1   AN INTEGRATIVE BIOLOGY APPROACH TO QUANTIFY THE BIODISTRIBUTION OF
2   AZIDOHOMOALANINE *IN VIVO*
3
4   Aya M. Saleh[a], Tyler VanDyk[a], Kathryn R. Jacobson[b], Sarah Calve[a,b,c] and Tamara L. Kinzer-Ursem[a,b,*]
5   [a]Weldon School of Biomedical Engineering, Purdue University, 206 S. Martin Jischke Dr, West
6   Lafayette, IN 47906
7   [b]Purdue University Interdisciplinary Life Science Program, 155 S. Grant Street, West Lafayette, IN 47907
8   [c]Paul M. Rady Department of Mechanical Engineering, University of Colorado – Boulder, 1111
9   Engineering Center, 427 UCB, Boulder, CO 80309
10   [*]Corresponding author: Prof. Tamara L. Kinzer-Ursem, Email: tursem@purdue.edu

11   **Running Title:** Quantification of Aha Biodistribution

12   **Total Character Count:** 63,278

13

Abstract

15

16   Identification and quantitation of newly synthesized proteins (NSPs) are critical to understanding

17 protein dynamics in development and disease. Probing the nascent proteome can be achieved using non-

18 canonical amino acids (ncAAs) to selectively label the NSPs utilizing endogenous translation machinery,

19 which can then be quantitated with mass spectrometry. Since its conception, ncAA labeling has been

20 applied to study many in vitro systems and more recently the in vivo proteomes of complex organisms such

21 as rodents. In vivo labeling is typically achieved by introducing ncAAs into diet, which requires extended

22 labeling times. We have previously demonstrated that labeling the murine proteome is feasible via injection

23 of azidohomoalanine (Aha), a ncAA and methionine (Met) analog, without the need for Met depletion.

24 With the ability to isolate NSPs without applying stress from dietary changes, Aha labeling can address

25 biological questions wherein temporal protein dynamics are significant. However, accessing this temporal

26 resolution requires a more complete understanding of Aha distribution kinetics in tissues. Furthermore,

27 studies of physiological effects of ncAA administration have been limited to gross observation of animal

28 appearance. To address these gaps, we created a deterministic, compartmental model of the biokinetic

29 transport and incorporation of Aha in mice. Parameters were informed from literature and experimentally.

30 Model results demonstrate the ability to predict Aha distribution and labeling under a variety of dosing

31 paradigms and confirms the use of the model as a tool for design of future studies. To establish the suitability

32 of the method for in vivo studies, we investigated the impact of Aha administration on normal physiology

33 by analyzing the plasma metabolome following Aha injection. We show that Aha administration does not

34 significantly perturb cellular functions as reflected by an unchanged plasma metabolome compared to non-

35 injected controls.

36 **Keywords:** Non-canonical amino acids, protein labeling, kinetics, compartment modeling, metabolomics

37

38   **Abbreviations:** Aha, azidohomoalanine; HCA, hierarchical clustering analysis; hpi, hours post injection;

39   KEGG, kyoto encyclopedia of genes and genomes; KNN, k-nearest neighbor; LC-MS/MS, liquid

40   chromatography tandem-mass spectrometry; LHS, Latin hypercube sampling, Met, methionine; MRM,

41   multiple reaction monitoring; ncAA, non-canonical amino acid; NSP, newly synthesized protein; ODE,

42   ordinary differential equation; PCA, principal component analysis, PRCC, partial rank correlation

43   coefficient; rF, relative fluorescence; SEf, standard error of fitting; fAha, free Aha; pAha, proteinous Aha;

44   Sysrv, systemic venous reservoir.

45

3

Author Summary

As the machinery of life, proteins play a key role in dynamic processes within an organism. As such, the response of the proteome to perturbation is increasingly becoming a critical component of biological and medical studies. Dysregulation of protein mechanisms following exposure to experimental treatment conditions can implicate physiological mechanisms of health and disease, elucidate toxin/drug response, and highlight potential targets for novel therapies. Traditionally, these questions have been probed by studying perturbations in total proteins following an experimental treatment. However, the proteome is expansive and noisy, often an early response can be indiscernible against the background of unperturbed proteins. Here, we apply a technique to selectively label newly synthesized proteins, which enables capturing early changes in protein behavior. We utilize an amino acid analog that naturally incorporates into proteins, and investigate the tissue distribution, protein labeling efficiency, and potential physiological impact of this analog in mice. Our results demonstrate that we can reproducibly predict protein labeling and that the administration of this analog does not significantly alter in vivo physiology over the course of our experimental study. We further present a computational model that can be used to guide future experiments utilizing this technique to study proteomic responses to stimuli.

Introduction

The use of non-canonical amino acid (ncAA) labeling for selective identification of newly synthesized proteins (NSPs) in mammalian cells was first introduced by Dieterich et al. in 2006 (1) and has since been applied to study several biological systems (see reviews (2, 3)). In this technique, an ncAA, typically a methionine (Met) analog, is introduced to the biological system of interest and incorporated into newly synthesized polypeptide chains using endogenous or engineered cellular translational machinery Distinction of nascent proteins from the constituent proteome is enabled by reactive chemical groups, such as azides and alkynes, which can be covalently modified via azide-alkyne cycloaddition (a click chemistry reaction) (2). As such, ncAA-labeled NSPs can be selectively conjugated to affinity or fluorescent tags for identification or visualization, respectively (1, 4). This technique has been successfully employed to probe protein dynamics in a variety of bacterial (5-7) and mammalian cells *in vitro* (8-10), as well as model organisms *in vivo*, including zebrafish (11) and *Xenopus* (12). More recently, ncAA labeling has also been shown to be effective in identifying NSPs in rodents (13-15). The expanding applications of ncAA labeling will enable previously inaccessible biological questions, wherein understanding the temporal dynamics of protein synthesis and turnover is critical, to be addressed.

For rodent proteome labeling, dietary administration of ncAA, typically enhanced with a Met-free diet, has been shown to achieve adequate labeling efficiency (13, 16). However, Met deprivation may affect normal physiology, particularly over longer labeling periods. Notably, the presence of Met in mammalian diet is essential for normal embryonic development (17-19), which constrains this method to studies of adult animals. In this regard, our group has previously demonstrated that labeling the adult and embryonic murine proteome can instead be achieved via systemic injection of ncAAs without the need for Met depletion (14, 20). Compared to feeding with an ncAA-enriched diet, the injection method achieves global proteome labeling in a shorter period of time, which enables the detection of proteins synthesized shortly after injection and proteins with high turnover rates (20). In addition, injections allow for accurate dosing

5

85   calculations, which eliminates the inherent variability of the feeding method due to fluctuations in feeding

86   patterns and intestinal absorption.

87   Despite the application of ncAA labeling in a number of studies to decipher complex cellular

88   processes in animal models (13, 15, 16, 21), understanding of the kinetics of ncAA distribution in tissues,

89   especially as it pertains to rates of protein incorporation and loss by degradation, is lacking. Determination

90   of the timescale of ncAA uptake by tissues following administration and the lag time before maximum

91   protein labeling are critical information for the design of robust temporal experiments to study the nascent

92   proteome. Predicting ncAA pharmacokinetics in murine models will also enable optimizing the dosing

93   regimen to attain the ideal concentrations to achieve sufficient protein labeling in the desired tissue over

94   the course of the study.

95   In addition to the lack of knowledge of ncAAs distribution kinetics *in vivo*, evaluation of the

96   physiological impact of ncAA administration to animals has been limited to examining changes in gross

97   behavior, physical appearance and body weight (13-15). A more robust analysis of the effect of ncAA

98   administration on the metabolome, and the corresponding implications for cellular function, is required to

99   confirm the suitability of the method for *in vivo* studies.

100   The aim of this study was to characterize the distribution kinetics of azidohomoalanine (Aha), a

101   widely used Met analog, in mice following subcutaneous injection, and to investigate the impact on normal

102   physiology. To study the biodistribution of Aha, we measured the concentration in the plasma, liver, kidney,

103   brain and skeletal muscle using liquid chromatography-tandem mass spectrometry (LC-MS/MS) over a

104   period of 24 h. This dataset was used to develop a deterministic compartment model of small molecule

105   biokinetics that characterizes the movement of freely diffusive Aha (fAha) throughout the mouse

106   circulatory system and into tissues. In addition, we used fluorescent western blotting to measure protein

107   labeling in these tissues during the same period of time. This second dataset was used to inform a model of

108   relative protein labeling as a function of fAha availability to characterize both the predicted labeling profile

6

109   for a given experimental treatment of Aha and the relative synthesis and turnover rates of Aha-labeled

110   proteins. We demonstrate that this model can be used to characterize nascent protein synthesis and turnover

111   within distinct tissues. Furthermore, we validated the capability of this model to predict NSP labeling under

112   more complex, multiple injection dosing paradigms, which can be a tool to guide the design of future

113   experiments utilizing Aha labeling.

114   We also probed the effect of Aha incorporation into NSPs and investigated whether Aha labeling

115   perturbs normal physiological functions. We compared the plasma metabolome 24 h after Aha injection to

116   that of non-injected mice to identify if metabolic pathways were dysregulated due to protein labeling with

117   Aha. Only ~ 1.3% of metabolites were differentially regulated in the injected mice, indicating that Aha

118   administration does not have a significant impact on normal physiology. Taken together, these results

119   provide a fundamental understanding of the interrelation between the distribution kinetics of the ncAA into

120   murine tissues and the associated degree of protein labeling, as well as the impact of ncAA injection on

121   physiological functions.

122   **Results and Discussion**

123   Effective protein labeling is critical to enrich Aha-labeled proteins with high signal-to-noise ratio

124   for accurate quantitative MS measurements and identification of newly synthesized proteins. Depending on

125   the tissue type and biological processes to be studied, multiple injections of Aha may be required to attain

126   a high enough degree of labeling that results in a suitable MS signal. Therefore, optimizing the dose and

127   frequency of Aha injections is critical for the appropriate design of labeling studies. In this regard, we

128   sought to describe the distribution and labeling kinetics of Aha with an experimentally informed

129   deterministic model. Our model was developed to describe: (1) the transport of freely diffusive Aha (fAha),

130   Aha that is yet to be incorporated into protein, in the plasma, (2) the circulatory exchange of fAha into

131   tissues of adult mice and (3) the degree of Aha incorporation into proteins (pAha) in specific tissues.

132   **Kinetic Model of Aha Biodistribution Describes Transport and Exchange**

7

133 To capture the dynamics of fAha distribution *in vivo*, an *in silico* model system of ordinary

134 differential equations (ODEs) was generated describing the physiological processes of small molecule

135 transport. Dosed fAha was introduced into the model at a non-localized reservoir, to mimic the injection

136 site of our subcutaneous dosing paradigm. From this reservoir, fAha enters the murine circulatory system,

137 at a rate that is a function of reservoir concentration and transport kinetics, and is distributed to distinct

138 tissue compartments (Figure 1).
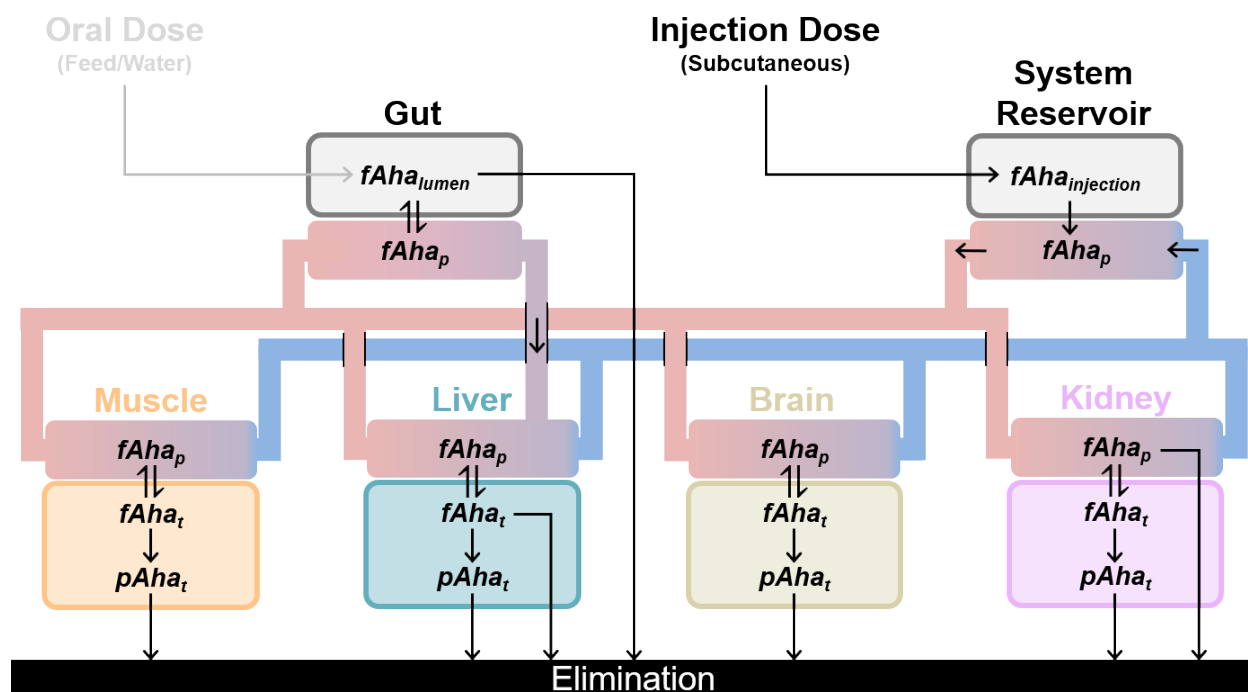
139
140 Figure 1. Biodistribution of fAha via transport and exchange. Introduced at a distinct injection site, fAha is
141 allowed to enter the circulation at a systemic venous reservoir. Driven by circulation, fAha is passed through
142 the arterial system (red) into distinct tissue compartments where exchange occurs at tissue specific rates.
143 Arrows indicate directional movement of fAha. Each tissue compartment consists of two sub-
144 compartments: plasma available for surface exchange (red to blue gradient) and an intracellular volume
145 (illustrated here as the bottom compartments: from left to right muscle, liver, brain, and kidney). A
146 mechanism for oral dosing (via plasma exchange with the gastrointestinal lumen) is illustrated, but not
147 included in this model.

148
149 Within each compartment, the time-dependent rate of change of the fAha plasma concentration

150 ($[fAha_p]$) available for exchange with each tissue can be described as a mass balance with two stages:

151 transport and exchange. The transport stage is governed by circulatory blood flow.

8

$$\left(\frac{d[fAha_p]_x}{dt}\right)_{transport} = \frac{Q_x}{V_x}([fAha_p]_{sysrv} - [fAha_p]_x) \qquad \textbf{Eq. 1}$$

152 where Q_x is the blood flow rate between tissue 'x' and a systemic venous reservoir (*sysrv*), and V_x is the

153 corresponding volume of plasma relevant to each tissue (Q/V represented as a lumped constant *qb* in

154 Supplements S2,S3). All kinetic parameters for circulatory transport were normalized by tissue mass to

155 compare relative perfusion rates between tissue compartments of differing size. Once localized to a tissue,

156 fAha in the plasma can also be exchanged across the cell membrane with the intracellular Aha concentration

157 ([*fAha_t*]) and is described as follows:

$$\left(\frac{d[fAha_p]_x}{dt}\right)_{exchange} = k_{e,x}[fAha_t]_x - k_{i,x}[fAha_p]_x \qquad \textbf{Eq. 2}$$

$$\left(\frac{d[fAha_t]_x}{dt}\right)_{exchange} = k_{i,x}[fAha_p]_x - k_{e,x}[fAha_t]_x \qquad \textbf{Eq. 3}$$

158 where $k_{i,x}$ and $k_{e,x}$ are the tissue specific import and export rates (min^{-1}·mg^{-1}) for fAha across the cell

159 membrane. The liver tissue, gut and kidney plasma compartments were assigned additional system removal

160 terms ($k_{r,x}$, Supplements S2,S3) accounting for excretion and metabolization of fAha. The two stages of

161 distribution were combined into a single system of ODEs, parameterized and bound within reasonable

162 ranges for a model of small molecule pharmacokinetics (Supplements S2-S5) (22-25).

163 **Kinetic Model of Protein Labeling Captures Aha Incorporation**

164 Within each tissue compartment, *fAha_t* is incorporated into proteins via protein synthesis. As a Met

165 analog, Aha is able to bind to methionyl-tRNA synthase, albeit at a much slower rate ($k_{cat}·K_m^{-1}$ Aha: 1.42E-3,

166 Met: 5.47E-1 μM^{-1}·s^{-1}) (26). Because the rate constant of Aha binding to the methionyl-tRNA synthase is

167 much slower than Met, and previous estimates of the amount of Aha incorporation into NSPs was less than

168 10% (14), we made a modeling assumption that [*fAha_t*] is negligibly depleted by incorporation into protein

169 (further justification, Supplement S2) and that recycling of Aha is nonexistent.

9

$$[fAha_t]_x \ll [pAha_t]_x \qquad \qquad \textbf{Eq. 4}$$

$$\left(\frac{d[fAha_t]_x}{dt}\right)_{synthesis} = -k_{s,x}[fAha_t]_x \approx 0 \ll \left(\frac{d[fAha_t]_x}{dt}\right)_{transport} \qquad \textbf{Eq. 5}$$

$$\left(\frac{d[pAha_t]_x}{dt}\right)_{synthesis} = k_{s,x}[fAha_t]_x - k_{d,x}[pAha_t]_x \qquad \textbf{Eq. 6}$$

170   where $k_{s,x}$ and $k_{d,x}$ are the tissue-specific (denoted for tissue 'x' as above) rate constants of incorporation of

171   Aha into proteins due to synthesis and loss of Aha-labeled proteins due to degradation. These equations

172   describing Aha labeling of proteins in each tissue were added to the biodistribution model establishing a

173   time resolved predictive model of tissue-specific protein labeling given a variety of input dosing paradigms.

174   **Experimental LC-MS/MS and Western Blotting Data Enables Parameter Fitting**

175   Model parameters were initialized and bound within reasonable ranges, informed from literature

176   and experimental measurements as described in the methods (Supplements S3-5) (22-25), then underwent

177   least squares regression to match experimentally measured data of Aha concentration and labeling. To

178   inform fitting, fAha concentration profiles in plasma and tissues were determined by injecting Aha

179   subcutaneously into mice at 0.1 mg·g⁻¹ total body weight and sacrificing 0.5-24 h post injection (hpi).

180   Accurate identification of fAha in each tissue was performed using LC-MS/MS multiple reaction

181   monitoring (Figure 2A). Additionally, the kinetics of Aha incorporation into tissue proteins were described

182   by examining the degree of protein labeling within each tissue over the duration of the study. To this end,

183   tissue homogenates were reacted with biotin-alkyne via copper-catalyzed click reaction, analyzed by

184   western blotting using a fluorescent streptavidin conjugate and the change in fluorescence intensity relative

185   to the non-injected controls was measured as an analog for pAha labeling (Figure 2B,C).

10
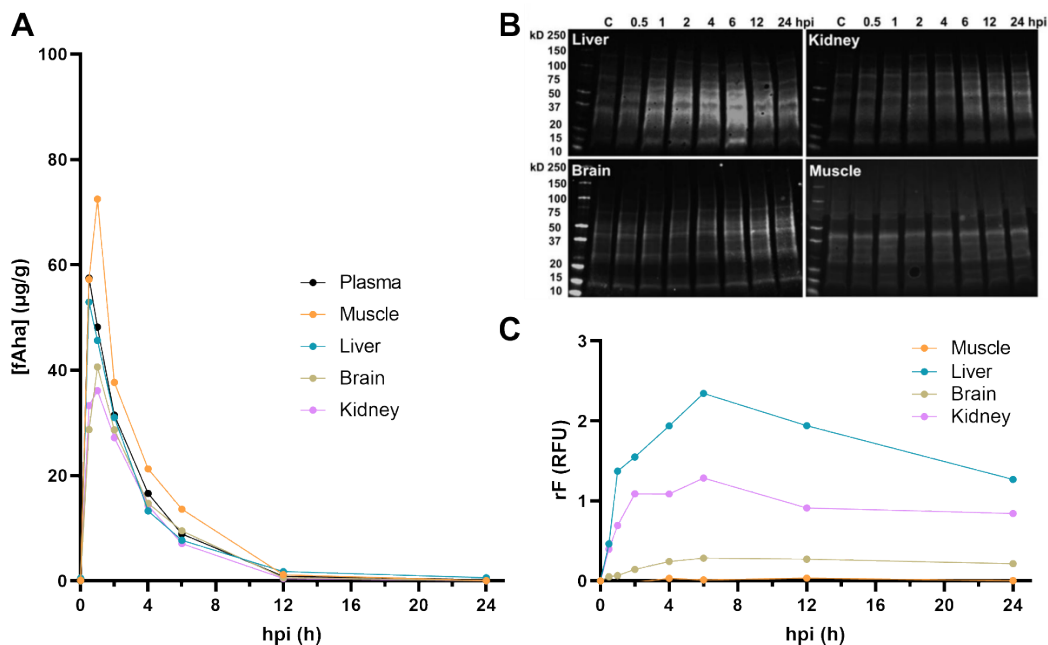
Figure 2. Free Aha concentration and protein incorporation and turnover kinetics in murine tissues were determined experimentally to inform the pharmacokinetics model. **(A)** The concentration profile of fAha in the plasma and different tissues. The amount of Aha (μg) measured by LC-MS/MS was normalized by the total plasma volume or tissue mass and averages were plotted over time. **(B)** Fluorescent western blots of the tissue homogenates of control non-injected samples (C) and samples collected 0.5 – 24 h post Aha injection (hpi). **(C)** Fluorescence intensity of western blot lanes were normalized to that of the respective control samples and averages were plotted as function of time (n=3 biological replicates).

The degree of fluorescent signal normalized relative to the background (rF) measured in relative fluorescence units (RFU) by semi-quantitative western blotting was assumed to be linearly proportional to the concentration of pAha using a fluorescent labeling factor, k_f. For each tissue (denoted 'x')

$$rF_x = \frac{signal - background}{background} = k_f[pAha_t]_x \qquad \textbf{Eq. 7}$$

$$\left(\frac{d(rF_x)}{dt}\right) = k_f k_{s,x}[fAha_t]_x - k_{d,x}(rF_x) \qquad \textbf{Eq. 8}$$

11

197   Aha Labeling Captures Relative Protein Synthesis and Turnover Dynamics in Murine Tissues

198   The concentration profiles of fAha in the plasma and tissues peaked between 0.5 and 1 hpi and was

199   mostly cleared from the system by 12 hpi, with the liver having the earliest peak compared to the other

200   tissues (Figure 3A-D). The early peak can be attributed to the high blood perfusion of the liver (27), which

201   likely results in faster distribution equilibrium of Aha into the liver compared to other tissues.

202   In all tissues, maximum protein labeling was observed around 6 hpi (Figure 3E-H). However, the

203   degree of labeling, represented by the maximum fold increase in fluorescence intensity compared to an

204   internal control, and the kinetics of protein incorporation and turnover varied considerably between tissues.

205   The liver showed the highest degree of labeling (Figure 3F) as well as the highest relative rates of Aha

206   incorporation and protein turnover (Table 1), whereas skeletal muscle had the lowest degree of labeling

207   (Figure 3E) and the slowest relative rates of incorporation and turnover (Table 1). Interestingly, the amount

208   of fAha (μg) per unit mass tissue (g) was higher in skeletal muscle than in liver (Figure 3A,B), indicating

209   that the low degree of labeling observed in skeletal muscle is predominantly due to a slow rate of muscle

210   Aha incorporation, implying a lower rate of Met incorporation and protein synthesis.

211   The observed differences in fluorescence intensities are in agreement with previous isotope labeling

212   studies that showed faster protein turnover rates in liver and kidney compared to brain and skeletal muscle

213   (28-30). Notably, our model estimated a protein half-life in the brain that is 2.7 times higher than the liver

214   (Table 1), in close agreement with the findings of Price et al. (9 h^{-1} and 3 h^{-1} days for brain and liver,

215   respectively) (29). The discrepancies between previously reported values and the half-lives values estimated

216   here can be attributed to the shorter timescale of our experimental setup as compared to stable isotope

217   labeling technique and the use of semi-quantitative western blotting measurements rather than MS. A more

218   robust quantitation of pAha concentrations using MS will be required for a more precise determination of

219   absolute protein kinetics in different murine tissues.

12
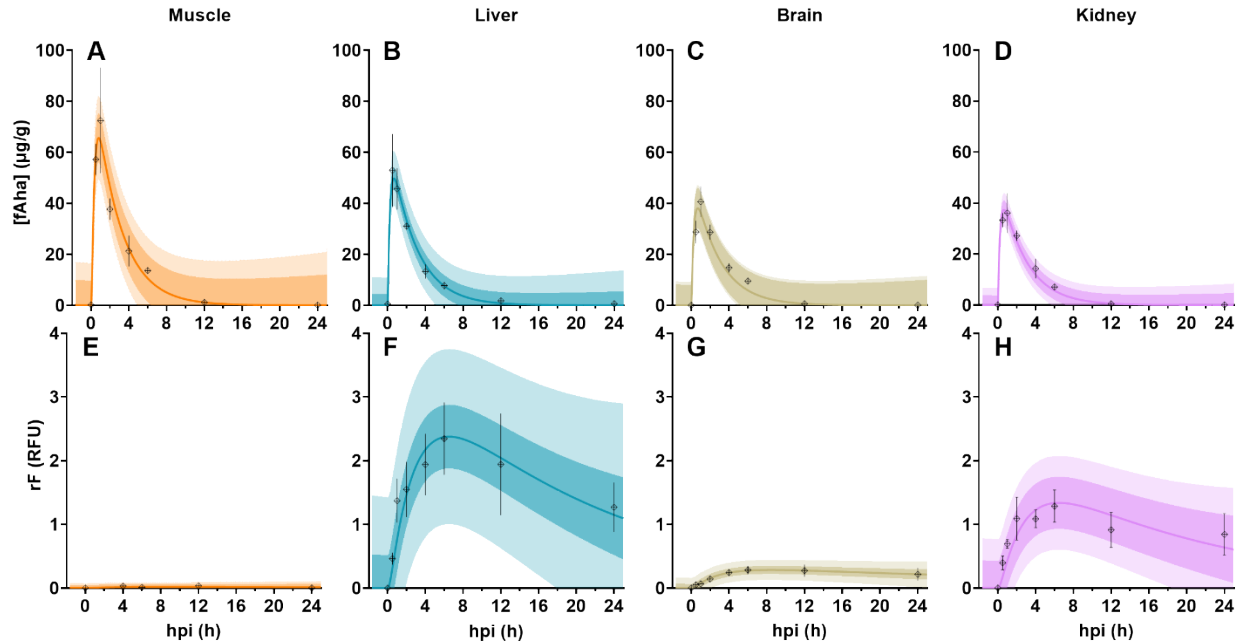
Figure 3. Aha kinetics and protein labeling following subcutaneous injection of Aha. **(A-D)** The concentration profiles of free Aha (fAha) in the plasma and different tissues. The amount of Aha (μg) measured by LC-MS/MS analysis was normalized by the total tissue mass and plotted over time. **(E-H)** Relative fluorescence (rF), as measured by western blotting, of proteins isolated from each tissue as a function of time. Filled points represent mean experimental measurement at each time point, error bars indicate experimental standard deviation (n = 3 biological replicates). Colored traces indicate best fit of model to each dataset, with darker and lighter shaded regions showing 95% prediction intervals for residual error calculated from mean experimental values and all experimental replicates, respectively.

13

Table 1. Protein incorporation and degradation estimated from Aha dynamics. Best fit parameter values for Aha incorporation, and protein turnover and half-lives in studied tissues.

Parameter, Units	Muscle	Liver	Brain	Kidney
Protein turnover rate, h^{-1}	2.23E-2	4.97E-2	2.12E-2	5.06E-2
Protein half-life, h	31.05	13.94	32.6	13.69
Aha Incorporation*, $RFU(\mu g\,pAha \cdot h)^{-1}$	5.36E-4	2.28E-2	6.06E-3	5.43E-2

*Relative incorporation rate $= k_f \cdot k_s \cdot mt^{-1}$, units include RFU per mass pAha.

Descriptive Statistics Support Model of Biodistribution and Labeling

Model validity was examined using the following metrics to investigate parameter stability and goodness of fit. First, a prediction interval of residuals was generated for each tissue studied, for both the biodistribution and protein labeling models. A 95% prediction interval ($PI_{\alpha=0.05}$) of residuals was calculated using all experimental replicates (Figure 3). A second, tighter prediction interval was calculated using the mean for each time point (Figure 3). For each tissue, the width of $PI_{\alpha=0.05}$ from the average line of best fit ($\hat{y}$) can be approximated using a naively informed forecast interval that assumes a normal distribution of residual error (31).

$$PI_{\alpha=0.05}(t) = \hat{y}(t) \pm 1.96 * \frac{\sum (y_i - \hat{y}_i)^2}{n} * \sqrt{1 + \frac{1}{n} + \frac{(t - \bar{t})^2}{\sum (t_i - \bar{t})^2}} \qquad \textbf{Eq. 7}$$

where, n is the total number of observations, (t_i, y_i) are the coordinates for each observation, t is the time point of the predicted residual, $\bar{t}$ is the average time of all experimental observations. These prediction intervals demonstrate a narrow range of residual error and capture all experimental means and >95% of experimental values, indicating an accurate predictive model.

Second, the covariance matrix of least squares regression was used to inform a standard error of fitting (SE_f) for all fitted parameters in the biodistribution model (Supplements S3-S5). SE_f describes variability of each parameter, but also reflects upon the definition of the model. Parameters with best fit values near to the constraints are less predictable; wide error would be indicative of a poorly constrained system. Error also increases with the number of fitted parameters, is inversely related to the quantity of data

14

250   points available for fitting, and is weighted by the metric used for optimization. Among all fit parameters,

251   we found relatively few SE_f values greater than best fit values by more than an order of magnitude (Figure

252   3, Supplements S3-5). Notably, we found that the widest error ranges in parameters that were the least well

253   characterized in the literature, specifically rates describing tissue import and export of fAha. Low SE_f

254   values, particularly among parameters within reported literature values, support a well characterized model,

255   although this metric is not fully sufficient to describe a complex non-linear regression.

256   A more thorough analysis was achieved using Monte-Carlo Latin hypercube sampling (LHS) to

257   perform efficient sampling of the input parameter space and correlation with partial rank correlation

258   coefficients (PRCCs). This global sensitivity analysis was performed against several metrics to probe the

259   influence of variation in parameter values on (1) model fitness (2) predicted Aha labeling levels in each

260   tissue (rF). To address the first of these, PRCC values, which vary between 1 (perfect positive correlation)

261   and -1 (perfect negative correlation), were calculated comparing variation in the input parameter values

262   against the sum of square errors for each fitting model. This analysis generates a PRCC value for each

263   parameter ($PRCC_f$) that characterizes its relative effect on the fitting metric, and therefore the fitting

264   process. $PRCC_f$ values higher in absolute magnitude indicate parameters that, when varied independently,

265   have the greatest influence on the fitting metric with values > 0.8 considered as indications of regions of

266   instability in the model (32, 33). All parameters fall safely below a $PRCC_f$ value of 0.3. Parameters with

267   the highest $PRCC_f$ values are those that influence the system removal of fAha, likely due to the rapid

268   metabolic profile exhibited following subcutaneous injection (Figure 3). $PRCC_f$ values for elimination and

269   import rates into tissues that work actively to remove fAha were > 0.2.

270   Taken together, $PRCC_f$ and SE_f values demonstrate that influential parameters are not those with

271   the widest error ranges (Figure 4, Supplements S5-S7). Furthermore, when comparing these values across

272   the parameter space, there is no one tissue that exhibits over-sensitive behavior by either metric, beyond

15

273   systemic elimination. These findings provide support for the model definition, boundary constraints, and

274   biological relevance of best fit parameters.
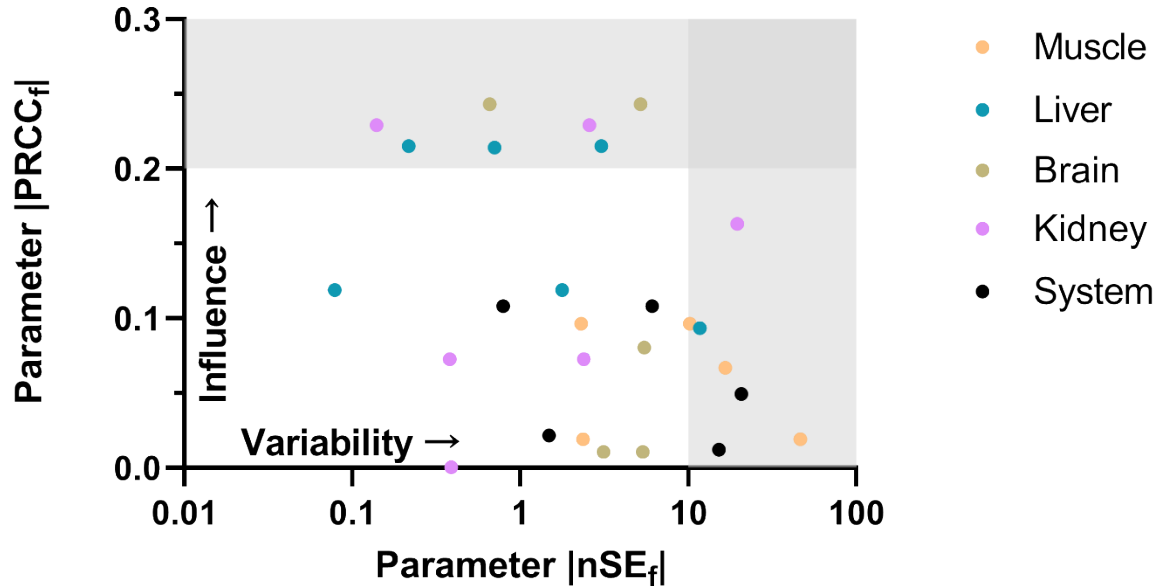
275
276   Figure 4. Sensitivity analysis for model parameter fitting. Each point represents a distinct parameter from
277   the model (colored by associated tissues) and coordinates indicate absolute values of normalized standard
278   error of fitting ($nSE_f = SE_f$/ best fit value) on the x-axis and $PRCC_f$ on the y-axis. Light gray shaded regions
279   indicate parameters that have either a relatively high model influence ($PRCC_f > 0.2$) or variability ($nSE_f >$
280   10). The dark grey shaded region indicates a domain where parameters would be considered unstable or
281   poorly constrained with both a high variability and influence on the model.
282
283   Further application of PRCC analysis allowed examination of the influence of parameter values

284   upon the predicted degree of Aha labeling in each tissue over time. All model parameters were varied within

285   their estimated fitting range (Supplements S5-S7) and sampled via LHS (n=10000). An rF value in each

286   tissue was predicted for each timepoint from 0-24 hpi and the resulting output was correlated to the input

287   parameter variation using PRCC analysis. Resulting traces elucidate kinetic parameters and physiological

288   mechanisms that most influence rF during different time domains (Figure 5). For example, immediately

289   after injection (0-4 hpi), the rate of absorption of Aha from initial injection site (*kaSubcu*) is a critical

290   mechanism driving rapid labeling efficiency in all four studied tissues. However, the importance of this

16

291   parameter declines over time, just as the influence of other parameters increases, particularly those

292   representing Aha degradation and elimination. The influence of liver transport is prominent in all four

293   tissues indicating the liver plays a major role in Aha kinetics. This suggests Aha may be primarily

294   metabolized and eliminated by the liver, rather than by eliminated via excretion. This possibility is

295   consistent with the known liver functions of toxin removal and amino acid metabolism (34).
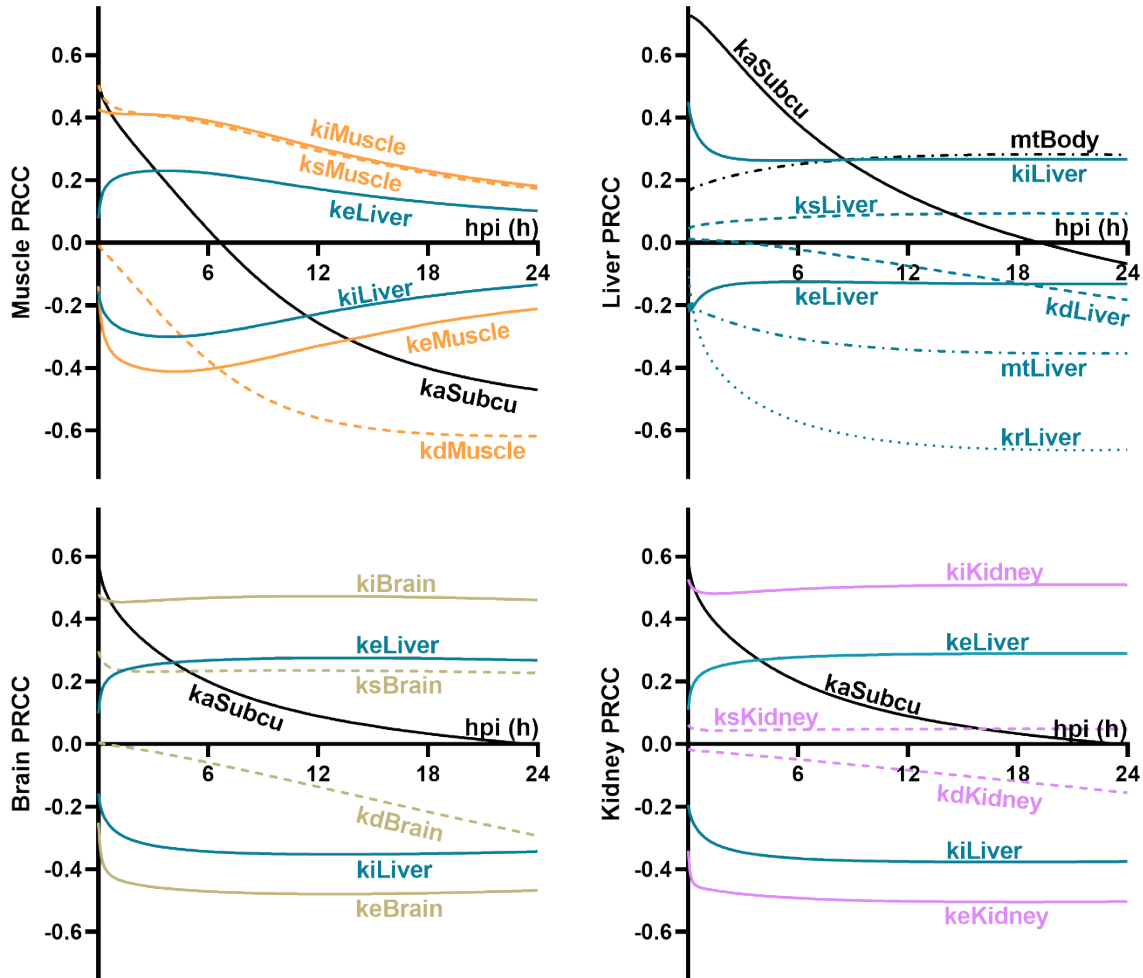

296
297   Figure 5.Sensitivity analysis of input parameter variation and Aha labeling in tissues over 24 hpi. Each plot
298   shows the time resolved PRCC values for a selection of parameters using predicted rF in each tissue as a
299   correlation metric. Parameters shown are either (1) descriptive of dynamics within the relevant tissue or (2)
300   related to a different tissue, but with a significant influence on the tissue of interest based upon PRCC
301   values > |0.2|. Trace colors indicate associated tissues (black = systemic param8883eter, blue = liver, orange
302   = muscle, green = brain, purple = kidney), trace patterns indicate parameter type (solid = fAha transport,
303   dotted = fAha elimination, dashed = pAha synthesis/degradation, dot-dashed = tissue mass).

17

304
305 **Predictive Simulations Accurately Capture Alternate Dosing Paradigms**

306   Attaining sufficient protein labeling is critical for accurate identification and quantitation of Aha-

307 labeled proteins using LC-MS/MS. For instance, note the near negligible labeling of skeletal muscle in

308 Figure 3. If muscle labeling is desired, a much higher Aha dose might be required to attain sufficient

309 labeling. This underlines the importance of tailoring the dosing regimen of Aha (*i.e.* amount per dose and

310 dosing frequency) to the tissue of interest and the biological question under investigation. Using the model

311 described above, fAha biodistribution and tissue protein labeling can be predicted for alternative dosing

312 regimens to aid future experimental design and predict labeling efficiency depending on the conditions of

313 a study. However, the model was based on the concentration of Aha in plasma and tissues over 24 h after a

314 single subcutaneous injection. While this data was sufficiently robust to develop a well-informed model of

315 Aha patterning after one subcutaneous dose, the use of this model to predict Aha content for longer time

316 scales or multiple dose paradigms would generate naïve forecasts. To address this gap, we validated the

317 predictions of our model against a data set that spans a longer time period and multiple injected doses. The

318 model was used to predict the rF labeling of brain and liver tissues for two alternative dosing paradigms

319 with either 12 h repeated doses (hrd) or 24 hrd over a 36 h period (Figure S7). As an internal control for

320 western blotting variation, these new experimental values were normalized by a shared time point with the

321 previous study (6 hours post initial injection) for each tissue (Figure 6).

18
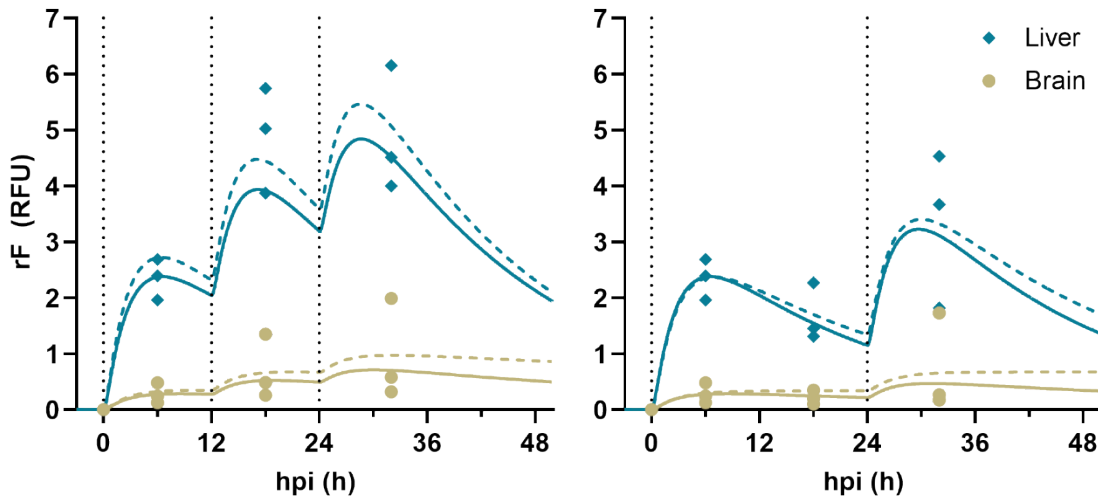
Figure 6. Model accurately predicts relative labeling in the brain and liver with multiple injection doses. rF experimental data (dots) and model predictions (lines) for **(A)** 12 hrd and **(B)** 24 hrd. For each tissue, experimental replicates from the repeated dose study are displayed at 6, 18, and 32 hpi as individual points. The solid line is the predicted trace using the best fit model from the original robust dataset. The dashed line is the model with parameter values refit to the data from the repeated dose experiments. The vertical dotted line indicates dose injection timepoints.

To determine the ability of the original model to predict Aha incorporation into proteins in various tissues, the data from each repeated dose study was used to refit the relative pAha synthesis rate and degradation for each tissue under each repeated dose paradigm. Relative to the parameter fit with the original experimental data, there was only a slight reduction in the standard error of regression (SE_{reg}), a goodness-of-fit metric, between the original and refit parameters in each tissue (Table 2). Additionally, among all refit parameters, a single parameter was adjusted beyond a single standard error of fit (SE_f) from the original best fit value (12 hrd liver $\Delta(k_f \cdot k_s) \approx +1.97SE$), and only the degradation rate in the brain changed by >20% (Table 2).

Table 2. Parameter and goodness-of-fit statistics related to the alternative dosing models in Figure 6.

Tissue	Parameter, Units	Pred. Value (SE$_f$)	Refit Value	%Δ
Liver, 12 hrd	$k_f \cdot k_s, RFU(\mu g\, fAha \cdot \mu g\, pAha \cdot h)^{-1}$	3.16E-4 (2.49E-5)	3.65E-4	+ 15.5%
	k_d, h^{-1}	8.29E-4 (1.80E-4)	8.67E-4	+ 4.6%
	SE_{reg}, RFU	1.975	1.684	- 14.7%
Brain, 12 hrd	$k_f \cdot k_s, RFU(\mu g\, fAha \cdot \mu g\, pAha \cdot h)^{-1}$	4.30E-4 (2.83E-5)	4.84E-5	+12.5%
	k_d, h^{-1}	3.54E-4 (1.11E-3)	1.37E-4	- 61.3%
	SE_{reg}, RFU	1.147	1.082	- 5.7%

19

Liver, 24 hrd	$k_f \cdot k_s, RFU(\mu g\, fAha \cdot \mu g\, pAha \cdot h)^{-1}$	3.16E-4 (2.49E-5)	3.02E-4	- 4.4 %
	k_d, h^{-1}	8.29E-4 (1.80E-4)	6.72E-4	- 18.9%
	SE_{reg}, RFU	1.558	1.524	- 2.2%
Brain, 24 hrd	$k_f \cdot k_s, RFU(\mu g\, fAha \cdot \mu g\, pAha \cdot h)^{-1}$	4.30E-4 (2.83E-5)	4.28E-5	- 0.5%
	k_d, h^{-1}	3.54E-4 (1.11E-3)	1.37E-9	- 99.9%
	SE_{reg}, RFU	0.953	0.916	- 3.9%

339

Aha Administration Does Not Perturb Normal Physiology in Mice

341   In addition to the characterization of Aha distribution kinetics in mice, the impact of Aha

342   administration on normal physiology must be qualified to establish the applicability of the method for *in*

343   *vivo* studies. To this end, the physiological impact of Aha incorporation into newly synthesized proteins

344   was evaluated using untargeted plasma metabolomic analysis. Since metabolites are the end products of

345   cellular biological processes, we reasoned that a lag time is expected between potential changes in protein

346   functions due to Aha incorporation and any associated effects on metabolism. Therefore, given that

347   maximum protein labeling occurred ~ 6 hpi (Figure 2B,C and Figure 3E-H), we analyzed the plasma

348   metabolome 24 hpi to identify any potential changes in metabolic pathways in response to Aha

349   incorporation.

350   LC-MS metabolomic analysis of the plasma identified a total of 1268 mass features (*i.e.*

351   metabolites). The peak area of each mass feature is proportional to the amount of the corresponding ion in

352   the sample and was used as a measurement for the relative abundance of each identified metabolite across

353   samples. Principal component analysis (PCA) revealed no distinct segregation between the control and Aha

354   mice, indicating that there were no global differences in the plasma metabolome between the two groups

355   (Figure 7A).

356   It should be noted that PCA cannot be performed in the presence of missing values. The occurrence

357   of missing values is common in untargeted metabolomic data, resulting from the presence of metabolites

358   with concentrations that are lower than the MS detection limit or due to technical reasons such as incomplete

20

359   ionization or inaccurate peak detection [35]. In our dataset, a total of 194 (2.5%) missing values were

360   detected across all samples. Since the percentage of missing values was low, it was assumed that the

361   potential impact of missing values is insignificant [36]. Therefore, the remaining 1112 mass features were

362   used for PCA (Figure 7A). To confirm the validity of this approach, PCA was also conducted on the dataset

363   after missing value imputation using the K-nearest neighbor (KNN) method and showed similar indistinct

364   grouping of the injected and control mice (Figure S8A).

365   In addition to PCA, unsupervised hierarchical clustering analysis (HCA) was conducted and a

366   heatmap was generated to examine variations in metabolic patterns between the Aha and control groups

367   (Figure 7B). HCA and heatmap visualization showed no clustering between the biological replicates of

368   each group and no distinct differential abundance patterns between the two groups. This result further

369   establishes that there are no substantial metabolic differences between control and injected mice. Similar to

370   PCA, HCA performed using the dataset imputed via the KNN method resulted in indistinct clustering of

371   the mice (Figure S8B).

372   Following global analysis using PCA and HCA, Student's *t*-test was employed to identify

373   metabolites that were differentially regulated between the two groups. In this analysis, a total of 15 out of

374   1112 metabolites were differentially abundant using a *p*-value of 0.05 and a fold change of ≥ 2 as cut-offs

375   (Figure 7C). Of the 15 metabolites, 3 were upregulated and 12 were downregulated in Aha-treated mice

376   compared to the control. Searching the 15 metabolites in the METLIN metabolite database using a mass

377   tolerance of 5 ppm did not identify any known metabolite. The presence of a large number of unknown

378   mass features is an intrinsic characteristic of untargeted metabolomic studies due to the complexity of the

379   mammalian metabolome and the lack of structure characterization of a large number of metabolites (35,

380   36). Yet, the fact that only ~1.3% of metabolites were dysregulated, and that these dysregulated metabolites

381   did not belong to any of the known major metabolic pathways, signifies that minimal metabolic alterations

382   occur due to Aha administration.

21
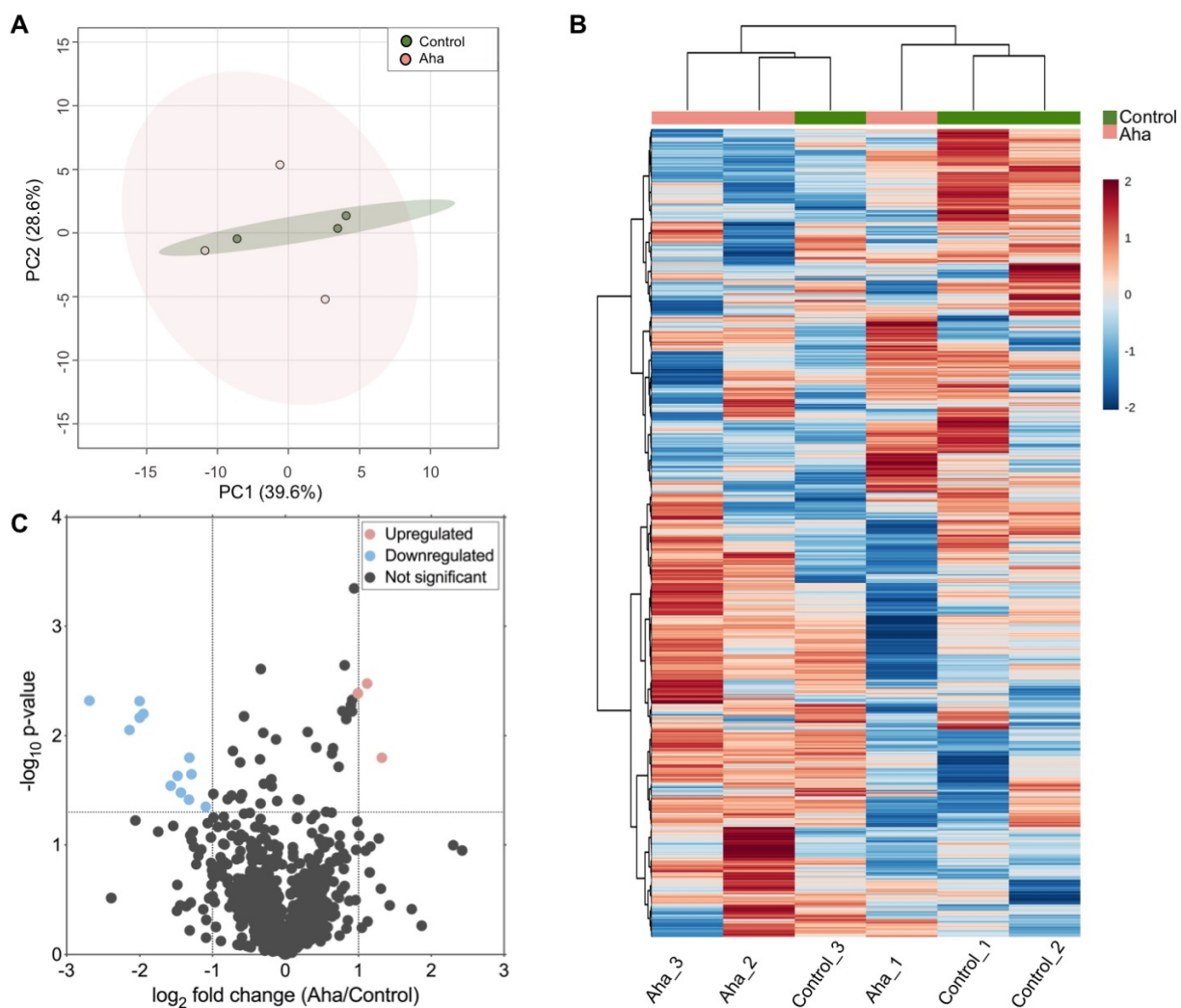
Figure 7. Aha administration does not significantly change the murine plasma metabolome. **(A)** Principal component analysis (PCA) showed no clear separation between control and Aha-treated groups in the first two components. Components 1 and 2 account for 39.6% and 28.6% of the total data variability, respectively. Green and red dots denote control and Aha samples, respectively. Green and red shaded areas represent the 95% confidence bands of the control and Aha samples, respectively. **(B)** Heat map of unsupervised hierarchal clustering analysis (HCA) of the identified metabolites show lack of clustering between replicates of each group. Color scale indicates metabolite abundance; blue: lowest, red: highest. **(C)** Volcano plot comparing the relative abundance of the identified metabolites between control and Aha groups according to statistical significance and fold change. Horizontal line indicates p-value = 0.05 and vertical lines indicate $\pm$ 2-fold change. Grey, red and blue circles denote equally-abundant, upregulated and downregulated metabolites, respectively.

Finally, to identify the metabolic pathways covered by the LC-MS analysis, analysis of equally-expressed metabolites was conducted using Metaboanalyst and the Kyoto Encyclopedia of Genes and

22

397   Genomes (KEGG) metabolic pathway database (37). Several metabolic pathways were detected, including

398   arachidonic acid metabolism, vitamin B6 metabolism, valine, leucine and isoleucine biosynthesis, galactose

399   metabolism, and cysteine and Met metabolism (Table S9). This result indicates that the LC-MS analysis

400   identified metabolites that belong to various metabolic pathways and that these pathways are not

401   significantly changed in mice injected with Aha.

402   Collectively, the metabolomic analyses demonstrate that Aha administration does not significantly

403   alter the murine plasma metabolome. This is in agreement with a recent study that investigated the

404   metabolic effect of growing *E.coli* in media supplemented with ncAAs (38). A major advantage of our

405   labeling technique is that it does not involve Met restriction or depletion as the case with other labeling

406   strategies that use a Aha-enriched Met-free diet. Met dietary restriction has been shown to alter the

407   metabolism in mouse models and in humans (39-41). Being a principle sulfur-containing amino acid, Met

408   restriction specifically alters Met and sulfur metabolism (40, 41). Notably, the results of the pathway

409   analysis identified several unchanged metabolites such as L-homocysteine, 5'-Methylthioadenosine, and 3-

410   sulfinoalanine that belong to cysteine (another sulfur-containing amino acid) and Met metabolism (Table

411   S9), indicating the advantage of the injection method with regards to its potential impact on metabolic

412   functions.

413   **Methods**

414   **Animal Model**

415   Animals used in these studies were derived from female age-matched wild-type C57BL/6 mice

416   (*Mus musculus*) purchased from The Jackson Laboratory. All experimental protocols were performed in

417   compliance with established guidelines and all methods were approved by Purdue Animal Care and Use

418   Committee (PACUC, protocols# 1209000723 and 1801001682). PACUC requires that all animal programs,

23

419   procedures, and facilities at Purdue University abide by the policies, recommendations, guidelines, and

420   regulations of the United States Department of Agriculture (USDA) and the United States Public Health

421   Service (USPHS) in accordance with the Animal Welfare Act and Purdue's Animal Welfare Assurance.

422   **Aha Injection, and Plasma and Tissue Collection**

423   L-azidohomoalanine (Aha; Click Chemistry Tools) was resuspended in $1 \times$ phosphate buffered

424   saline (PBS) to 10 mg·mL^{-1}, adjusted to pH 7.4, sterile filtered and stored at -20°C. All Aha injections were

425   administered subcutaneously at 0.1 mg·g^{-1} total mouse weight. Mice (n = 3, biological replicates) were

426   euthanized 0.5, 1, 2, 4, 6, 12 and 24 h post injection (hpi). Blood was harvested by cardiac puncture,

427   collected in EDTA-treated tubes and centrifuged at $1,500 \times g$ for 10 min at 4°C. The supernatant (plasma)

428   was transferred into a new tube, snap frozen in liquid nitrogen and stored at -80°C. Liver, brain, kidney and

429   hindlimb skeletal muscle tissues were dissected at each time point, snap frozen in liquid nitrogen and stored

430   at -80°C. Control plasma and tissues were collected as described above from non-injected mice (n = 3

431   biological replicates). For the validation of model predictive ability, two Aha dosing regimens were used:

432   (1) 12 h repeated doses (hrd) and (2) 24 hrd. Liver and brain tissues (n = 3 biological replicates) were

433   dissected as described above at 6, 18, and 32 hpi, snap frozen in liquid nitrogen and stored at -80°C.

434   **Sample Preparation for Aha Analysis**

435   For plasma sample preparation, 50 µL of plasma were mixed with 10 µL of $1 \times$ PBS, pH 7.4, and

436   5 µL of 100 ng·µL^{-1} L-α-aminobutyric acid (α-ABA; Sigma Aldrich) that was used as an internal standard.

437   12.5 µL of trichloroacetic acid (TCA; Sigma Aldrich) were added to the mixture to precipitate proteins.

438   The mixture was incubated for 10 min at 4°C and centrifuged at $16,000 \times g$ for 10 min at RT. The

439   supernatant was then mixed with 100% acetonitrile (ACN; Fisher Scientific) at a 1:1 ratio (v/v). The mixture

440   was transferred to an HPLC autosampler vial for LC-MS/MS analysis. For calibration curve generation,

24

441   Aha standards were prepared by mixing 50 µL of non-injected plasma with 10 µL of a known concentration

442   of Aha and 5 µL of α-ABA. Proteins were then precipitated with TCA and prepared for LC-MS/MS analysis

443   as described above.

444   For tissue sample preparation, tissues were rinsed with ice-cold $1 \times$ PBS, pH 7.4 to remove residual

445   blood and homogenized in ice-cold $1 \times$ PBS, pH 7.4 using a TissueRuptor (Qiagen). The final homogenate

446   weight was measured and converted to volume by using a homogenate density of 1 g·mL^{-1}. Samples were

447   then prepared for LC-MS/MS analysis as described for plasma by using 50 µL of the tissue homogenate.

448   The remaining plasma samples and tissue homogenates were snap frozen and stored at -80°C until use for

449   western blot and untargeted metabolomic analyses as described below.

450   **LC-MS/MS Targeted Analysis of Aha**

451   An Agilent 1260 Rapid Resolution liquid chromatography (LC) system coupled to an Agilent 6470

452   series QQQ mass spectrometer was used for Aha analysis (Agilent Technologies). An Intrada Amino Acid

453   2.0 mm x 150 mm, 3.0 µm column (Imtakt Corporation) was used for LC separation. The buffers were (A)

454   ACN, 0.3 % formic acid (FA; Sigma Aldrich) and (B) ACN/100 mM ammonium formate (20/80 v/v). The

455   linear LC gradient was as follows: time 0 min, 20 % B; time 5 min, 20 % B; time 11 min, 35 % B; time 20

456   min, 100 % B; time 22 min, 100 % B; time 22.5 min, 20 % B; time 30 min, 20% B. The flow rate was 0.3

457   mL·min^{-1}. Multiple reaction monitoring (MRM) was used for MS analysis. Data were acquired in a positive

458   electrospray ionization (ESI) model based upon parameters in Table 3. The jet stream ESI interface had a

459   gas temperature of 325°C, gas flow rate of 9 L·min^{-1}, nebulizer pressure of 35 psi, sheath gas temperature

460   of 250°C, sheath gas flow rate of 7 L·min^{-1}, capillary voltage of 3500 V in a positive mode, and nozzle

461   voltage of 1000 V. The delta electron multiplier voltage was 300 V. Agilent MassHunter Quantitative

462   Analysis software was used for data analysis (v.8.0).

463

464 Table 3. Multiple reaction monitoring (MRM) table for amino acid LC-MS/MS data acquisition

Compound name	Precursor ion (m/z)	Product ion (m/z)	Collision energy (eV)
Aha	145.1	101.3	5
Aha	145.1	71.3	10
Aha	145.1	58.3	40
Ala	90	44	15
Arg	175	116	18
Asn	133	87	12
Asp	134	88	14
Cys	122	76	15
Cys-Cys	241.1	152	15
Gln	147	84	22
Glu	148	130	12
Gly	76	30	15
His	156	110	19
Ile	132	86	15
Leu	132	86	15
Lys	147	84	20
Met	150	104	15
Phe	166	120	15
Pro	116	70	15

465
466 **Western Blot Analysis of Aha-Labeled Tissues**

467 Tissue homogenates were thawed and protein concentration was measured using the Pierce 660 nm

468 Protein Assay (ThermoFisher Scientific). 200 µg of tissue homogenate was alkylated with 40 mM

469 iodoacetamide for 30 min at RT in the dark with end-over-end rotation. Samples were then reacted for 2 h

470 at RT with the following click reagents: 50 µM biotin-alkyne (ThermoFisher Scientific), 5 mM tris(3-

471 hydroxypropyltriazolylmethyl)amine (THPTA; Click Chemistry Tools), 2 mM copper sulfate, 20 mM

472 aminoguanidine and 10 mM sodium ascorbate. Following the click reaction, proteins were precipitated by

473 adding ice-cold 100% acetone to the samples at a 4:1 ratio (v/v). Samples were incubated overnight at -

474 20ºC, centrifuged at 21,100 × g for 20 min at 4ºC, supernatants were discarded, and protein pellets were

475 vacuum-dried for 15 min at RT using a CentriVap (Labconco). Dried pellets were resuspended in 8 M urea

476 in 1× PBS and centrifuged at 16,000 × g for 15 min at RT to remove insoluble particles. The supernatants

26

477   were transferred into new tubes and protein concentration was measured using the Pierce 660 nm Protein

478   Assay (ThermoFisher Scientific). Proteins were resolved on 4 – 20% SDS-PAGE gels (BioRad), transferred

479   to a PVDF membrane (ThermoFisher Scientific) using the Trans-Blot Turbo Transfer System (BioRad) and

480   probed overnight at 4°C with IRDye 680 Streptavidin (LICOR) diluted 1:3000 in 1:1 TBST:Blocking

481   Buffer (BioRad). Membranes were imaged using an Azure Biosystems c600. Western blot images were

482   analyzed using ImageJ (National Institutes of Health) to calculate the mean fluorescence intensities of each

483   time point. The intensity of the control sample was used to normalize the intensity of each time point (n =

484   3 biological replicates per blot).

485   **Kinetic Modelling of Aha Distribution**

486   Simulations were run on a Lenovo Yoga with an Intel Core i7-8550U CPU @ 1.8 GHz and 8 GB

487   RAM. Simulations were performed using custom modeling scripts written in Python 3.6 (Supplement S1).

488   Systems of ordinary differential equations (Supplement S2) were solved using a flexible high order solver

489   from the SciPy python package (42). Most parameter values and ranges for fitting were informed from

490   reported literature values or experimental measurements from this study. For parameters related to an

491   experimental output ([fAha] or rF) without a reported literature value, an initial best estimate was selected

492   to produce a single time-step change one order of magnitude lower than the maximum recorded

493   experimental value. These parameters were then allowed to fit within a range of 1.5 orders of magnitude

494   from the initial estimate. Parameters were fit with a least squares minimization algorithm from 'Lmfit', a

495   prebuilt python library (43). All best fit values and boundary conditions can be found in the parameter tables

496   (Supplements S3-5).

497   **Parameter Sensitivity Analysis and Model Validation**

498   To effectively sample the input parameter space, Latin hypercube sampling (LHS) was utilized to

499   generate unique parameter sets (n = 10000), sweeping each parameter value through a range defined by the

500   boundary constraints from literature (Tables S3-S5) as previously detailed (32, 33). Global sensitivity

501   analysis was performed on the LHS generated parameter sets using partial rank correlation coefficient

502   (PRCC) analysis. This analysis quantifies the sensitivity of an output variable on the variation in input

503   parameter values (32, 33). Here, PRCCs were determined for each of the 19 fitted parameters (Table S3)

504   and 12 static parameters (Table S4) in the biodistribution model, as well as for all 8 parameters in the protein

505   incorporation model (Table S5). PRCCs were used to characterize the influence of each parameter on the

506   sum of square errors (SSE), the optimization metric for non-linear regression. Simulations used to inform

507   PRCCs were performed on the Brown Supercomputing Community Cluster at Purdue University (44), with

508   each simulation run on a single node with dual 12-core Intel Xeon Gold "Sky Lake" CPUs @ 2.60 GHz

509   and 96 GB of memory.

510   The standard error of fitting was determined for the 19 fitted parameters (Table S3) in the

511   biodistribution model and for all 8 fitted parameters (Table S5) in the protein incorporation model. Standard

512   error values were determined from the covariance matrix during non-linear regression using the built-in

513   functionalities of the 'Lmfit' python library (43).

514   **Plasma Sample Preparation for Untargeted Metabolomic Analysis**

515   The plasma metabolome of non-injected control samples (n = 3 biological replicates) and samples

516   collected 24 h post Aha injection (n = 3 biological replicates) was extracted by adding methanol:

517   chloroform: water (1:1:1 v/v) to 80 μL of each plasma sample. Samples were vortexed briefly and

518   centrifuged at 8,000 × g for 5 min at RT. The upper layer was transferred into a new tube and vacuum-dried

519   overnight at RT. The dried fraction was reconstituted in 75 μL 5% ACN and 0.1% FA. Reconstituted

520   samples were sonicated for 5 min, centrifuged at 16,000 × g for 8 min at RT, and the supernatants were

521   transferred to HPLC autosampler vials.

28

Untargeted LC-MS Metabolomic Analysis

522

523   Separations were performed on an Agilent 1290 UPLC system (Agilent Technologies). The

524   metabolites were analyzed using a Waters Acquity HSS T3 column (1.8 μm, 2.1 × 100 mm), with a mobile

525   phase flow rate of 0.45 mL·min^{-1}, where the mobile phase A and B were 0.1% FA in double distilled water

526   and ACN at a 1:1 ratio, respectively. Initial conditions were 100:0 A:B, held for 1 minute, followed by a

527   linear gradient to 20:80 at 16 min, then 5:95 at 22.5 min. Column re-equilibration was performed by

528   returning to 100:0 A:B at 23.5 min and holding until 28.5 min.

529   The mass analysis was obtained using an Agilent 6545 Quadrupole Time of Flight (Q-TOF) MS with ESI

530   capillary voltage +3.2 kV, nitrogen gas temperature 325 °C, drying gas flow rate 8.0 L·min^{-1}, nebulizer gas

531   pressure 30 psig, fragmentor voltage 130 V, skimmer 45 V, and OCT RF 750 V. MS data scans (m/z 70-

532   1000) were collected using Agilent MassHunter Acquisition software (v.B.06). Mass accuracy was

533   improved by infusing Agilent Reference Mass Correction Solution (G1969-85001). MS/MS was performed

534   in a data-dependent acquisition mode on composite samples.

Metabolomic Data Statistical Analysis

535

536   Peak deconvolution and integration were performed using Agilent ProFinder (v.10.0).

537   Bioinformatic analyses were performed using Agilent Mass Profiler Professional (v.13.1).

538   Chromatographic peaks were aligned across all samples. Peak areas were normalized by log$_2$-

539   transformation and applying a 75% percentile shift. Metabolites were filtered out if present in only one

540   sample. Furthermore, only metabolites present in all 3 replicates of either the control or Aha injected

541   samples were included. Statistical analysis was performed using unpaired student's t-test. Metabolites with

542     $P < 0.05$ and fold change ≥ 2 were considered significant. Peak annotations were performed using the

543   METLIN metabolite database, with a mass error of less than 5 ppm. Identifications were aided by MS/MS

544   spectra comparisons. Principal component analysis (PCA), hierarchal clustering analysis (HCA) and

545   metabolic pathway analysis were performed using MetaboAnalyst v.5.0.

546   **Conclusions**

547   Here, we report for the first time the biodistribution kinetics of the widely used Met analog, Aha,

548   in murine tissues, as well as the associated relative rates of incorporation of Aha into protein via protein

549   synthesis and loss via metabolism and protein turnover. These results showed that liver and kidney have

550   faster protein synthesis and turnover rates compared to brain and skeletal muscle, which is consistent with

551   previous studies that utilized isotope labeling (29). We also demonstrated that subcutaneous injection

552   allows for observing maximum protein labeling in a relatively short time (~ 6 h), which enables studying

553   proteins with shorter half-lives, in contrast to the traditional method of introducing the ncAA in diet or

554   using isotope-labeled amino acids. To support these findings, we developed a mathematical framework that

555   described the distribution kinetics of Aha in murine tissues and its relation to the degree of protein labeling

556   and computed the relative rates of protein synthesis and turnover. We further validated this framework for

557   predictive modeling of Aha labeling against an experimental dataset including two different repeated

558   injection dosing paradigms to demonstrate its efficacy as a tool for future experimental design. Finally, we

559   investigated the impact of Aha administration on the plasma metabolome and demonstrated that Aha

560   incorporation into cellular proteins does not have adverse effects on the normal physiology of mice. This

561   observation further confirms previous results from our group that demonstrated that ncAAs do not affect

562   the gross behavior nor the physical appearance of treated mice.

563   **Data Availability**

564   All relevant data are within the manuscript and its Supporting Information files. Model and code files can

565   be found at our lab's GitHub repository

566   (https://github.itap.purdue.edu/TamaraKinzerursemGroup/ncAABiokinetics)

567

568 **Acknowledgements**

574 **Conflicts of Interest**

575   The authors declare that they have no conflicts of interest.

576

References

1.   Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proceedings of the National Academy of Sciences of the United States of America. 2006;103:9482-7.

2.   Saleh AM, Wilding KM, Calve S, Bundy BC, Kinzer-Ursem TL. Non-canonical amino acid labeling in proteomics and biotechnology. Journal of Biological Engineering. 2019;13:43.

3.   Yuet KP, Tirrell DA. Chemical tools for temporally and spatially resolved mass spectrometry-based proteomics2014.

4.   Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, et al. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci. 2010;13(7):897-905.

5.   Bagert JD, Van Kessel JC, Sweredoski MJ, Feng L, Hess S, Bassler BL, et al. Time-resolved proteomic analysis of quorum sensing in Vibrio harveyi. Chemical Science. 2016;7:1797-806.

6.   Van Elsland DM, Bos E, De Boer W, Overkleeft HS, Koster AJ, Van Kasteren SI. Detection of bioorthogonal groups by correlative light and electron microscopy allows imaging of degraded bacteria in phagocytes. Chemical Science. 2016;7:752-8.

7.   Mahdavi A, Szychowski J, Ngo JT, Sweredoski MJ, Graham RLJ, Hess S, et al. Identification of secreted bacterial proteins by noncanonical amino acid tagging. Proceedings of the National Academy of Sciences of the United States of America. 2014;111:433-8.

8.   Bagert JD, Xie YJ, Sweredoski MJ, Qi Y, Hess S, Schuman EM, et al. Quantitative, time-resolved proteomic analysis by combining bioorthogonal noncanonical amino acid tagging and pulsed stable isotope labeling by amino acids in cell culture. Molecular and Cellular Proteomics. 2014;13:1352-8.

9.   Cohen LD, Zuchman R, Sorokina O, Müller A, Dieterich DC, Armstrong JD, et al. Metabolic Turnover of Synaptic Proteins: Kinetics, Interdependencies and Implications for Synaptic Maintenance. PLoS ONE. 2013;8:e63191.

10.   Kenney JW, Genheden M, Moon KM, Wang X, Foster LJ, Proud CG. Eukaryotic elongation factor 2 kinase regulates the synthesis of microtubule-related proteins in neurons. Journal of Neurochemistry. 2016;136:276-84.

11.   Hinz FI, Dieterich DC, Tirrell DA, Schuman EM. Noncanonical amino acid labeling in vivo to visualize and affinity purify newly synthesized proteins in larval zebrafish. ACS Chemical Neuroscience. 2012;3:40-9.

12.   Shen W, Liu HH, Schiapparelli L, McClatchy D, He Hy, Yates JR, et al. Acute Synthesis of CPEB Is Required for Plasticity of Visual Avoidance Behavior in Xenopus. Cell Reports. 2014;6:737-47.

13.   McClatchy DB, Ma Y, Liu C, Stein BD, Martinez-Bartolome S, Vasquez D, et al. Pulsed Azidohomoalanine Labeling in Mammals (PALM) Detects Changes in Liver-Specific LKB1 Knockout Mice. J Proteome Res. 2015;14(11):4815-22.

14.   Calve S, Witten AJ, Ocken AR, Kinzer-Ursem TL. Incorporation of non-canonical amino acids into the developing murine proteome. Scientific Reports. 2016;6:1-7.

619   15.   Alvarez-Castelao B, Schanzenbächer CT, Hanus C, Glock C, Dieck St, Dörrbaum AR, et
620   al. Cell-type-specific metabolic labeling of nascent proteomes in vivo. Nature Biotechnology.
621   2017;35:1196-201.
622   16.   McClatchy DB, Ma Y, Liem DA, Ng DCM, Ping P, Yates JR. Quantitative temporal
623   analysis of protein dynamics in cardiac remodeling. Journal of Molecular and Cellular Cardiology.
624   2018;121:163-72.
625   17.   Ikeda S, Sugimoto M, Kume S. Importance of methionine metabolism in morula-to-
626   blastocyst transition in bovine preimplantation embryos. Journal of Reproduction and
627   Development. 2012;58:91-7.
628   18.   Kudo M, Ikeda S, Sugimoto M, Kume S. Methionine-dependent histone methylation at
629   developmentally important gene loci in mouse preimplantation embryos. Journal of Nutritional
630   Biochemistry. 2015;26:1664-9.
631   19.   Tang S, Fang Y, Huang G, Xu X, Padilla-Banks E, Fan W, et al. Methionine metabolism
632   is essential for SIRT 1-regulated mouse embryonic stem cell maintenance and embryonic
633   development The EMBO Journal. 2017;36:3175-93.
634   20.   Saleh AM, Jacobson KR, Kinzer-Ursem TL, Calve S. Dynamics of Non-Canonical Amino
635   Acid-Labeled Intra- and Extracellular Proteins in the Developing Mouse. Cell Mol Bioeng.
636   2019;12(5):495-509.
637   21.   Evans HT, Bodea LG, Gotz J. Cell-specific non-canonical amino acid labelling identifies
638   changes in the de novo proteome during memory formation. Elife. 2020;9.
639   22.   Kirman CR, Hays SM, Aylward LL, Suh M, Harris MA, Thompson CM, et al.
640   Physiologically based pharmacokinetic model for rats and mice orally exposed to chromium.
641   Chem Biol Interact. 2012;200(1):45-64.
642   23.   Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter
643   values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.
644   24.   Streif JU, Hiller KH, Waller C, Nahrendorf M, Wiesmann F, Bauer WR, et al. In vivo
645   assessment of absolute perfusion in the murine skeletal muscle with spin labeling MRI. Magnetic
646   resonance imaging. J Magn Reson Imaging. 2003;17(1):147-52.
647   25.   Boswell CA, Mundo EE, Ulufatu S, Bumbaca D, Cahaya HS, Majidy N, et al. Comparative
648   physiology of mice and rats: radiometric measurement of vascular parameters in rodent tissues.
649   Mol Pharm. 2014;11(5):1591-8.
650   26.   Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Incorporation of azides into recombinant
651   proteins for chemoselective modification by the Staudinger ligation. Proceedings of the National
652   Academy of Sciences of the United States of America. 2002;99(1):19-24.
653   27.   Eipel C, Abshagen K, Vollmar B. Regulation of hepatic blood flow: the hepatic arterial
654   buffer response revisited. World J Gastroenterol. 2010;16(48):6046-57.
655   28.   McClatchy DB, Dong MQ, Wu CC, Venable JD, Yates JR. 15N metabolic labeling of
656   mammalian tissue with slow protein turnover. Journal of Proteome Research. 2007;6:2005-10.
657   29.   Price JC, Guan S, Burlingame A, Prusiner SB, Ghaemmaghami S. Analysis of proteome
658   dynamics in the mouse brain. Proceedings of the National Academy of Sciences. 2010;107:14508-
659   13.

33

660 30.   Hammond DE, Claydon AJ, Simpson DM, Edward D, Stockley P, Hurst JL, et al. Proteome
661 Dynamics: Tissue Variation in the Kinetics of Proteostasis in Intact Animals. Mol Cell Proteomics.
662 2016;15(4):1204-19.
663 31.   Hyndman R, Athanasopoulos G. Forecasting: principles and practice. 2nd ed: OTexts;
664 2018. 384 p.
665 32.   Kinzer-Ursem TL, Linderman JJ. Both ligand- and cell-specific parameters control ligand
666 agonism in a kinetic model of g protein-coupled receptor signaling. PLoS Comput Biol.
667 2007;3(1):e6.
668 33.   Romano DR, Pharris MC, Patel NM, Kinzer-Ursem TL. Competitive tuning: Competition's
669 role in setting the frequency-dependence of Ca2+-dependent proteins. PLoS Comput Biol.
670 2017;13(11):e1005820.
671 34.   Chiang j. Liver Physiology: Metabolism and Detoxification. In: McManus L, Mitchell R,
672 editors. Pathobiology of Human Disease, Academic Press; 2014. p. 770-1782.

673 35.   Yang Z, Nakabayashi R, Okazaki Y, Mori T, Takamatsu S, Kitanaka S, et al. Toward better
674 annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites
675 from Oryza sativa (rice) by using MS/MS and NMR analyses. Metabolomics. 2014;10(4):543-55.
676 36.   Chaleckis R, Meister I, Zhang P, Wheelock CE. Challenges, progress and promises of
677 metabolite annotation for LC-MS-based metabolomics. Curr Opin Biotechnol. 2019;55:44-50.
678 37.   Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives
679 on genomes, pathways, diseases and drugs. Nucleic Acids Research. 2017;45:D353-D61.
680 38.   Steward KF, Eilers B, Tripet B, Fuchs A, Dorle M, Rawle R, et al. Metabolic Implications
681 of Using BioOrthogonal Non-Canonical Amino Acid Tagging (BONCAT) for Tracking Protein
682 Synthesis. Front Microbiol. 2020;11:197.
683 39.   Mentch SJ, Mehrmohamadi M, Huang L, Liu X, Gupta D, Mattocks D, et al. Histone
684 Methylation Dynamics and Gene Regulation Occur through the Sensing of One-Carbon
685 Metabolism. Cell Metab. 2015;22(5):861-73.
686 40.   Gao X, Sanderson SM, Dai Z, Reid MA, Cooper DE, Lu M, et al. Dietary methionine
687 influences therapy in mouse cancer models and alters human metabolism. Nature.
688 2019;572(7769):397-401.
689 41.   Olsen T, Ovrebo B, Haj-Yasein N, Lee S, Svendsen K, Hjorth M, et al. Effects of dietary
690 methionine and cysteine restriction on plasma biomarkers, serum fibroblast growth factor 21, and
691 adipose tissue gene expression in women with overweight or obesity: a double-blind randomized
692 controlled pilot study. J Transl Med. 2020;18(1):122.
693 42.   Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy
694 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-72.
695 43.   Newville M, Stensitzki T, Allen DB, Rawlik M, Ingargiola A, Nelson A. Lmfit: Non-
696 Linear Least-Square Minimization and Curve-Fitting for Python2016 June 01, 2016.
697 ascl:1606.014 p.
698 44.   McCartney G, Hacker T, Yang B. Empowering Faculty: A Campus Cyberinfrastructure
699 Strategy for Research Communities. Educause Review. 2014.

700

701   **Supporting Information Captions**

702   **S1 Text.** Code and model files.

703   **S2 Text.** Systems of ODEs for Aha model.

704   **S3 Table.** Distribution model fitting parameters.

705   **S4 Table.** Distribution model static parameters.

706   **S5 Table.** Incorporation model fitting parameters.

707   **S6 Figure.** Representative western blots informing Figures 2 and 3.

708   **S7 Figure.** Representative western blots informing Figure 6.

709   **S8 Figure.** Metabolomic analysis results following missing value imputation.

710   **S9 Table.** Selected metabolic pathways identified by Metaboanalyst analysis of the untargeted LC-MS

711   analysis of Aha metabolome.

712   **S10 Table.** Untargeted LC-MS Aha metabolomic analysis raw data.