Protein turnover models for LC–MS data of heavy water metabolic labeling

Rovshan G. Sadygov (iD)

Corresponding author: Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555, USA.
Tel.: (409)772-3287; E-mail: rgsadygo@utmb.edu

Abstract

Protein turnover is vital for cellular functioning and is often associated with the pathophysiology of a variety of diseases. Metabolic labeling with heavy water followed by liquid chromatography coupled to mass spectrometry is a powerful tool to study *in vivo* protein turnover in high throughput and large scale. Heavy water is a cost-effective and easy to use labeling agent. It labels all nonessential amino acids. Due to its toxicity in high concentrations (20% or higher), small enrichments (8% or smaller) of heavy water are used with most organisms. The low concentration results in incomplete labeling of peptides/proteins. Therefore, the data processing is more challenging and requires accurate quantification of labeled and unlabeled forms of a peptide from overlapping mass isotopomer distributions. The work describes the bioinformatics aspects of the analysis of heavy water labeled mass spectral data, available software tools and current challenges and opportunities.

Keywords: protein turnover, evolution of deuterium-enriched mass isotopomers, nonlinear models of time course data, rate constant estimation from metabolic labeling with heavy water followed by liquid chromatography – mass spectrometry (LC–MS)

Introduction

The cellular proteome is in dynamic equilibrium – proteins are continuously synthesized and degraded. The equilibrium is sustained by the proteostasis network, which maintains proteins in the appropriate abundance, folding state, concentration and location [1]. The network contributes to a cell's ability to respond to external and internal stimuli. Proteostasis is essential for healthy cellular functioning, and it is often associated with ageing [2] and disease states such as nonalcoholic fatty liver disease, neurodegenerative diseases [3] and cancers [4]. Protein turnover has been studied for decades using various techniques such as radioactive labeling [5], small molecule inhibitors of synthesis or degradation and fusion proteins tagged with constructs such as green fluorescent protein (GFP) [6].

Metabolic labeling with isotope tracers has historically been used to study protein turnover [7, 8]. Advancements in high resolution and accuracy mass spectrometry (MS) have turned proteomics into a powerful technique to study protein turnover in large scale [9]. Studies using cell cultures and stable isotopic labeling by amino acids in cell culture (SILAC) [10] labeling can produce turnover rates of several thousand proteins [11]. The proteome dynamics of several cell cultures have been studied using this approach [12–14]. The studies have greatly contributed to the wealth of information on proteome dynamics and its regulation. However, the inference of

in vivo protein turnover from that of cell cultures is not straightforward [15, 16]. As discussed below, heavy amino acids have also been used to label living organisms. Protein turnover studies using SILAC-based labeling have recently been reviewed [17].

Two types of stable isotope tracers have been often used for *in vivo* [18–21] protein turnover studies: heavy (nonradioactive) essential amino acids, such as $^{13}C_6$-Lys, $^{13}C_6$-Leu, $^2H_3$-Leu and atom-based tracers, such as ^{15}N and ^2H. For bioinformatics purposes, metabolic labeling with heavy amino acids has advantages. The labeling results in predictable shifts in masses of peptides. The isotope profiles of labeled and unlabeled peptides are well separated, thus readily allowing the determination of their relative abundances (RAs). The approach has been successfully applied in a number of *in vivo* studies analyzing protein turnover in zebra fish [22], *Caenorhabditis elegans* [23], *Drosophila* [24] and mice [25, 26]. The labeling with $^{13}C_6$-Leu has recently been used to study protein turnover in human ventricular cerebrospinal fluid of patients who suffered a subarachnoid hemorrhage [27]. In labeling with heavy amino acids, the quantification is only possible for peptides that contain the labeled amino acid. Often, organisms recycle the essential amino acids from the catabolism of proteins. A recent study developed a mathematical model that accounted for the reusage (from protein degradation) of an essential amino acid, Lys, in metabolic labeling with $^{13}C_6$-Lys diet

[25]. A longer time period may be required for reaching an equilibrium of labeling precursors (free Lys). In addition, labeling with the amino acid diet is expensive, and additional time is necessary to adapt to the diet.

The labeling with atom-based agents such as ^{2}H or ^{15}N often results in composite isotope profiles of labeled and unlabeled species. Therefore, the computational and mass informatics techniques for the data analyses from these labeling approaches are more complex. Stable isotope labeling of mammals uses a ^{15}N-enriched diet such as blue-green algae *Arthrospira platensis* (also referred to as spirulina) for labeling [19, 28, 29]. Protein turnover of the murine liver, blood and brain proteins has been analyzed [30] using this labeling method. The analysis of rat brain proteome turnover has identified extremely long-lived proteins [31, 32].

Labeling with deuterium in heavy water supplied in drinking water is cost-effective and easy to use. Heavy water in concentrations that lead to >20% enrichment (in body weight) is toxic for many species [33]. Until recently, the concentration of deuterium was chosen such that to result in 8–10% body water enrichment (BWE) [34]. These enrichment levels were required for the quantification of labeling in liquid chromatography and mass spectrometry (LC–MS). However, the recent advances in mass analyzers and bioinformatics data analysis techniques allow the detection and quantification of labeling resulting from only 2–3% BWE. Since heavy water labeling is also used in clinical settings, the development of techniques using lower deuterium concentrations and shorter labeling duration is ongoing. In heavy water labeling, deuterium is incorporated into C–H bonds of nonessential amino acids (NEAAs) [35], which are synthesized endogenously. Since essential amino acids are obtained from the diet, they are largely unlabeled. Therefore, to observe protein turnover in an isotope profile, a peptide that contains NEAAs should be chosen. Since several ubiquitous amino acids (such as Ala or Gly) are NEAAs, practically all peptides resulting from a tryptic digest can be used in quantification.

Recent reviews have discussed analytical [36], instrumentation [37] and practical [38] aspects of experiments using metabolic labeling followed by LC–MS to study *in vivo* protein turnover. This review focuses on recent developments of bioinformatics techniques and tools for the analysis of mass spectral data from heavy water metabolic labeling. It emphasizes techniques for time course modeling, peak detection and integration, variance estimation and settings for peptide identifications from MS/MS and protein sequence databases. Current bioinformatic and technical challenges are also discussed.

Throughout the text, protein turnover rate, degradation rate constant and fractional synthesis rates will be used interchangeably. Briefly, protein turnover rate is defined as the time necessary for both degradation and synthesis of half of the present protein pool. Under the steady state assumption (used in discussions here), protein turnover rate is determined from the degradation rate constant, k_{deg}. The fractional synthesis rate, which will be defined below, is obtained from the depletion of the RA of the monoisotope, normalized to the asymptotic enrichment.

Methods

Figure 1 shows a general workflow of labeling, data acquisition and analyses to extract protein degradation rate constants. After initial intraperitoneal (IP) injection with heavy water (the IP dose is determined from body-weight), living organisms are provided with deuterium-enriched water. The IP injection rapidly (within hours [39]) raises deuterium enrichment in the body water to about 3–4%. A continuous supply of drinking water enriched with deuterium maintains the BWE. The deuteriums are incorporated into amino acids and subsequently into proteins. At specific time points of labeling, tissue samples are collected. For a baseline comparison, tissues from unlabeled animals are also collected. At each time point, the tissues are processed using a routine MS-based proteomics workflow: protein extraction/separation, protein digestion, peptide separation, LC–MS–MS/MS and peptide/protein identification from tandem mass spectra and protein sequence databases [19]. BWE levels are determined from blood serum. As proteins are degraded and resynthesized, they incorporate deuterium-enriched NEAAs. The enrichment leads to a change of isotope envelopes of intact peptides in MS1. Traditionally, the time course of the label incorporation is modeled using the evolution (from unlabeled to plateau enrichment) of the isotope profiles of peptides [35, 39–41]. Alternatively, the deuterium incorporation can also be monitored in the isotope profiles of fragment ions [42, 43] or immonium ions [44].

Data set

Unless stated otherwise, this work uses a publicly available data set [45] obtained from murine liver labeled with heavy water. Normal diet-fed LDLR$^{-/-}$ (low-density lipoprotein receptor) mice were labeled with heavy water for up to 21 days. Liver tissues were collected at 0, 3, 7, 11, 15, 21 days of labeling. The time point selection was made based on empirical observations. Recently, we have developed a time point selection algorithm to achieve optimal sampling for proteins in a sample [46]. At each labeling time point, one mouse was sacrificed. Proteins were extracted and fractionated using SDS-PAGE. The proteins of Band 4 were used in this work. Tryptic peptides were separated using reversed-phase liquid chromatography and analyzed in Q Exactive™ Plus Mass Spectrometer (Thermo Scientific, CA). The mass profiles of intact peptides (MS1) were recorded in Orbitrap at the resolution of 70 000 (200 *m/z*). The peptides were fragmented in a data-dependent mode (DDA) using high-energy collision. The resulting fragment ions were recorded at a resolution of
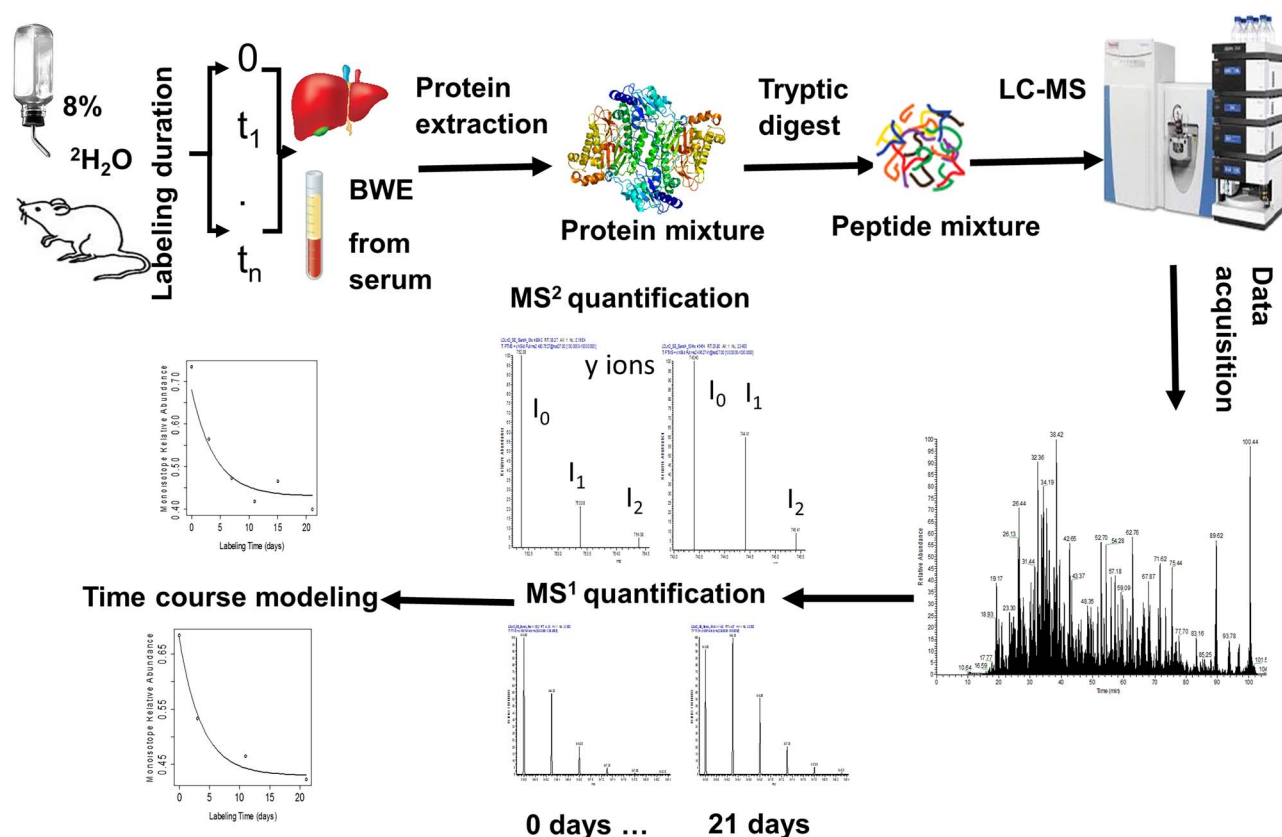
Figure 1. A workflow of rate constant estimation. After initial IP injection, mice are provided deuterium-enriched drinking water. At predefined durations of labeling, mice are sacrificed, tissues and blood are collected. BWE is determined from the enrichment of free amino acids in blood serum. Proteins are extracted, digested, and peptides are analyzed in LC–MS–MS/MS. MS/MS spectra are searched to identify peptides. Peptide sequences, elution times, m/z and charge state are used to detect and quantify isotope envelopes. The time course of label incorporation into peptides (MS1) or (less frequently) into fragment ions (MS/MS) is used to extract peptide/protein degradation rate constants.

15 000. Further details of the experimental data are in the original publication [45].

Database search for peptide/protein identification

Mascot [47] database search engine for peptide/protein identification from tandem mass spectra was used. The search parameters were as follows: the fixed modification of Cys carbamidomethylation and variable modifications of Met (oxidation) and Lys (acetylation); tryptic enzymatic specificity with two missed cleavages; the mass accuracies of precursors and fragments were 15 ppm and 40 ppm, respectively. Importantly, the precursor mass accuracy tolerance included up to two ^{13}C peaks. This setting specifies that a precursor chosen for fragmentation could potentially be from the monoisotope, first, or the second heavy mass isotopomers. As a result of labeling, heavy mass isotopomers become more abundant, and the monoisotope becomes depleted. In these cases, often, a ^{13}C peak (a heavy mass isotopomer) is chosen for fragmentation in a mass spectrometer in the DDA mode.

Other database search engines such as SEQUEST [48], ProLuCID [49] or MS-GF+ [50] may be used. The ability to search peptide sequences based on the precursor mass

of a heavy mass isotopomer (often referred to as ^{13}C peak) is important and present in these algorithms. It should, however, be noted that except for the precursor mass shift, the algorithms do not seem to account for fragmentation patterns (product ions) generated from heavy mass isotopomers of a peptide.

Extracting degradation rate constants from deuterium incorporation

This and the next sections describe theoretical techniques used to extract protein rate constants from LC–MS data of heavy water metabolic labeling. This section summarizes approaches that model the time course of the monoisotopic RA. As the proteins incorporate deuterium from the heavy water, monoisotopic RA depletes. The time course is modeled as an exponential decay to yield the decay (rate) constant. These approaches work with RAs, and they require complete isotope profiles. The next section addresses the time dynamics of ratios of abundances of mass isotopomers. Using a probabilistic framework of isotope distributions, time courses of several ratios can be obtained via closed formulas [51]. The results allow characterizing heavy water incorporation dynamics using partial isotope profiles.

The time course of mass isotopomer profiles of precursor ions as they incorporate the labeling agent was originally modeled in mass isotopomer distribution analysis (MIDA) [52]. It was suggested that the fractional synthesis, $f(t)$, can be described as:

$$f(t) = \frac{I_n(t) - I_n(0)}{I_n^{asymp} - I_n(0)} = 1 - e^{-kt} \tag{1}$$

where $I_n(t)$ is the RA of the nth mass isotopomer ($n = 0$ for the monoisotope) at time point t, I_n^{asymp} is the RA of the nth isotope at the plateau of enrichment. k is termed as a fractional synthesis rate and also as the degradation rate constant. Equation (1) is fit to the experimental data of RAs to determine the rate constant. A derivation of Equation (1) for the monoisotopic RA is provided in the Supplementary File.

In metabolic labeling with heavy water, there are a certain number of hydrogens (mostly in sidechains of NEAAs) for a peptide that are available for deuterium substitution (N_{EH}). The RA of the monoisotopic peak, $I_0(t)$, is expressed as [51]:

$$I_0(t) = I_0(0)\left(1 - \frac{p_X(t)}{1 - p_H}\right)^{N_{EH}} \tag{2}$$

p_H is the natural abundance of deuterium, $p_X(t)$ is the enrichment of the peptide with deuterium (in excess of the natural abundance), atom percent excess (APE) [53], at the labeling duration time t. When the labeling of a protein reaches its plateau, $(p_X(t) + p_H)$ becomes equal to the BWE, p_W. Hence:

$$I_0^{asympt} = I_0(0)\left(1 - p_W/\left(1 - p_H\right)\right)^{N_{EH}} \tag{3}$$

Equations (1) and (3) provide the time course equation for the monoisotopic RA:

$$I_0(t) = I_0^{asymp} + \left(I_0(0) - I_0^{asymp}\right)e^{-kt} \tag{4}$$

Equation (4) assumes that the label incorporation into NEAAs is faster than protein synthesis [54]. Equation (4) can also be obtained as a solution to a one-compartment kinetic model of labeling [55].

The experimental units (e.g. mice) are not expected to have the same BWE levels. The dependence on the BWE is in the asymptotic term, I_0^{asympt}. The data are transformed to the same asymptote. For example, if at labeling time point t, the observed RA is $I_0^1(t)$, and the corresponding BWE level results in asymptotic enrichment I_1^{asymp}, then normalization to a different BWE with an asymptotic enrichment level of I_2^{asymp} is:

$$I_0^2(t) = I_2^{asymp} + \left(I_0^1(t) - I_1^{asymp}\right)$$
$$\times \left(I_0(0) - I_2^{asymp}\right)/\left(I_0(0) - I_1^{asymp}\right)$$

As is seen from the formula, for $I_0^2(t)$, which replaces $I_0^1(t)$ in data analysis, the asymptotic enrichment is I_2^{asymp}. It is assumed that the degradation rate constants are the same in two subjects. The result is obtained using Equation (1).

The variance of the rate constant from Equation (4) is estimated using the Delta method [56]. If a homoscedastic variance, $\sigma_0^2 = Var(I_0(t))$ is assumed in measuring $I_0(t)$ at all-time points, then the variance of the estimated rate constant is [57]:

$$Var(k) = \sigma_0^2 / \sum_{i=1}^{N} t_i^2 \left(I_0(t_i) - I_0^{asymp}\right)^2 \tag{5}$$

σ_0^2 is related to the experimental error of measuring mass isotopomer abundances. It is estimated from the mean squared error (MSE) of the fit.

Another form of transformation of Equation (1) leads to a linear regression fit for the log of the fractional synthesis rate [58]:

$$kt = -\log\left(\frac{I_0(t) - I_0^{asymp}}{I_0(0) - I_0^{asymp}}\right)$$

The linear least squares solution of the above equation from a sample with N number of time points is:

$$k = \sum_{i=1}^{N} t_i \log\left(\frac{I_0(0) - I_0^{asymp}}{I_0(t_i) - I_0^{asymp}}\right) / \sum_{i=1}^{N} t_i^2 \tag{6}$$

The variance of the rate constant from the linear regression is given by Equation (5), as well. In the linear model, the variance of the measurements of $I_0(t_i)$, σ_0^2, is also determined from the MSE. The variance will determine the confidence interval of the estimated value. Both models (linear and nonlinear regression fits) are known to be sensitive to outliers. For model comparison, correlations, R^2 coefficient of determination and residual sum of squares are used.

A stochastic, data-driven model, Gaussian process for extracting peptide/protein turnover rates [59] has been developed. In this model, the RA is modeled as a multivariate Gaussian random variable with exponentially decaying mean, $\vec{\mu}$ and variance–covariance matrix, $\sum$:

$$\vec{I_0} \sim MVN\left(\vec{\mu}, \Sigma\right); \vec{I_0}, \vec{\mu} \in R^N;$$

where MVN stands for multivariate normal distribution, N is the number of data points (time points at which

heavy isotope levels have been measured). The components of $\vec{\mu}$ are determined from the following equation:

$$\mu\,(t_i) = I_0^{asymp} + \left(I_0(0) - I_0^{asymp}\right) e^{-kt_i}$$

The covariance matrix, $\sum$, is an N by N matrix defined as:

$$\sum (i,j) = K\,(i,j) + \sigma_\varepsilon^2 \delta\,(i,j)$$

where δ is the Kronecker's delta, σ_ε^2 is the variance of white Gaussian noise. $K(s,t)$ is the Ornstein–Uhlenbeck kernel:

$$K\,(i,j) = \sigma_\gamma^2 e^{\left(-|t_i - t_j|k\right)}/k$$

σ_γ^2 is the variance of the model fluctuations. The parameters of this empirical Bayes model are determined via maximization of the log-likelihood function, $\mathcal{L}(\vec{\theta})$, $\vec{\theta} = (k, \sigma_\gamma, \sigma_\varepsilon)$:

$$\log\left(\mathcal{L}\left(\vec{\theta}\right)\right) = -0.5 * \left\{\ln\left(\det\left(\Sigma\right)\right) - \left(\vec{y} - \vec{\mu}\right)^T\right.$$
$$\left. \times\, \Sigma^{-1} \left(\vec{y} - \vec{\mu}\right) - n * \ln\left(2\pi\right)\right\} \quad (7)$$

$\vec{y}$ denote the actual experimental data of RA of the monoisotope. This model accounts for the correlations between observations that are present in time course data (measurements at nearby time points are expected to be correlated more than those at the far time points). The rate constant in this model is obtained by the maximization of $\log(\mathcal{L}(\vec{\theta}))$.

Often, protein half-lives (τ) are used in the analysis. It is defined as:

$$\tau = \ln(2)/k$$

The variance of the half-life is obtained from that of the rate constant using Delta method:

$$\text{var}\,(\tau) = \text{var}\,(k)\,\ln^2(2)/k^4$$

Isotope patterns and deuterium incorporation

A recent study [51] has shown that protein turnover can be obtained from the abundances of only two mass isotopomers. The ratio of the abundances of the first heavy mass isotopomer, $A_1(t)$, to that of the monoisotope, $A_0(t)$, is:

$$\frac{A_1\,(t)}{A_0\,(t)} = N_{EH} \frac{p_X\,(t)}{(1 - p_H)\,(1 - p_H - p_X\,(t))} + \frac{A_1\,(0)}{A_0\,(0)} \quad (8)$$

The corresponding ratio of the abundances of the second heavy mass isotopomer, $A_2(t)$, and the monoisotope,

$A_0(t)$ is:

$$\frac{A_2\,(t)}{A_0\,(t)} = \frac{A_2\,(0)}{A_0\,(0)} - \frac{A_1\,(0)}{A_0\,(0)} \frac{p_H N_{EH}}{(1 - p_H)} + \left(\frac{p_H}{1 - p_H}\right)^2$$
$$\times \frac{N_{EH}\,(N_{EH} + 1)}{2} - \left(\frac{p_X\,(t) + p_H}{1 - p_H - p_X\,(t)}\right)^2 \frac{N_{EH}\,(N_{EH} + 1)}{2}$$
$$+ \frac{N_{EH}\,(p_X\,(t) + p_H)}{(1 - p_H - p_X\,(t))} \frac{A_1\,(t)}{A_0\,(t)} \quad (9)$$

The APE, $p_X(t)$, can be determined from the raw abundances of two mass isotopomers e.g. from Equation (8):

$$p_X\,(t) = \frac{(A_1\,(t)\,/A_0\,(t) - A_1\,(0)/A_0\,(0))\,(1 - p_H)^2}{N_{EH} + (1 - p_H)\,(A_1\,(t)\,/A_0\,(t) - A_1\,(0)/A_0\,(0))} \quad (10)$$

From $p_X(t)$, it is possible to reconstruct the time course using Equation (2), and then determine the rate constant, Equation (4). Equations (8) and (9) can also be used for the simultaneous determination of $p_X(t)$ and N_{EH}.

It should be noted that fast-Fourier transforms and binomial distributions are also used to estimate deuterium APE in hydrogen–deuterium exchange experiments to study protein structure [60–62]. However, in these experiments, enrichment in deuterium is high, and they result in separated isotope profiles. The techniques presented in this section apply to the cases, when on average less than one hydrogen in a peptide may be exchanged with deuterium.

Results and discussions

Proteins/peptides are identified from tandem mass spectra LC–MS experiments using database search engines. For time course modeling, proteins that have been observed in at least four different time points of labeling are recommended [63]. Peptide sampling in DDA is semi-stochastic. In replicate experiments, peptides are chosen for fragmentation in some but not in all experiments. In addition, in heavy water metabolic labeling, it is observed that the number of identified proteins/peptides declines with the increased duration of labeling. This is shown in Figure 2, which depicts the total number of spectra and those peptide-spectrum matches that passed the homology and identity thresholds. The results are from Mascot [47]. Shown are the averages from two technical replicates at each labeling time. As seen from the figure, the number of identified peptides steadily declines with the increased labeling duration. This result has consistently been observed in other studies. The reasons for this phenomenon are unclear. It does not seem to be explainable only by the higher abundances of the heavy mass isotopomers after the labeling.

Peptide identifications that pass the false discovery rate (FDR) [64] thresholds are used for quantification. Each peptide identification item provides peptide sequence, identification score(s), elution time, charge
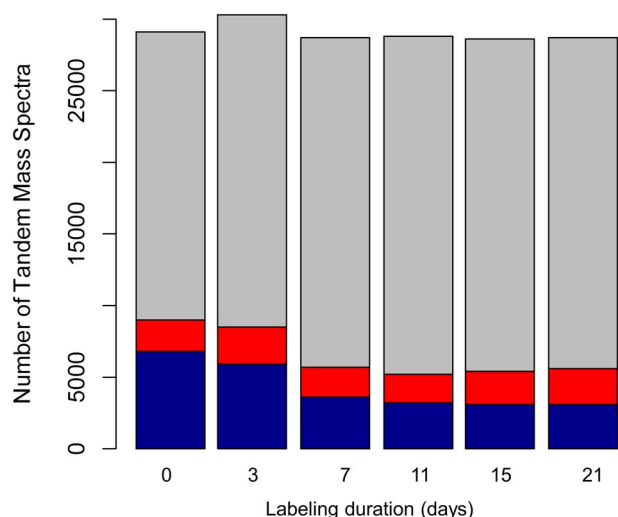
Figure 2. The number of PSMs passing the FDR threshold tends to decrease with the increased incorporation of deuterium. Shown are the statistics from Mascot for six labeling durations with the average of two technical replicates at each duration. The blue and red columns are the (averaged number of) PSMs that passed Mascot's identity and homology thresholds, respectively. The gray columns are the number of unidentified spectra.

state and peptide mass-to-charge ratio, m/z. Targeted feature detection is performed using the m/z, elution time, charge state and peptide sequence. Figure 3 shows the isotope envelopes (at the apex of the elution profile) of LNEINEK, a peptide of mouse CPSM (carbamoyl-phosphate synthase [ammonia], mitochondrial) protein, at four different labeling durations. As the peptide incorporates deuterium, its isotope envelope changes. The time course of the RA of monoisotope of this peptide is used in the exponential decay model to extract the degradation rate constant, Equation (4). The time course of the RAs of this peptide is shown in Figure 4 (empty black circles).

The dependence of the label incorporation on the essential amino acids is modeled using a concept of a number of hydrogens available for deuterium substitution in metabolic labeling with heavy water, N_{EH}. The number is assumed to be the sum of numbers of hydrogens sites that are accessible to the deuterium for each amino acid. The N_{EH} numbers of amino acids have been experimentally determined only for mouse [65]. The same numbers have been used in the analysis of data from other species e.g. in the analysis of human plasma protein turnover [63]. In general, it can be expected that the N_{EH} numbers will differ between organisms and possibly for the same organism but under different conditions/environments. The recent developments [51] allow the simultaneous determination of relative labeling and N_{EH} numbers from raw abundances of three mass isotopomers of a peptide, Equations (8) and (9).

Table 1 shows the rate constant estimations using three approaches aforementioned for 10 CPSM peptides, each of which has 13 hydrogens available for deuterium substitution ($N_{EH} = 13$). As seen from the table, the results

from various rate constant estimation techniques are close. However, it should be noted that the linear model will have invalid results for some data points. It happens when the measured monoisotopic RA is less than the asymptotic RA, which is determined theoretically. These data points occasionally occur, and they are omitted in the rate constant calculations in the linear model. The nonlinear models are more flexible, and these data points are included. The linear model has been employed in experiments that estimated the rate constants from one metabolic labeling point [66].

As seen from Table 1, the rate constants for peptides exhibit variability. For all but one (LTSIDKWFLYKmR, which exhibited low abundance in MS1) of the peptides shown in the table and Figure 4, isotope envelopes are well isolated and without interferences from the co-eluting contaminants. The goodness-of-fit parameters (such as the Pearson correlation and MSE) of the time course model [Equation (4)] are high. Two potential reasons for this outcome are the spectral accuracy (SA) of mass analyzers and the accuracy of the applied N_{EH} parameter. The SA (the accuracy in the determination of RAs of mass isotopomers [67, 68]) of modern mass analyzers is affected by various factors. Thus, in Orbitrap mass analyzers, trap overfilling (space charge effect) may lead to systematic errors in spectral patterns [69]. It has also been reported that the SA is affected by ion losses that occurred during the long scan times required for high mass resolution [67]. The RAs of mass isotopomers exhibit variations during the chromatographic elution. The RAs measured from samples of unlabeled experimental units often disagree with corresponding theoretical values. The insufficient accuracy of SA contributes to the variability in rate constant estimations. The N_{EH} parameter is another factor potentially contributing to the variabilities in the rate constant estimations. It has been experimentally determined only for mouse.

A potential source of variability in peptide degradation rate constants may be the amino acid composition and differences in the biogenesis of NEAAs. NEAAs can be grouped into families based on the first amino acid synthesized in each family. Correlations between the number of amino acids from each family and the rate constants were small and mostly positive, Supplementary Table S2. Thus, data from more than 950 proteins and 13 000 peptides (LDLR$^{-/-}$ mouse strain) showed that the highest median value (from proteins), 0.23, of correlations was observed for peptides that contained at least one of Glu or Gln amino acids. Thus, peptides that contained the mentioned amino acids tended to have higher rate constants than the median (protein) rate. The results for other combinations of amino acids are presented in Supplementary Table S2. The absolute values of the median correlations were small.

Although the recent improvements in increased mass accuracy and resolution have improved peptide
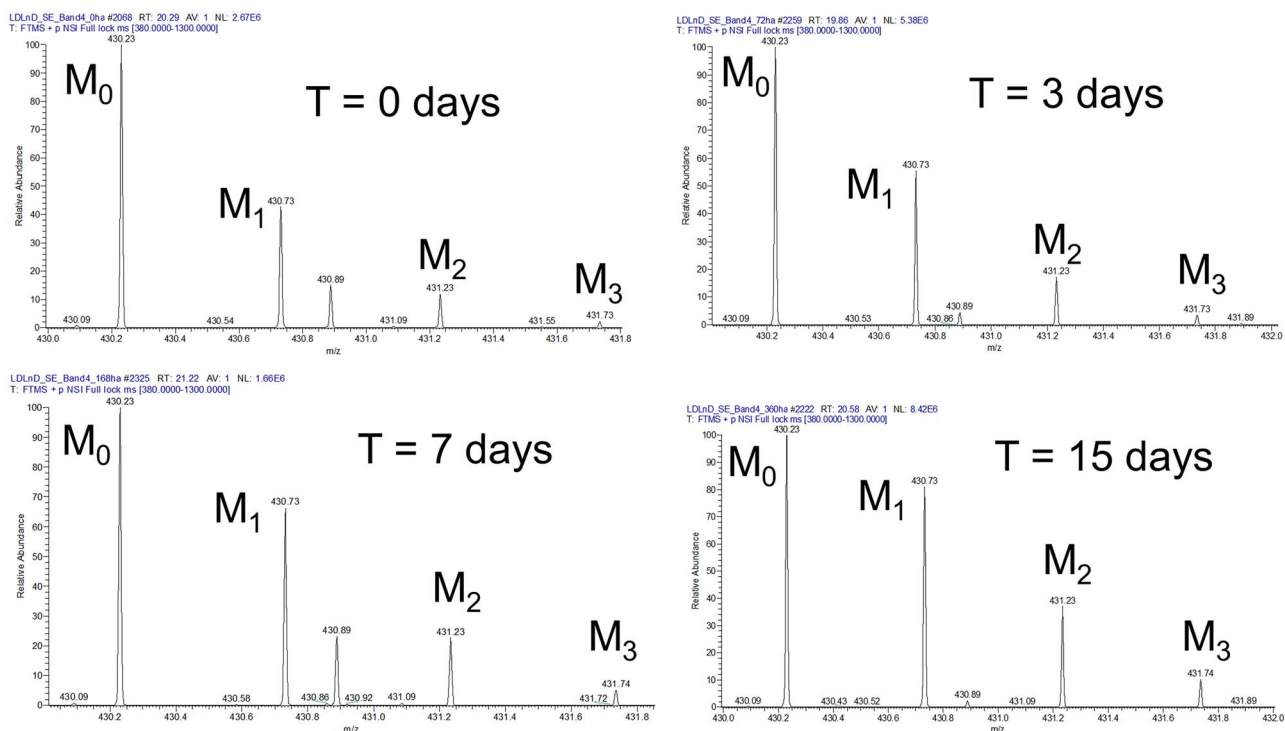
Figure 3. Deuterium incorporation during metabolic heavy water labeling leads to the increase in RAs of heavy mass isotopomers. Shown (from MS1 spectra) are the abundances of the first four mass isotopomers (M_0 through M_3) of CPSM peptide, LNEINEK, at three time points of labeling, 3, 7 and 15 days and the natural isotope distribution (time points are shown on the figure). An increase in the RAs of the heavy mass isotopomers can be seen from the figure. M_n refers to the nth mass isotopomer.

**Table 1.** Linear and nonlinear least squares fits produce close results of rate constants

Peptide	Equation (4), rate constant/SD	Equation (6), rate constant/SD	Equation (7), rate constant/SD	Figure 4 symbols*
LNEINEK	0.095/0.008	0.088/0.01	0.096/ 0.025	○
YMESDGIK	0.134/0.02	0.150/0.026	0.136/ 0.039	×
RTSVNVVR	0.066/0.012	0.059/0.014	0.071/ 0.019	△
YmESDGIK	0.136/0.017	0.118/0.025	0.139/ 0.035	+
MRDILNMDK	0.074/0.006	0.072/0.006	0.075/ 0.023	✳
mRDILNMDK	0.11/0.007	0.107/0.007	0.112/ 0.017	▽
mRDILNMDK^{+3}	0.111/0.003	0.108/0.004	0.113/ 0.006	★
mRDILNmDK	0.117/0.012	0.128/0.016	0.119/ 0.028	◇
TLGVDFIDVATK	0.173/0.032	0.049/0.051	0.170/ 0.028	⊠
LTSIDKWFLYKmR	0.146/0.175	0.057/0.187	0.042/ 0.848	⊕

Shown are the amino acid sequences, model fit results from 10 peptides of murine CPSM protein. All peptides have 13 hydrogens available for deuterium substitution in metabolic labeling. Shown are the rate constants in the units of day^{-1}. Peptide^{+3} – the peptide is +3 charged, the rest of the peptides are +2 charged. m denotes oxidized methionine. *Symbols used to designate the experimental data for peptides in Figure 4. The peptides are listed in the order of the RAs of their monoisotopes from Figure 4.

identifications, the increased resolution has not improved the SA in MS1. The high mass resolution separates mass profiles of the co-eluting species with close m/z values. Therefore, the high mass resolution reduces the ambiguity in assigning peaks. However, for quantitative techniques using metabolic stable isotope labeling, an important quantity is the SA. It has been reported that the increase in mass resolution of popular mass analyzer, Orbitrap, leads to a decrease in SA [70, 71]. The reduction in SA results from the increased transient times in which the lower abundance signals decay faster than the high abundance signals [71]. Recent analyses of heavy water labeling data from two different bioinformatics tools revealed that the accuracy of rate constant estimation

with the increased mass resolution consistently reduces after the resolution of 60 000–70 000 [70, 72]. In addition, as a result of the longer scanning times, the number of identified peptides, the first step in the data analysis, also decreases. These factors should be considered in setting parameters of mass analyzers in experiments for measuring protein turnover rates.

The biological variability is inherent in the rate constant estimations as animals are sacrificed at each labeling time point to collect tissues (Figure 1). However, the biological variability does not seem to create systematics shifts in rate constants. Figure S1 of the Supplementary file shows this. For a protein, it was assumed there is a distribution (Normal) of rate
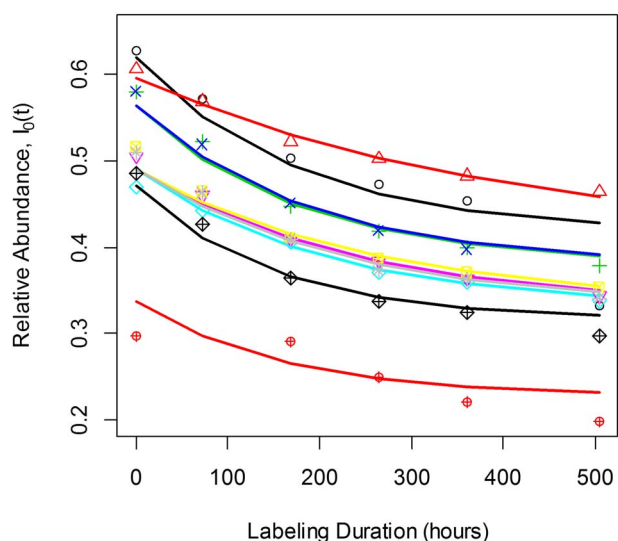
Figure 4. Time courses of label incorporation for well-quantified peptides are similar. Shown are the time courses of the monoisotopic RAs for 10 peptides (Table 1) of protein CPSM. For all peptides, the number of hydrogens available for deuterium substitution in metabolic labeling with heavy water is 13. Solid lines are theoretical fits. The experimental data points for each peptide are denoted by symbols designated in Table 1. The crossing of fitting curves was caused by the inconsistency in the estimation of monoisotopic RA of unlabeled RTSVNVVR (data points marked with red triangles).

constants for biological units. For each time point of labeling, the Normal distribution was sampled for a rate constant. The rate constant was used to generate $I_0(t)$ for the labeling time point t. The simulated time course data were used to 'estimate' the rate constant. The rate constant was then compared with the mean of the Normal distribution. The standard deviation (SD) of the normal distribution was equal to 25% of the mean. As Supplementary Figure S1 shows, there was no systematic bias between the computed and 'true' rate constants.

In addition to the variability in protein turnover rates between biological units, potentially technical variability may be introduced because of the different BWE levels. Differences in BWEs can be theoretically modeled as was shown above in the formula for $I_0^2(t)$. Normally, BWEs between different biological units are close. Large differences in BWEs may affect rate constant estimations, as small values of BWE result in subtle changes, especially for slow turnover proteins at shorter labeling durations. The quantification of these changes will be more affected by the SA of mass analyzers.

Bioinformatics tools for analyses of heavy water labeling data

Currently, there are three freely available bioinformatics tools for estimating proteome dynamics from metabolic labeling with heavy water and LC–MS, Table 2. ProTurn [73] is a Java-based software. It also uses R environment for statistical computing [74]. It is available online. DeuteRater [72] is developed in Python and is freely available via GitHub. It uses combinatorial distribution

analysis originally developed in MIDA [52]. The latter is proprietary software [75].

Our lab has developed (in C/C++ language) a freely available (via GitHub) algorithm, d2ome [45]. For the peptides/proteins that pass the user-defined stringency thresholds, the algorithm performs a mass accuracy-based peak detection and integration using the chromatographic elution profiles of peptides [76]. Theoretically, computed isotope distributions [77] are used to model incorporation at the initial and asymptotic time points.

A recent report [78] used a Dirichlet distribution to model a composite spectrum of labeled and unlabeled peptides. The mixing coefficient was determined using a Bayesian approach. In this approach, a peptide's isotope distribution is assumed to be a mixture of unlabeled and fully labeled (obtained by convolving isotope distributions of free amino acids) peptides. The coefficient is used to compute the ratio of the heavy to light peptides. The isotope distributions of labeled amino acids were determined from the free amino acids in blood. The model was implemented in the environment of Stan probabilistic programming language [79].

The algorithms can be viewed as assembled from three modules: data preprocessing, peak detection and integration and rate constant estimation. ProTurn and d2ome use the standard input files (mzid and mzML) to read sequence, m/z, charge, elution times of proteins/peptides that pass identification thresholds such as protein/peptide FDRs. DeuteRater uses a user-defined input to specify protein, peptide sequence, precursor m/z and elution time window. Using a statistical control (FDR) ensures that only peptides/proteins that are confidently identified will be processed for protein turnover in high throughput experiments. On the other hand, user-selected input of peptides/proteins provides more flexibility for the analyses of chosen peptides/proteins.

As seen below, an important element of the algorithms is peak detection and integration. IsotoQuant [39] module of ProTurn extracts ion chromatograms of peptides using retention time and mass isolation window of ±100 ppm. d2ome detects peaks based on the monoisotopic mass and the isotope profile. It determines the apex of the elution profile of the monoisotope. Peaks that are larger than 20% of the apex abundance are retained. DeuteRater uses a vector-based approach for isotope profile representation. It starts with the user input elution scan. It retains scans that differ from the neighboring scan by no more than $\log_{10}(1.2)$. The angle between the vectors (from two consecutive scans) is defined as the differences. DeuteRaters integrate the peaks in the user-defined elution window. Both DeuteRater and d2ome allow the user to set mass accuracy for peak detection. The option is useful as the mass accuracy of the mass analyzers continues to improve. d2ome computes and reports the statistics on the mass accuracy distribution of identified peptides. The information may guide to setting the mass accuracy for each experiment.

Table 2. Software tools for rate constant estimation using LC–MS data of metabolic labeling with heavy water

Software	Availability	Programing language	Input data
MIDA [51]	Proprietary	–	–
ProTurn [79]	http://heartproteome.org/proturn	Java and R	mzid and mzml
DeuteRater [59]	GitHub	Python	user chosen peptide list and mzML
d2ome [44]	GitHub	Visual Studio C/C++	mzid and mzML

All three algorithms use nonlinear least squares [80] to fit the label incorporation to an exponential function. It has been reported that the choice of the optimization algorithm does not affect the results of rate constant extraction [73]. ProTurn and DeuteRater generate plots of experimental data and the corresponding theoretical fit. ProTurn and d2ome determine the median of rate constants of peptides as a protein turnover rate. DeuteRater uses the combined time courses of the fractional synthesis of all peptides of a protein to determine protein turnover rate. The computation of the rate constant of each peptide explicitly reveals the variability in the protein turnover rate. Using a combined fractional synthesis of all peptides at each time point of labeling may potentially increase the number of quantified proteins when peptides are identified in some but not in all time points. DeuteRater can also use the mass shifts instead of RAs.

Observation of label incorporation into fragment ions in MS/MS

As aforementioned, most current techniques for determining protein turnover use the isotope profiles of intact peptides in MS1. However, it is also possible to monitor label incorporation into fragment ions in MS/MS. It has been demonstrated [42, 43] that label incorporation into a single amino acid (Arg) can qualitatively be monitored as a function of the labeling duration. Another study [44] showed that it is possible to quantitatively estimate protein turnover rate from the labeling dynamics of an immonium ion of Pro. Figure 5 shows the violin plot of RAs of the monoisotope of y_6 ion of peptide YMESDGIK. The distributions are from RAs of multiple fragmentation events of the peptide during its chromatographic elution. The RAs exhibit considerable fluctuations. Applying the current quantitative modeling techniques to extracting rate constants from fragment ions produced results that underestimated those from the isotope distribution of intact peptides [42]. There are potential advantages of signal estimation in MS/MS. The technique obviates the need for peak detection and quantification in MS1; therefore, interferences from co-eluting species are negligible. It also avoids effects from space charging.

Techniques using amino acid-based stable isotope tracers have employed fragment ion-based quantification for rate constants [3, 81, 82]. The quantification utilizes labeled and unlabeled fragment ions from different tandem mass spectra to estimate the rate constant. Recently, selected reaction monitoring has
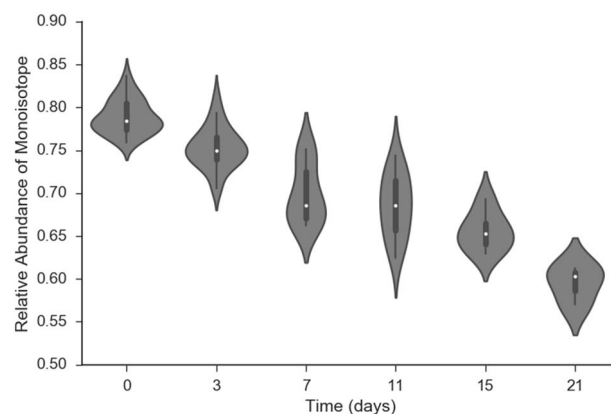
Figure 5. The monoisotopic RA of a fragment ion shows a characteristic depletion of metabolic labeling with heavy water. Violin plot of the RA of the monoisotopic peak of y_6 fragment ion of the peptide, YMESDGIK, at six time points of labeling. The white circles are the median values of RAs at specific time points.

been applied to study protein turnover by quantifying fragment ions. [83–85].

The number of peptide entries reduces after the goodness-of-fit filtering

Despite ubiquitous deuterium incorporation into peptides, only a fraction of the available peptide entries is useful for rate constant extraction [73]. Only 43% of peptides were retained after the filtering based on the goodness-of-fit metric for mouse strain, C57BL/6JRj strain (control group) [73]. The results are similar to those that were observed in another study [45] in a smaller data set. From the experience of data analyses, one of the main reasons for the poor goodness-of-fit metrics is the co-elution of the target peptides with the contaminants. Bioinformatic approaches for rate constant estimations from partial isotope profiles have the potential to improve rate constant estimations by obviating the necessity to quantify the complete isotope profile [43, 51]. Theoretically, it is possible to estimate the label incorporation from RAs of two mass isotopomers [51].

Detection of RA changes

Accurate estimation of the rate constant from mass spectral data requires adequate sampling of time points on the axis of the metabolic labeling duration. The magnitude of changes in isotope profiles of a peptide is theoretically determined by four factors: BWE, N_{EH}, protein turnover rate and labeling duration. The higher the rate, the earlier (the metabolic labeling time) are the isotope profile changes. Also, high levels of BWE lead to faster
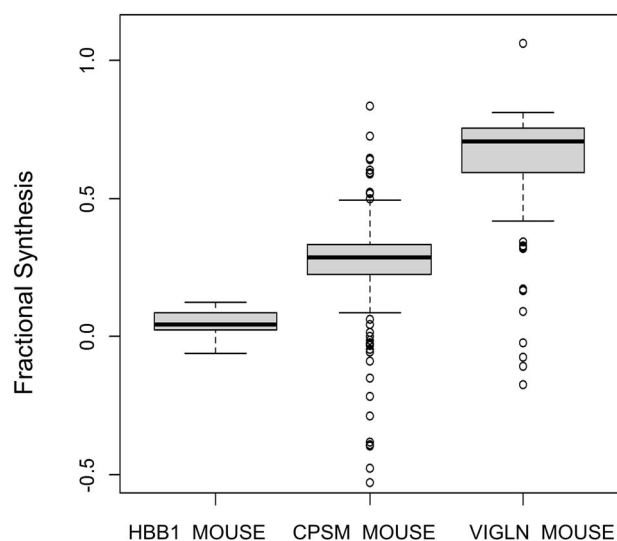
Figure 6. Three days of labeling result in measurable changes of isotope profiles. Boxplots of fractional syntheses of three proteins after three days of labeling are shown. Shown are the fractional syntheses for several peptides of each protein. For the same labeling duration, proteins with faster degradation rate constants (shorter half-lives) such as Vigilin (80 unique peptides) exhibit substantially higher fractional synthesis and measurable (by MS) changes in the isotopic profiles. For peptides of hemoglobin subunit beta-1, HBB1 (20 unique peptides), small changes are observed for isotope distributions of most of its peptides. Most of peptides of CPSM (324 unique peptides) also exhibit measurable fractional synthesis rates.

changes in the isotope profiles of peptides. However, mass spectrometers have their own specifics. As was noted, the SA of modern mass spectrometers does not allow accurate measurements of small changes in isotope profiles. In addition, the signal-to-noise ratio of MS is often dependent on the protein/peptide abundance. It is shown with the example of three proteins in Figure 6. The figure shows the boxplots of the fractional synthesis of three proteins: CPSM, high-density lipoprotein binding protein (VIGLN) and hemoglobin subunit beta-1 (HBB1). As seen from the figure, although the first two proteins show large fractional syntheses after three days of labeling, the last protein incorporates small amounts of deuterium. Using Equation (10), the estimate of the APE for CPSM peptides is 0.0039 after three days of labeling ($p_W = 0.0302$). The RA of ^{2}H in nature is approximately 1.5 $\times$ 10^{-4}. Thus, with the current SA of mass spectrometers, it is possible to estimate deuterium incorporations about 25 times that of the natural abundance. The calculated half-lives of the above proteins are (7.1 ± 2.1), (1.9 ± 0.5) and (48 ± 11) days, respectively.

Results from murine heart proteome studies using other labeling techniques and software

This section presents the results of murine proteome turnover studies using two different labeling techniques and analysis tools. A recent study [25] analyzed mouse (C57BL/6JRj strain) heart proteome turnover using ^{13}C$_6$-Lys labeling. Previously, murine heart proteome from this mouse strain was analyzed using heavy water labeling and ProTurn [73]. The data were also analyzed with

d2ome. Supplementary Figure S2 shows the scatter plot of protein turnover rates from d2ome and ProTurn. Pearson's and Spearman's correlations between the rates from the two methods for proteins quantified by three or more peptides (in both tools) were 0.73 and 0.93, respectively. Data points for these proteins are shown in blue in Supplementary Figure S2. Among these proteins, the biggest relative difference was observed for long-chain fatty acid transport protein 1 (Uniprot identification, Q60714, Uniprot entry name S27A1_MOUSE). The computed rates were 0.054 (d2ome) and 0.234 (Pro-Turn). The basic element of quantification is the determination of monoisotopic RA. Supplementary Table S1 shows the quantitative information for peptides of the protein S27A1_MOUSE from two software tools. It is noted that due to different peptide identification engines, the number of identified peptides was larger in ProTurn (the study used ProLuCID [49]). The table only shows the peptides that have been quantified in both tools. The time course of monoisotopic RAs of one peptide from the two bioinformatics tools and their fits are shown in Supplementary Figure S3. The figures show that the main reason for the different rates is the peak quantification in MS1. It should be noted that in d2ome, the high variability of peptide rate constants results in a high SD of protein rate constant. Thus, the SD of the rate constant of S27A1_MOUSE was 0.044, which resulted in a high coefficient of variation (CV), 82%. The corresponding CV from ProTurn was high as well and equal to 70%.

Three proteins (from 7 interacting proteins with scores above 0.7 in STRING [86]) that strongly interact with S27A1_MOUSE were also quantified in the sample. The interacting proteins are shown in Supplementary Figure S4. CPT1B_MOUSE (Q924X2), ACSL1_MOUSE (P41216) and CD36_MOUSE (Q08857) had similar turnover rates computed in d2ome and ProTurn. The computed turnover rates (in units of day^{-1}) for these proteins were (0.047; 0.055), (0.036; 0.046) and (0.147; 0.128), respectively. Turnover rates from d2ome (the first number) and ProTurn (the second number) are shown in parenthesis for each protein. Since the rate constants between proteins differed, it was not possible to hypothesize about a possible rate constant of S27A1_MOUSE based on those of its interacting proteins.

Protein turnover of heart proteome of C57BL/6JRj mouse strain using ^{13}C$_6$-Lys metabolic labeling has quantified 1164 proteins [25]. The scatter plot of turnover rates from heavy water labeling [73], and ^{13}C$_6$-Lys labeling [25] is shown in Figure 7. Pearson's and Spearman's correlations between the turnover rates were 0.64 and 0.90, respectively. The turnover rates obtained with ^{13}C$_6$-Lys labeling tended to be higher in general. The medians of computed half-lives were 9.3 and 4.4 days for heavy water labeling and ^{13}C$_6$-Lys labeling, respectively. The medians of half-lives of proteins (719 proteins) that were common from quantification with heavy water labeling (estimated using d2ome) and ^{13}C$_6$-Lys labeling were 9.3 and 5.4 days, respectively.
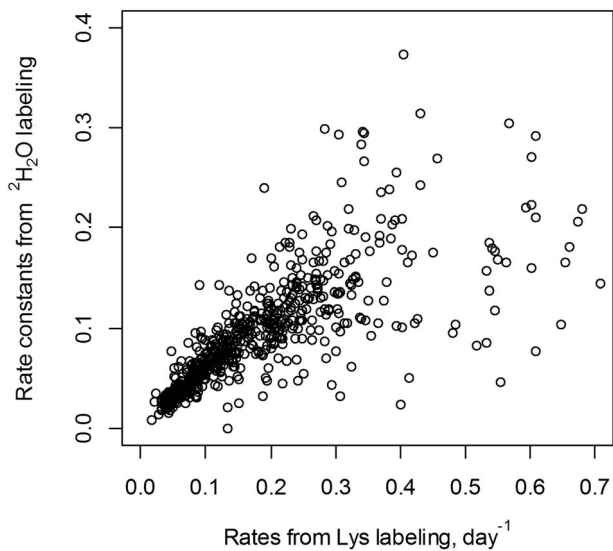
Figure 7. Turnover rates of murine (C57BL/6JRj strain) heart proteins measured using $^{13}C_6$-Lys labeling were, in general, higher than those computed using heavy water labeling. The *x*- and *y*-axes show the turnover rates obtained using heavy water labeling (computed using d2ome) and $^{13}C_6$-Lys labeling, respectively.

The median of the half-lives in the original publication of heavy water labeling [73] (computed using ProTurn) for this mouse strain in the control sample was 5.8 days. The scatter plot of rate constants between ProTurn and $^{13}C_6$-Lys labeling results is shown in Supplementary Figure S5. As seen from the plot, the rate constants computed from $^{13}C_6$-Lys labeling were higher in general. The median half-lives of proteins common (913 proteins) to heavy water labeling (ProTurn) and $^{13}C_6$-Lys labeling were 6.6 and 4.6 days, respectively. It is noted here that a study [87] of murine heart proteome (C57BL/6 strain) using metabolic $^2H_3$-Leu labeling reported median half-live of 9.1 days (5-month-old female mice). The comparisons of turnover rates in the above-mentioned experiments are confounded by biological (such as age and gender) and environmental (housing and diet) factors.

Perspectives

Heavy water is inexpensive and easy to use labeling agent to study protein turnover *in vivo*. Dynamics of thousands of proteins of liver [39, 45, 88], heart [34, 73, 89], skeletal muscle [2, 58] and blood [90] has been reported. Heavy water is toxic in high concentrations (>20%); therefore, low concentrations (2–3% BWE) are often used. Deuterium from heavy water is incorporated into NEAAs and labels virtually all peptides/proteins. Due to the low enrichment, the labeling generates composite isotope profile – an overlap of labeled and unlabeled peptides. This is different from labeling with heavy amino acids, which result in well-separated isotope profiles. Therefore, SA of mass analyzers is particularly important for accurate measurements of deuterium incorporation. Bioinformatics techniques [43, 51] have been developed to accurately model the enrichment. Accurate

estimation of label incorporation from semi-quantitative isotope profiles will increase the output of protein turnover studies.

Bioinformatics methods that estimate rate constants from two or three mass isotopomers will reduce the need to detect and quantify mass isotopomers from the entire isotope envelope. Furthermore, the developments of techniques that can estimate the number of hydrogens available for deuterium substitution on the fly will make it possible to apply the technology to a broader spectrum of species rigorously. One potential approach for this problem is to use the systems of two equations for two variables (deuterium enrichment and NEH) that are provided from the time courses of abundances of heavy mass isotopomers [51].

The number of identified proteins/peptides decreases with the increased labeling duration, even when the possible fragmentations from heavy mass isotopomers are included in database search engines. Sample processing and bioinformatics techniques that provide relatively consistent protein/peptide identification numbers from heavy water labeled samples will increase the proteome coverage.

Protein turnover rates obtained with applications of heavy water labeling and LC–MS generally agree with typical values in literature. However, it has been suggested that the precursor pool of amino acids may be diluted [35] due to the amino acid recycling. The dilution of labeled amino acids by the amino acids of catabolized tissue proteins has been modeled for experiments using stable isotope-labeled essential amino acid, Lys [25, 91, 92]. In the case of the NEAAs, the effect has not been studied. It cannot be ruled out that the recycling contributes to the variability observed in estimations of rate constants of peptides from the same protein.

Rate constant estimations from the fragment ions provide information about label incorporation into individual amino acids. It potentially provides data for studying specific metabolic pathways. Technique potentially avoids the contaminations from the co-eluting species and charge spacing.

Only a portion of the identified peptides currently produces results that pass the thresholds of the goodness-of-fit of time course models. It is similar to techniques that use heavy amino acids for labeling [87, 91]. Enhancements in SA, protein/peptide separation techniques and bioinformatics will improve the protein turnover estimations.

Conclusion

This review discussed bioinformatics aspects of the data processing for the estimation of protein degradation rate constants. Three modeling techniques have been compared. The nonlinear least squares model is more flexible for fitting data than the linear model. For the interpretation of the data, it was emphasized that the SA of the current mass spectrometers is still improving.

New bioinformatics techniques are needed to improve the efficiency of peptide data usage. Currently, only a portion of the identified peptides produces quantification results that pass the threshold of data analysis. New technical developments that allow rate constant estimation from partial isotope profiles will be helpful in the analyses of overlapping spectral profiles resulting from co-elutions.

Key Points

- Heavy water metabolic labeling followed by LC–MS is an easy to use, cost-effective, safe and powerful tool for high-throughput and large-scale studies of *in vivo* protein turnover.
- Incomplete labeling with heavy water results in a composite spectrum of labeled and unlabeled forms of a peptide. Accurate formulas provide the time evolution of mass isotopomer abundances.
- Linear, nonlinear and stochastic models produce similar rate constants for protein turnover.
- Calculated peptide degradation rate constants show variabilities, even for the same value of the parameter: the number of hydrogens accessible to deuterium in metabolic labeling.
- As the label duration (incorporation) increases, the numbers of identified and quantified peptides decrease.

Supplementary data

Supplementary data are available online at https://academic.oup.com/bib.

Data availability

All data used in this study are available in Supplementary data and referenced publications.

Funding

References

1. Balch WE, Morimoto RI, Dillin A, *et al.* Adapting proteostasis for disease intervention. *Science* 2008;**319**:916–9.
2. Miller BF, Baehr LM, Musci RV, *et al.* Muscle-specific changes in protein synthesis with aging and reloading after disuse atrophy. *J Cachexia Sarcopenia Muscle* 2019;**10**:1195–209.
3. Sato C, Barthelemy NR, Mawuenyega KG, *et al.* Tau kinetics in neurons and the human central nervous system. *Neuron* 2018;**97**:1284–1298.e7.
4. Labbadia J, Morimoto RI. The biology of proteostasis in aging and disease. *Annu Rev Biochem* 2015;**84**:435–64.
5. Arias IM, Doyle D, Schimke RT. Studies on the synthesis and degradation of proteins of the endoplasmic reticulum of rat liver. *J Biol Chem* 1969;**244**:3303–15.
6. Eden E, Geva-Zatorsky N, Issaeva I, *et al.* Proteome half-life dynamics in living human cells. *Science* 2011;**331**:764–8.
7. Schoenheimer R. *The Dynamic State of Body Constituents*. New York: Hafner Pub. Co., 1964.
8. Wilkinson DJ. Historical and contemporary stable isotope tracer approaches to studying mammalian protein metabolism. *Mass Spectrom Rev* 2018;**37**:57–80.
9. Schwanhausser B, Busse D, Li N, *et al.* Global quantification of mammalian gene expression control. *Nature* 2011;**473**:337–42.
10. Ong SE, Blagoev B, Kratchmarova I, *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;**1**:376–86.
11. Cambridge SB, Gnad F, Nguyen C, *et al.* Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *J Proteome Res* 2011;**10**:5275–84.
12. Larance M, Ahmad Y, Kirkwood KJ, *et al.* Global subcellular characterization of protein degradation using quantitative proteomics. *Mol Cell Proteomics* 2013;**12**:638–50.
13. Kristensen AR, Gsponer J, Foster LJ. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol Syst Biol* 2013;**9**:689.
14. Zecha J, Meng C, Zolg DP, *et al.* Peptide level turnover measurements enable the study of Proteoform dynamics. *Mol Cell Proteomics* 2018;**17**:974–92.
15. Rahman M, Sadygov RG. Predicting the protein half-life in tissue from its cellular properties. *PLoS One* 2017;**12**:e0180428.
16. Dufner DA, Bederman IR, Brunengraber DZ, *et al.* Using 2H2O to study the influence of feeding on protein synthesis: effect of isotope equilibration in vivo vs. in cell culture. *Am J Physiol Endocrinol Metab* 2005;**288**:E1277–83.
17. Ross AB, Langer JD, Jovanovic M. Proteome turnover in the spotlight: approaches, applications, and perspectives. *Mol Cell Proteomics* 2021;**20**:100016.
18. Claydon AJ, Beynon R. Proteome dynamics: revisiting turnover with a global perspective. *Mol Cell Proteomics* 2012;**11**:1551–65.
19. Rauniyar N, McClatchy DB, Yates JR, 3rd. Stable isotope labeling of mammals (SILAM) for in vivo quantitative proteomic analysis. *Methods* 2013;**61**:260–8.
20. Hammond DE, Claydon AJ, Simpson DM, *et al.* Proteome dynamics: tissue variation in the kinetics of proteostasis in intact animals. *Mol Cell Proteomics* 2016;**15**:1204–19.
21. Hsieh EJ, Shulman NJ, Dai DF, *et al.* Topograph, a software platform for precursor enrichment corrected global protein turnover measurements. *Mol Cell Proteomics* 2012;**11**:1468–74.
22. Nolte H, Holper S, Housley MP, *et al.* Dynamics of zebrafish fin regeneration using a pulsed SILAC approach. *Proteomics* 2015;**15**:739–51.
23. Visscher M, De Henau S, Wildschut MHE, *et al.* Proteome-wide changes in protein turnover rates in C. elegans models of longevity and age-related disease. *Cell Rep* 2016;**16**:3041–51.
24. Schober FA, Atanassov I, Moore D, *et al.* Stable isotope Labeling of amino acids in flies (SILAF) reveals differential phosphorylation of mitochondrial proteins upon loss of OXPHOS subunits. *Mol Cell Proteomics* 2021;**20**:100065.
25. Fornasiero EF, Mandad S, Wildhagen H, *et al.* Precisely measured protein lifetimes in the mouse brain reveal differences across tissues and subcellular fractions. *Nat Commun* 2018;**9**:4230.
26. Heo S, Diering GH, Na CH, *et al.* Identification of long-lived synaptic proteins by proteomic analysis of synaptosome protein turnover. *Proc Natl Acad Sci U S A* 2018;**115**:E3827–36.

27. Lehmann S, Hirtz C, Vialaret J, *et al.* In vivo large-scale mapping of protein turnover in human cerebrospinal fluid. *Anal Chem* 2019;**91**:15500–8.

28. Wu CC, MacCoss MJ, Howell KE, *et al.* Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem* 2004;**76**:4951–9.

29. Zhang Y, Reckow S, Webhofer C, *et al.* Proteome scale turnover analysis in live animals using stable isotope metabolic labeling. *Anal Chem* 2011;**83**:1665–72.

30. Price JC, Guan S, Burlingame A, *et al.* Analysis of proteome dynamics in the mouse brain. *Proc Natl Acad Sci U S A* 2010;**107**:14508–13.

31. Savas JN, Toyama BH, Xu T, *et al.* Extremely long-lived nuclear pore proteins in the rat brain. *Science* 2012;**335**:942.

32. Toyama BH, Savas JN, Park SK, *et al.* Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* 2013;**154**:971–82.

33. Kushner DJ, Baker A, Dunstall TG. Pharmacological uses and perspectives of heavy water and deuterated compounds. *Can J Physiol Pharmacol* 1999;**77**:79–88.

34. Kasumov T, Dabkowski ER, Shekar KC, *et al.* Assessment of cardiac proteome dynamics with heavy water: slower protein synthesis rates in interfibrillar than subsarcolemmal mitochondria. *Am J Physiol Heart Circ Physiol* 2013;**304**:H1201–14.

35. Busch R, Kim YK, Neese RA, *et al.* Measurement of protein turnover rates by heavy water labeling of nonessential amino acids. *Biochim Biophys Acta* 2006;**1760**:730–44.

36. Holmes WE, Angel TE, Li KW, *et al.* Dynamic proteomics: in vivo proteome-wide measurement of protein kinetics using metabolic labeling. *Methods Enzymol* 2015;**561**:219–76.

37. Kasumov T, Willard B, Sadygov RG, *et al.* Proteome dynamics with heavy water – instrumentations, data analysis, and biological applications, recent advances in proteomics research. In: Magdeldin S (ed). *Proteomics/Book 2.* London, UK: IntechOpen, 2015, 21–55.

38. Miller BF, Reid JJ, Price JC, *et al.* CORP: the use of deuterated water for the measurement of protein synthesis. *J Appl Physiol (1985)* 2020;**128**:1163–76.

39. Kim TY, Wang D, Kim AK, *et al.* Metabolic labeling reveals proteome dynamics of mouse mitochondria. *Mol Cell Proteomics* 2012;**11**:1586–94.

40. Kasumov T, Ilchenko S, Li L, *et al.* Measuring protein synthesis using metabolic (2)H labeling, high-resolution mass spectrometry, and an algorithm. *Anal Biochem* 2011;**412**:47–55.

41. Dufner D, Previs SF. Measuring in vivo metabolism using heavy water. *Curr Opin Clin Nutr Metab Care* 2003;**6**:511–7.

42. Borzou A, Sadygov VR, Zhang W, *et al.* Proteome dynamics from heavy water metabolic labeling and peptide tandem mass spectrometry. *Int J Mass Spectrom* 2019;**445**:116194.

43. Wang B, Sun G, Anderson DR, *et al.* Isotopologue distributions of peptide product ions by tandem mass spectrometry: quantitation of low levels of deuterium incorporation. *Anal Biochem* 2007;**367**:40–8.

44. Angel TE, Naylor BC, Price JC, *et al.* Improved sensitivity for protein turnover quantification by monitoring Immonium ion Isotopologue abundance. *Anal Chem* 2019;**91**:9732–40.

45. Sadygov RG, Avva J, Rahman M, *et al.* d2ome, software for in vivo protein turnover analysis using heavy water Labeling and LC-MS, reveals alterations of hepatic proteome dynamics in a mouse model of NAFLD. *J Proteome Res* 2018;**17**:3740–8.

46. Sadygov VR, Zhang W, Sadygov RG. Timepoint selection strategy for in vivo proteome dynamics from heavy water metabolic labeling and LC-MS. *J Proteome Res* 2020;**19**:2105–12.

47. Perkins DN, Pappin DJ, Creasy DM, *et al.* Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;**20**:3551–67.

48. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 1994;**5**:976–89.

49. Xu T, Park SK, Venable JD, *et al.* ProLuCID: an improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J Proteomics* 2015;**129**:16–24.

50. Kim S, Pevzner PA. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* 2014;**5**:5277.

51. Sadygov RG. Partial isotope profiles are sufficient for protein turnover analysis using closed-form equations of mass Isotopomer dynamics. *Anal Chem* 2020;**92**:14747–53.

52. Hellerstein MK, Neese RA. Mass isotopomer distribution analysis: a technique for measuring biosynthesis and turnover of polymers. *Am J Physiol* 1992;**263**:E988–1001.

53. MacCoss MJ, Wu CC, Matthews DE, *et al.* Measurement of the isotope enrichment of stable isotope-labeled proteins using high-resolution mass spectra of peptides. *Anal Chem* 2005;**77**:7646–53.

54. Daurio NA, Zhou H, Chen Y, *et al.* Examining targeted protein degradation from physiological and analytical perspectives: enabling translation between cells and subjects. *ACS Chem Biol* 2020;**15**:2623–35.

55. Guan S, Price JC, Ghaemmaghami S, *et al.* Compartment modeling for mammalian protein turnover studies by stable isotope metabolic labeling. *Anal Chem* 2012;**84**:4014–21.

56. Casella G, Berger RL. *Statistical Inference.* Australia: Thomson Learning, 2002.

57. Huang L, Kim D, Liu X, *et al.* Estimating relative changes of metabolic fluxes. *PLoS Comput Biol* 2014;**10**:e1003958.

58. Wilkinson DJ, Franchi MV, Brook MS, *et al.* A validation of the application of D(2)O stable isotope tracer techniques for monitoring day-to-day changes in muscle protein subfraction synthesis in humans. *Am J Physiol Endocrinol Metab* 2014;**306**:E571–9.

59. Rahman M, Previs SF, Kasumov T, *et al.* Gaussian process modeling of protein turnover. *J Proteome Res* 2016;**15**:2115–22.

60. Hotchko M, Anand GS, Komives EA, *et al.* Automated extraction of backbone deuteration levels from amide H/2H mass spectrometry experiments. *Protein Sci* 2006;**15**:583–601.

61. Na S, Lee JJ, Joo JWJ, *et al.* deMix: decoding deuterated distributions from heterogeneous protein states via HDX-MS. *Sci Rep* 2019;**9**:3176.

62. Yamamoto T, Yamagaki T, Satake H. Development of software for the in-depth analysis of protein dynamics as determined by MALDI mass spectrometry-based hydrogen/deuterium exchange. *Mass Spectrom* 2020;**8**:S0082.

63. Lam MP, Wang D, Lau E, *et al.* Protein kinetic signatures of the remodeling heart following isoproterenol stimulation. *J Clin Invest* 2014;**124**:1734–44.

64. Kall L, Storey JD, MacCoss MJ, *et al.* Posterior error probabilities and false discovery rates: two sides of the same coin. *J Proteome Res* 2008;**7**:40–4.

65. Commerford SL, Carsten AL, Cronkite EP. The distribution of tritium among the amino acids of proteins obtained from mice exposed to tritiated water. *Radiat Res* 1983;**94**:151–5.

66. Ilchenko S, Haddad A, Sadana P, *et al.* Calculation of the protein turnover rate using the number of incorporated (2)H atoms and proteomics analysis of a single Labeled sample. *Anal Chem* 2019;**91**:14340–51.

67. Su X, Lu W, Rabinowitz JD. Metabolite spectral accuracy on orbitraps. *Anal Chem* 2017;**89**:5940–8.

68. Ilchenko S, Previs SF, Rachdaoui N, *et al*. An improved measurement of isotopic ratios by high resolution mass spectrometry. *J Am Soc Mass Spectrom* 2013;**24**:309–12.

69. Gorshkov MV, Fornelli L, Tsybin YO. Observation of ion coalescence in Orbitrap Fourier transform mass spectrometry. *Rapid Commun Mass Spectrom* 2012;**26**:1711–7.

70. Sadygov RG. High-resolution mass spectrometry for in vivo proteome dynamics using heavy water metabolic labeling. *Int J Mol Sci* 2020;**21**:7821.

71. Hofmann AE, Chimiak L, Dallas B, *et al*. Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures. *Int J Mass Spectrom* 2020;**457**:116410.

72. Naylor BC, Porter MT, Wilson E, *et al*. DeuteRater: a tool for quantifying peptide isotope precision and kinetic proteomics. *Bioinformatics* 2017;**33**:1514–20.

73. Lau E, Cao Q, Ng DC, *et al*. A large dataset of protein dynamics in the mammalian heart proteome. *Sci Data* 2016;**3**:160015.

74. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2019.

75. Hellerstein MK, Neese RA. Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. *Am J Physiol* 1999;**276**:E1146–70.

76. Sadygov RG. Using SEQUEST with theoretically complete sequence databases. *J Am Soc Mass Spectrom* 2015;**26**:1858–64.

77. Sadygov RG. Poisson model to generate isotope distribution for biomolecules. *J Proteome Res* 2018;**17**:751–8.

78. O'Brien JJ, Narayan V, Wong Y, *et al*. Precise estimation of in vivo protein turnover rates. *bioRxiv* 2020; 2020.2011.2010.377440.

79. Carpenter B, Gelman A, Hoffman MD, *et al*. Stan: a probabilistic programming language. *J Stat Softw* 2017;**76**:32.

80. Byrd RH, Lu PH, Nocedal J, *et al*. A limited memory algorithm for bound constrained optimization. *Siam J Sci Comput* 1995;**16**: 1190–208.

81. Bateman RJ, Munsell LY, Chen X, *et al*. Stable isotope labeling tandem mass spectrometry (SILT) to quantify protein production and clearance rates. *J Am Soc Mass Spectrom* 2007;**18**: 997–1006.

82. Paterson RW, Gabelle A, Lucey BP, *et al*. SILK studies - capturing the turnover of proteins linked to neurodegenerative diseases. *Nat Rev Neurol* 2019;**15**:419–27.

83. Holman SW, Hammond DE, Simpson DM, *et al*. Protein turnover measurement using selected reaction monitoring-mass spectrometry (SRM-MS). *Philos Trans A Math Phys Eng Sci* 2016;**374**:20150362.

84. Tomazela DM, Patterson BW, Hanson E, *et al*. Measurement of human surfactant protein-B turnover in vivo from tracheal aspirates using targeted proteomics. *Anal Chem* 2010;**82**: 2561–7.

85. Lee AY, Yates NA, Ichetovkin M, *et al*. Measurement of fractional synthetic rates of multiple protein analytes by triple quadrupole mass spectrometry. *Clin Chem* 2012;**58**:619–27.

86. Szklarczyk D, Morris JH, Cook H, *et al*. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;**45**:D362–8.

87. Dai DF, Karunadharma PP, Chiao YA, *et al*. Altered proteome turnover and remodeling by short-term caloric restriction or rapamycin rejuvenate the aging heart. *Aging Cell* 2014;**13**: 529–39.

88. Lee K, Haddad A, Osme A, *et al*. Hepatic mitochondrial defects in a nonalcoholic fatty liver disease mouse model are associated with increased degradation of oxidative phosphorylation subunits. *Mol Cell Proteomics* 2018;**17**:2371–86.

89. Shekar KC, Li L, Dabkowski ER, *et al*. Cardiac mitochondrial proteome dynamics with heavy water reveals stable rate of mitochondrial protein synthesis in heart failure despite decline in mitochondrial oxidative capacity. *J Mol Cell Cardiol* 2014;**75**: 88–97.

90. Golizeh M, Lee K, Ilchenko S, *et al*. Increased serotransferrin and ceruloplasmin turnover in diet-controlled patients with type 2 diabetes. *Free Radic Biol Med* 2017;**113**:461–9.

91. Rolfs Z, Frey BL, Shi X, *et al*. An atlas of protein turnover rates in mouse tissues. *Nat Commun* 2021;**12**:6778.

92. Chepyala SR, Liu X, Yang K, *et al*. JUMPt: comprehensive protein turnover modeling of in vivo pulse SILAC data by ordinary differential equations. *Anal Chem* 2021;**93**:13495–504.