U-Asp], and [U-Gln]

[U-Ala]



[U-Asp]



[U-Gln]



## Appendix 1.3

# Experimental Data and Excel Sheet from Secondary <sup>13</sup>C-MFA Experiment

Mass isotopomer data (corrected for natural isotope a	abundances)									
	Ala t0	Ala t1	Ala t2	Asp t0	Asp t1	Asp t2	GIn t0	Gln t1	Gln t2	Labeled solution only
Ala260 (M0)	62.5%	49.7%	24.8%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%
Ala260 (M1)	0.5%	0.5%	0.7%	1.0%	0.9%	0.9%	0.9%	0.9%	0.9%	0.8%
Ala260 (M2)	2.6%	3.5%	5.2%	7.0%	7.0%	6.9%	7.0%	7.0%	7.0%	7.0%
Ala260 (M3)	34.4%	46.3%	69.3%	91.7%	91.8%	91.8%	91.8%	91.8%	91.8%	91.8%
Asp418 (M0)	0.2%	0.2%	0.2%	72.8%	67.2%	53.2%	0.3%	0.2%	0.2%	0.2%
Asp418 (M1)	0.4%	0.3%	0.4%	-0.2%	0.1%	0.0%	0.4%	0.4%	0.4%	0.4%
Asp418 (M2)	1.9%	1.9%	2.2%	0.5%	0.6%	0.9%	2.4%	2.2%	2.1%	1.9%
Asp418 (M3)	9.0%	9.0%	9.0%	2.5%	3.0%	4.2%	8.9%	9.0%	9.0%	9.1%
Asp418 (M4)	88.4%	88.4%	88.2%	24.5%	29.2%	41.6%	88.0%	88.2%	88.3%	88.4%
Gln431 (M0)	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	75.1%	65.2%	45.7%	0.3%
Gln431 (M1)	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	-0.3%	-0.1%	0.0%	0.1%
Gln431 (M2)	0.1%	0.1%	0.0%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
Gln431 (M3)	0.5%	0.6%	0.5%	0.4%	0.4%	0.4%	0.1%	0.2%	0.3%	0.4%
Gln431 (M4)	4.7%	4.6%	4.6%	4.5%	4.5%	4.6%	1.2%	1.6%	2.5%	4.7%
Gln431 (M5)	94.5%	94.5%	94.7%	94.9%	94.8%	94.8%	23.8%	33.1%	51.5%	94.5%
Volume of Sample [uL]:	100									
Volume of Standard [uL]:	150									
	Ala			Asp			Gln			
Concentration of Species in Std. [mM]:	7.24			4.66			4			
Concentration of Species in Sample [mM]:	18.4	10.8	3.6	18.9	14.6	8.1	18.0	11.3	5.1	
OD 600 measurement at time of sample:	0.032	0.430	0.912	0.111	0.365	0.881	0.184	0.454	0.901	
Dry weight biomass present at time of sample (Cmol)	0.00005	0.00069	0.00146	0.00018	0.00058	0.00141	0.00029	0.00073	0.00144	
Growth Conditions/ Assumptions:										
1 OD 600 = 0.32 g dry weight biomass/L										
125 mL flask										
25 g/ Cmol dry weight biomass										
Cx (mCmol/L)	0.41	5.50	11.68	1.42	4.68	11.28	2.35	5.81	11.53	
In(Cx [Cmol])	-7.81	-5.20	-4.45	-6.56	-5.37	-4.48	-6.05	-5.15	-4.46	5
time (hr)	0	12	15	0	9	15	0	27	48	