Compartment Modeling for Mammalian Protein Turnover Studies by Stable Isotope Metabolic Labeling

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2012

Impact Factor: 5.64 · DOI: 10.1021/ac203330z · Source: PubMed

5 AUTHORS, INCLUDING:

John Price
Brigham Young University - Provo Main Campus
29 PUBLICATIONS **1,305** CITATIONS
SEE PROFILE

Sina Ghaemmaghami
University of Rochester
27 PUBLICATIONS **4,864** CITATIONS
SEE PROFILE

Alma L Burlingame
University of California, San Francisco
622 PUBLICATIONS **26,570** CITATIONS
SEE PROFILE

analytical chemistry

Article

pubs.acs.org/ac

Compartment Modeling for Mammalian Protein Turnover Studies by Stable Isotope Metabolic Labeling

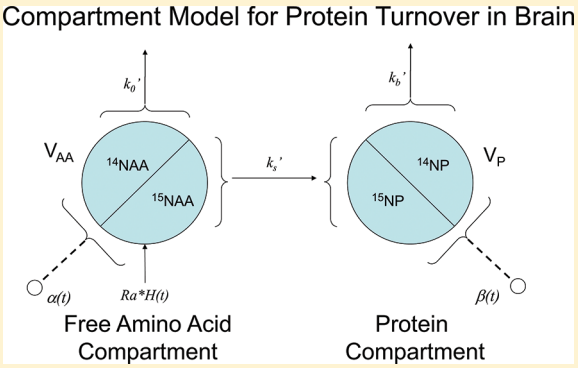
Shenheng Guan,*[,†] John C. Price,[‡] Sina Ghaemmaghami,[‡,§] Stanley B. Prusiner,[‡,§] and Alma L. Burlingame[†]

[†]Department of Pharmaceutical Chemistry and Mass Spectrometry Facility, University of California, San Francisco, California 94158-2517, United States

[‡]Institute for Neurodegenerative Diseases and [§]Department of Neurology, University of California, San Francisco, California 94143-0518, United States

Supporting Information

ABSTRACT: Protein turnover studies on a proteome scale based on metabolic isotopic labeling can provide a systematic understanding of mechanisms for regulation of protein abundances and their transient behaviors. At this time, these large-scale studies typically utilize a simple kinetic model to extract protein dynamic information. Although many high-quality, protein isotope incorporation data are available from those experiments, accurate and additionally useful protein dynamic information cannot be extracted from the experimental data by use of the simple kinetic models. In this paper, we describe a formal connection between data obtained from elemental isotope labeling experiments and the well-known compartment modeling, and we demonstrate that an appropriate application of a compartment model to turnover of proteins from mammalian tissues can indeed lead to a better fitting of the experimental data.



Quantitative proteomics studies typically provide static comparison of protein abundance differences of the system under study. However, proteins are constantly synthesized and degraded in a living organism depending on their cellular localization and functionality even if their concentrations are kept constant. Although proteins are synthesized by a canonical pathway, protein synthesis rates are subjected to precise transcriptional and translational controls and are influenced by the availability of amino acids. A cell also possesses a number of pathways to remove unwanted protein from the system to modulate the appropriate protein abundances. Abnormal activities of protein degradation pathways have been associated with the onset of numerous diseases. It is likely that all neurodegenerative disorders, including Alzheimer's (AD), Parkinson's, and Huntington's diseases as well as the frontotemporal dementias, may result from the inability of cells to remove the prion forms of etiologic proteins, which subsequently assemble into insoluble aggregates.[1] Understanding the molecular mechanisms of those neurodegenerative diseases may be facilitated by direct measurement of protein turnover in healthy and diseased tissues.

To study protein turnover, it is necessary to apply metabolic isotope labeling methods to trace the amino acid incorporation into a protein. For single-cell organisms, dynamic stable isotope labeling with amino acids in culture (dynamic SILAC) experiments[2] utilize a single or a few isotopically labeled amino acids to measure protein turnover on a proteome scale. The abundance ratio of labeled and unlabeled peptide ions can be directly used to construct peptide incorporation curves. Even in cell cultures, measurement of free amino acids in their intracellular compartments is difficult due to the complex dynamics in protein degradation and tRNA recycling. It is expected that protein turnover in a complex organism, such as a mammal, has much more complicated kinetics.

Recently, we have traced the global protein turnover in mammalian tissues by use of ^{15}N stable isotope labeling and high-resolution mass spectrometry detection.[3] Unlike typical dynamic SILAC experiments, all nitrogen atoms in the food source are labeled with the heavy nitrogen isotope. The advantages of using ^{15}N labeling include easy production of a natural food source (blue-green algae, *Spirulina platensis*) for mammalian models with near purity of heavy isotope and high dynamic range of isotope incorporation measurement. A disadvantage of single-element labeling compared to single amino-acid labeling is that more sophisticated data-processing software is needed to extract protein turnover information from complex spectra containing many isotopic peaks. A data-processing platform that fulfills this requirement has been described elsewhere.[4] Using a combination of careful execution of optimized experimental protocols and robust data-processing

Received: December 15, 2011
Accepted: March 23, 2012
Published: March 23, 2012

   dx.doi.org/10.1021/ac203330z | *Anal. Chem.* 2012, 84, 4014−4021

algorithms, we were able to measure protein turnover rate constants for ~1700 proteins in mouse tissues. In order to obtain global proteome dynamics in the mammalian samples, we used a simple empirical kinetic model to capture the essence of protein turnover in different tissues. A delayed exponential function was proposed to describe the simple kinetic model.[3] The delay time parameter was included to account for the delay in delivering amino acids from the digestion of the food source. Although the simple kinetic model or the simple model functional form of amino acid incorporation curve fits the general structure of the peptide/protein incorporation curves, there are discrepancies in details of the experimental data so that the simple functional form cannot fit well. In this contribution, we describe the use of more sophisticated compartment kinetic models in an attempt to explain the observed discrepancies. Although the concept of compartment modeling for protein turnover was developed in the 1950s,[5] we have not found adequate literature providing a clear description for its applications to large-scale dynamic proteomics studies. Here, we attempt to establish a consistent relationship between the experimentally observed data with a compartment model with minimal parameters. The newly developed compartment models can provide the functional forms that can be used for more satisfactory fitting to experimentally observed protein incorporation curves, compared to the simple, widely used one-exponential function. More importantly, use of compartment models allows us to extract an individual protein's turnover rate constant and additional turnover rate constants, such as the free amino acid compartment turnover rate constant, by measuring only the incorporation curve of the protein of interest.

THEORY

One-Compartment Model. Before examining compartment models for isotope labeling of proteins, we investigated the kinetic information that can be extracted from compartment modeling by measuring relative concentration of labels (Figure 1). The following derivation of the kinetic solution of a single-compartment model is not novel, but it sets the stage for a better understanding of more complex compartment models. In the one-compartment model shown in Figure 1, let R_A be the flow rate of pure label A into a compartment (or a pool) initially containing no label, with a volume of V, starting at t = 0. There are two key assumptions of our current compartment modeling. The first assumption is that a label mixes instantly in a compartment. The second assumption is that the size of a compartment does not change during the experiment. In the model, the same volume of mixed material must flow out of the compartment. The concentration of label, [A], in the compartment can be described by the differential equation

$$\frac{d[A]}{dt} = \frac{R_A}{V}([A]_T - [A]) \quad (1)$$

in which $[A]_T$ is the concentration of pure label A. If we measure the relative concentration of A in the compartment, the above equation becomes

$$\frac{d\alpha}{dt} = \frac{R_A}{V}(1 - \alpha) \quad (2)$$

in which $\alpha = [A]/[A]_T$ is the relative (or fractional) concentration of label A. The solution of the equation is

$$\alpha(t) = [1 - e^{-(R_A/V)t}] \quad (3)$$

Therefore, the rate constant, $k_A = R_A/V$, is the relative volumetric flow rate of label A and the rate constant is inversely proportional to the size of the compartment. This derivation highlights a limitation of an experiment based on measurement of relative concentration of labels. Independent experiments must be carried out in order to obtain independent information about mass or volumetric flow rates or the sizes of compartments.

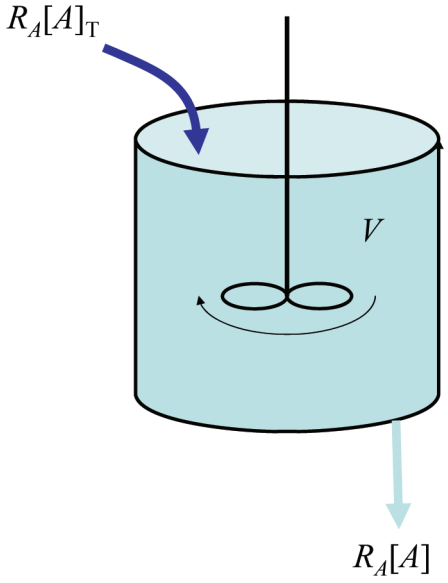


Figure 1. One-compartment model. $R_A[A]_T$ and $R_A[A]$ are the quantities (moles if the concentrations are in molar units; R_A is in liters per unit time) flowed into and out of the compartment per unit of time, respectively.

Two-Compartment Model of Isotope Labeling. We now consider ^{15}N-labeled amino acids in a two-compartment model (Figure 2). We assign [AA] as the total concentration of amino acids in the free amino acid pool, [P] as the total concentration of amino acids in the protein-bound pool, V_{AA} as the volumetric size of the free amino acid pool, V_P as the volumetric size of the protein-bound pool, [^{14}NAA] as the concentration of ^{14}N amino acids in the free amino acid pool, [^{15}NAA] as the concentration of ^{15}N amino acids in the free

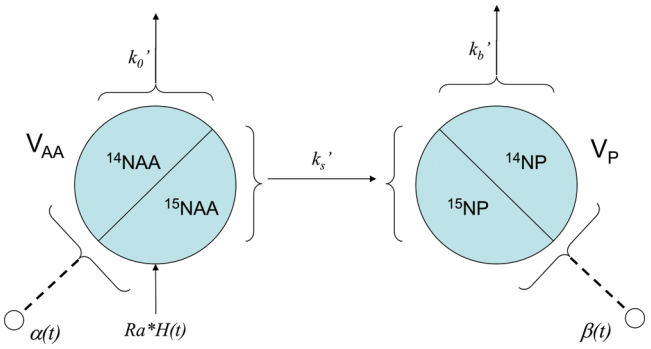


Figure 2. Two-compartment model for protein turnover in brain. $H(t)$ is the step function of t; $H(t) = 0$ for $t < 0$ and $H(t) = 1$ for $t \geq 0$. V_{AA} and V_P are the volumetric sizes of the free amino acid pool and the protein-bound pool, respectively. If it is assumed that there is not an isotope effect, the rate constants are the same for both unlabeled and labeled amino acids. α and β are the fractional concentrations of ^{15}N-labeled amino acids in the free amino acid pool and the protein-bound pool, respectively.

dx.doi.org/10.1021/ac203330z | Anal. Chem. 2012, 84, 4014–4021

amino acid pool, [^{14}NP] as the concentration of ^{14}N amino acids in the protein-bound pool, and [^{15}NP] as the concentration of ^{15}N amino acids in the protein-bound pool.

The model neglects the contribution of the breakdown of a protein of interest into the free amino acid pool because an individual protein's contribution is likely to be small. Assume that, at time t = 0, ^{14}N food source is replaced by ^{15}N food source with constant amino acid influx of R_a. According to the model, eqs 4 and 5 can be established for ^{14}N amino acids

$$V_{AA}\frac{d[^{14}NAA]}{dt} = -(k_s' + k_{0a}')[^{14}NAA] \quad (4)$$

$$V_P\frac{d[^{14}NP]}{dt} = k_s'[^{14}NAA] - k_b'[^{14}NP] \quad (5)$$

and eqs 6 and 7 can be established for ^{15}N amino acids

$$V_{AA}\frac{d[^{15}NAA]}{dt} = -(k_s' + k_{0a}')[^{15}NAA] + R_a[AA] \quad (6)$$

$$V_P\frac{d[^{15}NP]}{dt} = k_s'[^{15}NAA] - k_b'[^{15}NP] \quad (7)$$

Under a homeostatic condition (compartment sizes are constant), the total concentration of the free amino acid pool [AA] and that for the protein-bound pool [P] are constants:

$$[^{14}NAA] + [^{15}NAA] = [AA] \quad (8)$$

$$[^{14}NP] + [^{15}NP] = [P] \quad (9)$$

We can derive the following relationship due to mass conservation

$$R_a = k_{0a}' + k_s' \quad (10)$$

by use of eqs 4 and 6, and

$$\frac{[AA]}{[P]} = \frac{k_b'}{k_s'} \quad (11)$$

by use of eqs 5 and 7. The differential equations (eqs 4−7) are also subjected to the initial conditions (the experimental procedure):

$$[^{15}NAA]_{t=0} = 0 \quad (12)$$

$$[^{15}NP]_{t=0} = 0 \quad (13)$$

Introduction of the variables α and β

$$\alpha = \frac{[^{15}NAA]}{[AA]} \quad (14)$$

$$\beta = \frac{[^{15}NP]}{[P]} \quad (15)$$

transforms eqs 6 and 7 into

$$\frac{d\alpha}{dt} = -k_0\alpha + k_0 \quad (16)$$

$$\frac{d\beta}{dt} = k_b\alpha - k_b\beta \quad (17)$$

with the initial conditions of

$$\alpha(0) = 0 \quad (18)$$

$$\beta(0) = 0 \quad (19)$$

Here, $k_0 = (k_s' + k_0')/V_{AA}$ is the lumped rate constant for amino acids removal from the free amino acid pool, and $k_b = k_b'/V_P$ is the protein degradation rate constant. Just as illustrated in the one-compartment model, it is clear that those rate constants are the relative volumetric flow rate constants normalized by their respective compartment volume sizes. The solution of eq 16 is an exponential recovery function, and the solution of eq 17 is that driven by the exponential recovery. Equations 16 and 17 can be easily solved by use of Laplace transform and the solution of interest as

$$\alpha(t) = 1 - e^{-k_0 t} \quad (20)$$

$$\beta(t) = 1 + \frac{k_b e^{-k_0 t}}{k_0 - k_b} - \frac{k_0 e^{-k_b t}}{k_0 - k_b} \quad (21)$$

The half-life times for the free amino acid pool and the protein pool can be found as ln (2)/k_0 and ln (2)/k_b, respectively. Depending on the compartment model and where the label concentrations are measured, some parameters cannot be uniquely determined. The investigation of whether parameters can be determined is called identifiability analysis.[5a] In this model, we should be able to extract protein degradation and free amino acid removal rate constants by measuring the amino acid relative concentration in the protein-bound pool alone (β).

Relationship between Compartment Models and Experimental Data. Now, we need to establish a connection between the relative ^{15}N-labeled amino acid concentration in protein (β, eq 15) and the experimentally observed data. Each amino acid contains at least one nitrogen atom; some amino acids contain two (asparagine, glutamine, lysine, and tryptophan), three (histidine), or four (arginine) nitrogen atoms. The relative concentration of ^{15}N-labeled amino acid in a peptide is proportional to the relative concentration of ^{15}N atoms in the peptide if all amino acids incorporate into proteins with the same probability. This is, of course, an approximation because there is some delay for nonessential amino acids to become available compared to the essential amino acids. The proportional constant is 1.3577 for an averagine[6] (the elemental composition of an "average" amino acid in an unlabeled proteome is $C_{4.9384}H_{7.7583}N_{1.3577}O_{1.4773}S_{0.0417}$) . The proportional constant cancels out when the relative fractional concentration is computed (eq 15). In our data processing pipeline,[4] we computed the ^{15}N distribution for every peptide ion identified by a database search. An ^{15}N distribution is the relative fraction of the peptide ion containing a certain number of ^{15}N atoms.

Let $F = \{f_{ij}\}$ be the ^{15}N distribution matrix, in which row index i = 1, 2, ..., N_{IT} is the incorporation time points and column index j = 0, 1, ..., N_{NA} is the number of ^{15}N atoms incorporated. j = 0 corresponds to the peptide with zero ^{15}N incorporation, and $j = N_{NA}$ is the maximal ^{15}N incorporation number. The ^{15}N distribution, of a given incorporation time point i, is converted to relative ^{15}N fraction in the peptide by

$$RF_i = \left[\sum_{j=0}^{N_{NA}}\left(\frac{j}{N_{NA}}\right)\frac{f_{ij}}{\sum_{j=0}^{N_{NA}} f_{ij}}\right] \quad (22)$$

The number N_{NA} is the possible nitrogen atoms that can be metabolically labeled. j/N_{NA} is the relative ratio of ^{15}N isotopes over the total exchangeable nitrogen atoms. The $f_{ij}/\Sigma_{j=0}^{N_{NA}} f_{ij}$ can

dx.doi.org/10.1021/ac203330z | Anal. Chem. 2012, 84, 4014−4021

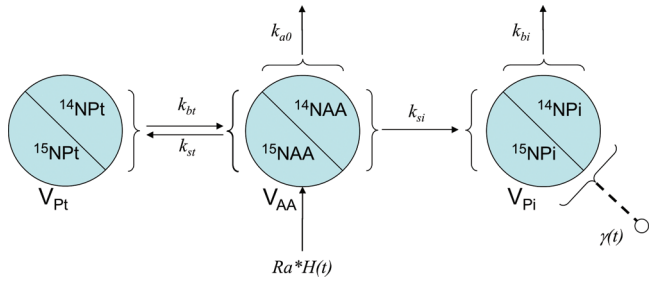


Figure 3. Three-compartment model for protein turnover in liver. The compartments are the free amino acid pool (^{14}NAA/^{15}NAA), the total protein pool ($^{14}NP_t/^{15}NP_t$), and the protein-of-interest pool ($^{14}NP_i/^{15}NP_i$). V_{AA}, V_{P_t}, and V_{P_i} are the volumetric sizes of the free amino acid pool, the total protein-bound pool, and the individual protein-bound pool, respectively.

be considered as the probability of the peptide ion having j number of ^{15}N atoms. The relative fraction of ^{14}N is simply $1 - RF_i$ for the incorporation time point i. In this work, a peptide incorporation curve is the function of the relative ^{15}N fraction, RF, against the incorporation time. Recently, Kasumov et al.[7] used the total moles percent enrichment (MPE) quantity for their ^{2}H labeling experiments, that is analogous to our definition of relative isotope fraction.

Three-Compartment Model. As described in the Results section, the two-compartment model fit the peptide/protein incorporation relative fraction curves well for proteins in mouse brain. However, the two-compartment model fit poorly to proteins in the liver. An examination of the relative fraction incorporation curves from liver protein revealed that the curves cannot be fitted well even with the general two-exponential functions (see Results section for detailed discussion). Therefore, we added an additional total protein compartment to construct a three-compartment model, shown in Figure 3.

Inclusion of the total protein pool is justified because the amino acid recycling by protein degradation and protein export in liver must be considered. The detailed derivation of the three-pool model functional form is given in the Supporting Information. A full three-compartment model accounting for both total protein export and total protein degradation (illustrated in Figure S1 of the Supporting Information) requires definition of five rate constants. However, we found the fitting of the five-parameter system was not robust because the optimization procedure can easily be trapped in a local minimum. The value for the rate constants for both total protein degradation into free amino acid (k_{bt}) and total protein export (k_{0t}) should be small because they are normalized by the large volumetric size of the total protein pool. Indeed, the five-rate constant values for good fits showed that the total protein export rate constant was typically the smallest one. Thus, we removed the rate constant for total protein export from consideration in order to reduce the degree of parametric dimension, resulting in a four-rate-constant system (Figure 3). The relative concentration of ^{15}N-labeled amino acid in each protein of interest is

$$\gamma(t) = 1 + y_u e^{-ut} + y_v e^{-vt} + y_{k_{bi}} e^{-k_{bi}t} \tag{23}$$

in which

$$u = \frac{(k_{st} + k_{0a} + k_{bt}) - \sqrt{(k_{st} + k_{0a} + k_{bt})^2 - 4k_{0a}k_{bt}}}{2} \tag{24}$$

$$v = \frac{(k_{st} + k_{0a} + k_{bt}) + \sqrt{(k_{st} + k_{0a} + k_{bt})^2 - 4k_{0a}k_{bt}}}{2} \tag{25}$$

$$y_u = \frac{k_{0a}k_{bi}(u - k_{bt})}{(u - v)(u - k_{bi})u} \tag{26}$$

$$y_v = \frac{k_{0a}k_{bi}(v - k_{bt})}{(v - u)(v - k_{bi})v} \tag{27}$$

$$y_{k_{bi}} = \frac{k_{0a}(k_{bi} - k_{bt})}{(u - k_{bi})(v - k_{bi})} \tag{28}$$

The solution contains three exponentials. A general form of three-exponential function $[y(t) = 1 + y_1e^{-k_1t} + y_2e^{-k_2t} + y_3e^{-k_3t}]$ requires six parameters for its definition. Here only four parameters—k_{st}, total protein synthesis rate constant; k_{bt}, total protein degradation rate constant; k_{0a}, amino acid outflow rate constant; and k_{bi}, individual protein degradation rate constant—completely define our three-compartment model. In other words, the system has four degrees of parametric freedom. When the function was used to fit the experimental incorporation curves, the fitness was similar to the fittings with the full five-parameter function containing the total protein export rate constant.

RESULTS

All data are from experiments previously described.[3] The data processing for the ^{15}N distribution of peptide ions was described.[4] Fitting the compartment model functional forms with ^{15}N distribution data was done similarly as described.[4] Briefly, the relative ^{15}N fraction was calculated according to eq 22 and the peptide incorporation curves were constructed from the relative ^{15}N fraction data points against the incorporation times. The relative ^{15}N fraction incorporation curves of the peptide ions belonging to a protein were aggregated to the protein ^{15}N incorporation curve by use of a similarity selection procedure after the shared peptides were removed. In order to guarantee the quality of the protein incorporation curves, we only considered the curves obtained by aggregation of at least two peptide curves and that contained at least four nonzero data points out of nine incorporation times. Still, there were 705 and 797 protein curves to consider in brain and liver, respectively. The Supporting Information contains the protein ^{15}N relative fraction incorporation curves along with the fitted rate constants.

dx.doi.org/10.1021/ac203330z | Anal. Chem. 2012, 84, 4014−4021

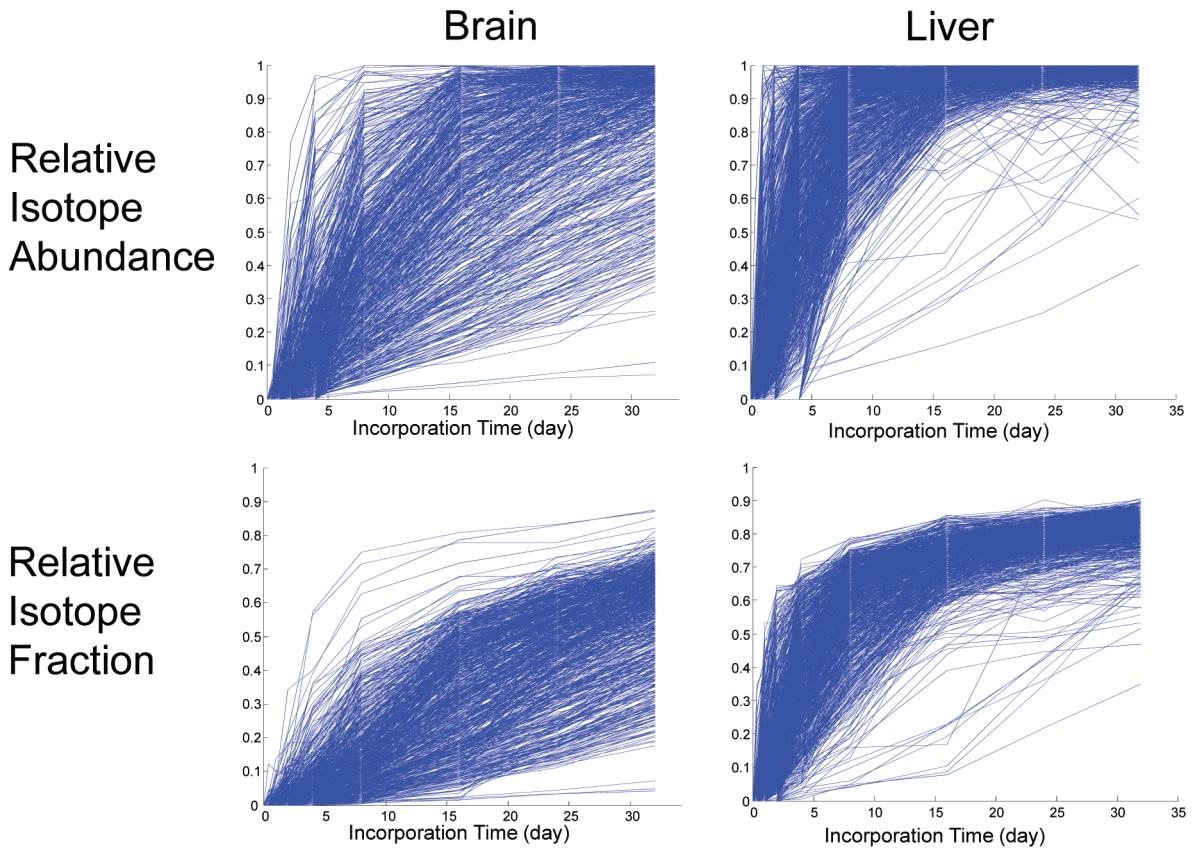


Figure 4. Peptide incorporation curves constructed by use of relative isotope abundance (RIA; top panels) or by use of relative isotope fraction, defined by eq 22 (bottom panels). Peptide curves are from brain proteins (left panels) and from liver proteins (right panels). The relative isotope abundance (RIA) definition can be found in ref 8 or in Supporting Information of ref 4.

Relative Isotope Fraction. As discussed under Theory, relative ^{15}N fraction (substituting for ^{15}N-labeled amino acid concentration) provides a direct connection of compartment models to experimental observation. An additional advantage of using relative isotope fraction instead of the commonly used relative isotope abundance (RIA) (defined in eq 1 of ref 8) is illustrated in Figure 4. For proteins with rapid turnover, the incorporation curves using relative isotope abundance (RIA) can quickly reach saturation. This is quite common for many proteins in liver (Figure 4, upper right). In comparison, no proteins attain saturation with the incorporation curves using relative isotope fraction (Figure 4, lower panels).

Brain Proteins: Two-Compartment Modeling. The relative ^{15}N fraction incorporation curves for brain proteins were subjected to a nonlinear regression procedure to fit the theoretical function of eq 21 for the two-compartment model. Compared to the empirical delayed exponential function [$\beta(t) = 1 - e^{-k_0(t-t_0)}$, for which k_0 is the rate constant and t_0 is the time delay] to fit the relative isotope abundance proposed in ref 3, the two-compartment function fit better for almost all of the 700 protein relative fraction incorporation curves from brain, although both models use two fitting parameters (Figure 5).

Shown in Figure 5 are two curve fittings of the same protein (phosphatidylethanolamine-binding protein 1, P70296). The fitting was much more precise for the two-compartment model (Figure 5, lower trace). The rate constant obtained from the two-compartment model (0.0373 day^{-1}) is approximately 40% smaller than that of 0.0563 day^{-1} from the delayed exponential fit (Figure 5, upper trace).

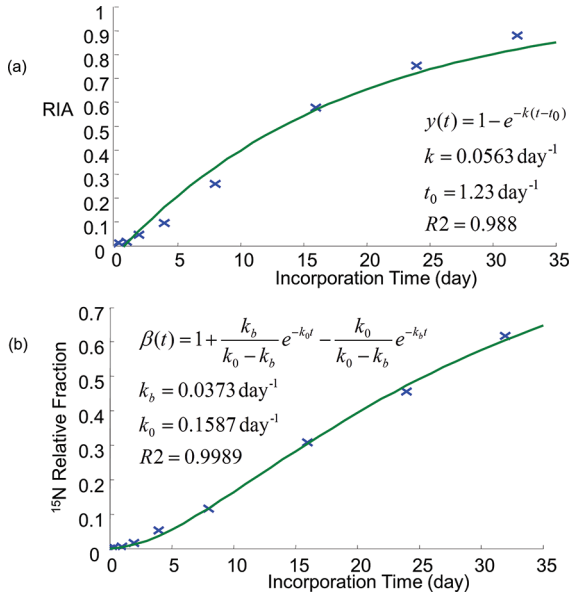


Figure 5. Fitting protein incorporation curves (a) by use of a delayed exponential function for relative isotope abundance and (b) by use of the two-compartment model function of eq 21 for relative isotope fraction.

The improvement in fitting of the two-compartment model over the delayed exponential function for all the selected brain proteins is illustrated in Figure S2 of the Supporting Information. The majority (>70%) of the two-compartment model fitting had

dx.doi.org/10.1021/ac203330z | Anal. Chem. 2012, 84, 4014−4021

R^2 values of 0.99 or better. In the case of the delayed exponential fit, a smaller portion (<40%) can be fitted equally or better.

It is interesting to examine the extreme cases of the functional form of eq 21. If the protein degradation rate constant k_b is much smaller than the amino acid removal constant k_0, or $k_b \ll k_0$, the solution (eq 21) reduces to a single-exponential function (one compartment).

$$\beta(t) \approx 1 - e^{-k_b t} \tag{29}$$

One additional disadvantage for using the delayed exponential function is the sensitivity of the fit with early incorporation time points. If the delay factor is positive, the delayed exponential function becomes negative when the incorporation time is small. This is, of course, inconsistent with the definition of relative isotope abundance, which should have a possible range from zero to one. One can consider eliminating the early incorporation time points; however, such a choice is subjective and a data processing procedure based on such input is not robust.

Any nonlinear fitting requires initial parametric values. The fitting of the two-compartment model is robust. For all the protein curves tested, the fitting converged to the global minimum with the same initial values for both rate constants. It is perhaps not surprising here because there are only two degrees of parametric freedom in this system.

Liver Proteins: Three-Compartment Modeling. When attempting to fit the ^{15}N relative fraction incorporation curves for liver proteins with the two-compartment model function of eq 21, a high degree of deviation was observed (Figure S3 of the Supporting Information). In fact, the two-compartment model fitted more poorly than the delayed exponential function, which is equivalent to a general one-exponential function $(1 - y_0 e^{-kt})$, in which y_0 and k are the fitting parameters). Using a general two-exponential function $[\gamma(t) = 1 + y_1 e^{-k_1 t} + y_2 e^{-k_2 t}$ with four parameters] did not produce better fitting (data not shown). Introduction of the third total protein compartment produced much-improved fitting for most proteins in liver, even when the number of the fitting parameters was kept at four. A typical example is shown in Figure 6, in which the ^{15}N relative fraction incorporation curve for the protein transitional endoplasmic reticulum ATPase (Q01853) was constructed from 37 of its peptide incorporation curves. In a liver protein turnover study with ^{14}C labeling,[9] a three-exponential function was found to be more appropriate.

The two-compartment model fitted poorly with the protein degradation rate constant of 0.137 day^{-1} and the three-compartment model fitted nearly perfectly with the data for the protein degradation rate constant of 0.317 day^{-1}. It is interesting to note that the total protein degradation rate constant (k_{bt}) is very small. This perhaps should not be surprising because the total protein pool size is very large compared to that for free amino acids in a tissue, as discussed for the single-compartment model. The statistical analysis of all total protein degradation rate constants for liver proteins supports this statement (Figure S4 of the Supporting Information). The median value of the total protein degradation rate constants for all the liver proteins was approximately 0.04 day^{-1}, which is at least 1 order of magnitude smaller than the other rate constants. Some total protein degradation rate constants had negative values. Of course, these negative values cannot be correct. But from an experimental point of view, this can be explained by a situation in which the variation caused by observation errors is larger than the expected values.

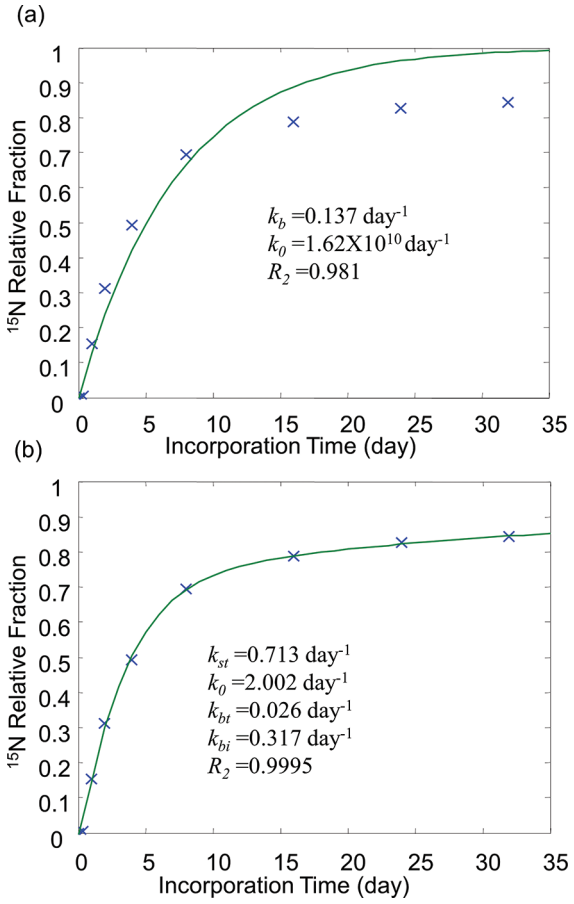


Figure 6. Illustration of curve fitting for a liver protein, transitional endoplasmic reticulum ATPase: (a) two-compartment model function fitting with two rate constants and (b) three-compartment model function fitting with four rate constants.

DISCUSSION

In the Theory section, we introduced a series of compartment models with increasing sophistication in order to fit the experimentally observed data better. The most important development here is the realization that the relative isotopically labeled fraction can serve as a direct connection between compartment modeling and experimental data. The relative isotope fraction can be accurately extracted from a single survey scan spectrum from its isotopic peaks without the need to introduce an external factor for its normalization. In a typical dynamic SILAC experiment, if the detected peptide contains a single labeled amino acid, the relative isotope abundance is equal to the relative fraction concentration. However, if the peptide of interest contains more than one labeled amino acid, the relative isotope fraction is a more appropriate indication of the relative amino acid concentration of the peptide instead of the relative isotope abundance.

In our previous analysis[3] of ^{15}N labeling data, relative isotope abundance (or "labeled population") and relative mass shift of isotope incorporation information were treated as separate identities. The relative mass shift was shown to have biphasic characteristics. The fast initial increase of the mass shift indicates the quick priming of the free amino acid compartment. This observation strongly supports the use of the "step" input function of free ^{15}N-labeled amino acids. More

dx.doi.org/10.1021/ac203330z | Anal. Chem. 2012, 84, 4014−4021

importantly, the relative isotope fraction naturally integrated both relative isotope abundance and relative mass shift into a single identity, allowing the experimental observable to be connected to compartment models in a more straightforward way. Since relative isotope fraction is directly related to the relative concentration of labeled amino acids in a compartment, the modeling strategy of this study is established on a concrete theoretical foundation.

In a typical large-scale protein turnover study, a large number of high-quality protein incorporation curves are available because each can be constructed by use of many (sometimes hundreds) peptide incorporation curves. The compartment models presented here seem able to provide optimal functional forms for fitting the experimental data. For brain proteins, the two-compartment model with only two parameters provided superior fitting compared to any two-parameter functional forms with single-compartment models, demonstrating the importance of the correct functional forms. For liver proteins, we attempted to produce a three-compartment model with a minimal number of parameters. A more general three-compartment model with five rate constants, including that for total protein export, is described in the Supporting Information. But with the five-parameter system, it was difficult to produce satisfactory fits to the experimental data because it is more sensitive to the initial parameter estimates. The four-parameter system was more robust and produced substantial improvement over the less sophisticated models.

The goal of this study is to provide compartment models that have the following two characteristics. First, parameters in the models should imply certain biological significance, such as the relative flow rate constant of free amino acids out of the free amino acid compartment. Second, the models should fit the experimental data well with a minimal number of parameters. It is the second requirement that distinguishes the more sophisticated models from the simpler ones. For example, the four-parameter/three-compartment model can provide equal or better fitting for brain proteins than the two-parameter/two-compartment model. However, the improvement of fitting is negligible. To reduce from the four-parameter/three-compartment model to the two-parameter/two-compartment model, one can simply assume the total protein compartment has negligible impact on individual protein turnover (setting the relative flow rates into and out of the total protein compartment to zero). This is a reasonable assumption since in brain tissue nearly all proteins turn over much more slowly than in liver. On the other hand, the total protein turnover can not be ignored for liver proteins. The three-compartment model is necessary to produce satisfactory fits to the experimental protein turnover data from liver, and the number (4) of parameters can no longer be reduced without significant sacrifice of fitness.

We have also compared the rate constants obtained by fitting relative isotope fraction with compartment modeling against those from fitting relative isotope abundance with the empirical delayed exponential function. In general, the correlations were strong and the ranking order of magnitude was kept in the two data sets. Therefore, the functional analysis given previously[3] remains valid. But the values of the rate constants do differ. For brain proteins, the range and values of k_b from the two-compartment model were smaller than those from the delayed exponential fit. This may result from the reduction of outliers in the two-compartment modeling, which separates the contribution of amino acid removal better than the simple delayed exponential model. This effect was more dramatic for the liver proteins. For the majority of the liver proteins, the individual protein degradation rate constants exhibited a wider range. The delayed exponential modeling seems incapable of accurately extracting individual degradation rate constants with small values for liver proteins. The compartment modeling by use of the relative amino acid concentration (proportional to the relative isotope fraction) has a disadvantage of being unable to separate the compartment size from its rate constant. For example, under the present modeling framework, a small relative isotope fraction can result from either a small portion of completely labeled proteins or all proteins labeled with a small level of labeling. Additional theoretical and informatics development is needed to explain the two extreme cases.

The current methodology, including the experimental execution and data analysis, has its limitations. An animal tissue is a complicated system, and what we are attempting to obtain is the averaged protein turnover information. The dynamics of digestion of amino acids from food and circulation delivery to cells in the targeted tissue is folded into the amino acid and protein compartment turnover rate constants. Therefore, the turnover rate constants obtained from the current modeling method can still be contaminated with some systematic errors.

■ ASSOCIATED CONTENT

Supporting Information

Additional text with differential equations and their solutions for ^{15}N-labeled amino acids in a three-compartment model and four figures showing three-compartment model for both total protein export and total protein degradation, brain and liver protein incorporation curve fitting, and histogram of total protein degradation rate constants (pdf); table of protein ^{15}N relative fraction incorporation curves along with the fitted rate constants (xls). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail sguan@cgl.ucsf.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NIH NCRR Grants RR01614 (to A.L.B.) and RR019934 (to A.L.B.) and by NIH Grants AG010770 (to S.B.P.), AG021601 (to S.B.P.), AG02132 (to S.B.P.), and AG031220 (to S.B.P.), as well as gifts from the G. Harold and Leila Y. Mathers Charitable Foundation, Sherman Fairchild Foundation, John Douglas French Alzheimer's Foundation, Lincy Foundation, and Rainwater Charitable Foundation.

■ REFERENCES

(1) Colby, D. W.; Prusiner, S. B. *Nat. Rev. Microbiol.* **2011**, *9*, 771–777.

(2) (a) Claydon, A. J.; Beynon, R. J. *Methods Mol. Biol.* **2011**, *759*, 179–195. (b) Doherty, M. K.; Hammond, D. E.; Clague, M. J.; Gaskell, S. J.; Beynon, R. J. *J. Proteome Res.* **2009**, *8* (1), 104–112. (c) Pratt, J. M.; Petty, J.; Riba-Garcia, I.; Robertson, D. H.; Gaskell, S. J.; Oliver, S. G.; Beynon, R. J. *Mol. Cell. Proteomics* **2002**, *1* (8), 579–591.

(3) Price, J. C.; Guan, S.; Burlingame, A.; Prusiner, S. B.; Ghaemmaghami, S. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (32), 14508−14513.
(4) Guan, S.; Price, J. C.; Prusiner, S. B.; Ghaemmaghami, S.; Burlingame, A. L. *Mol. Cell. Proteomics* **2011**, *10*, No. M111.010728.
(5) (a) Carson, E. R.; Cobelli, C.; Finkelstein, L. *The Mathematical Modeling of Metabolic and Endocrine Systems*; John Wiley & Sons: New York, 1983; p 394. (b) Wolfe, R. R.; Chinkes, D. L. *Isotope Tracers in Metabolic Research*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, 2005; p 474.
(6) Senko, M. W.; Beu, S. C.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (4), 229−233.
(7) Kasumov, T.; Ilchenko, S.; Li, L.; Rachdaoui, N.; Sadygov, R. G.; Willard, B.; McCullough, A. J.; Previs, S. *Anal. Biochem.* **2011**, *412* (1), 47−55.
(8) Vogt, J. A.; Schroer, K.; Holzer, K.; Hunzinger, C.; Klemm, M.; Biefang-Arndt, K.; Schillo, S.; Cahill, M. A.; Schrattenholz, A.; Matthies, H.; Stegmann, W. *Rapid Commun. Mass Spectrom.* **2003**, *17* (12), 1273−1282.
(9) Garlick, P. J.; Waterlow, J. C.; Swick, R. W. *Biochem. J.* **1976**, *156* (3), 657−663.