Protein degradation and growth dependent dilution substantially shape mammalian proteomes

Andrew Leduc,[1,✉] & Nikolai Slavov[1,2,✉]

[1]Departments of Bioengineering, Biology, Chemistry and Chemical Biology, Single Cell Proteomics Center, and Barnett Institute, Northeastern University, Boston, MA 02115, USA
[2]Parallel Squared Technology Institute, Watertown, MA 02472, USA
✉ Correspondence: leduc.an@northeastern.edu or nslavov@northeastern.edu
∈ Code & Data: github.com/SlavovLab/Protein-Clearance

Cellular protein concentrations are controlled by rates of synthesis and clearance, which includes protein degradation and dilution due to growth. Thus, cell growth rate may influence the mechanisms controlling variation in protein concentrations. To quantify this influence, we analyzed the growth-dependent effects of protein degradation within a cell type (between activated and resting human B-cells), across human cell types and mouse tissues. The results indicate that growth-dependent dilution can account for 40 % of protein concentration changes. Furthermore, we find that the variation in protein degradation rates is sufficient to account for up to 50 % of the variation in protein concentrations at slow growth as contrasted with 7 % in growing cells. Thus, our model unifies previous observations with our new results and highlights larger than previously appreciated contributions of protein degradation to protein variation at slow growth, both across proteoforms and tissue types.

Introduction

Protein concentrations vary across the proteome within a cell and across cell types and states. This variation is controlled by protein synthesis and clearance. Protein synthesis rates reflect both mRNA abundance and translation efficiency, while clearance rates include protein degradation, secretion, and dilution from cell growth. Protein synthesis rates have been been shown to be the dominant factor in a few model systems representing growing cells[1–3]. Yet, the relative importance of protein synthesis and clearance may depend on the rate of cell growth as hypothesized previously[3].

At low growth rates, proteins accumulate according to their degradation rate, producing an inverse relationship between degradation rates and protein concentrations. As the growth rate increases, this accumulation is attenuated because proteins become diluted faster than they degrade. However, reduced accumulation may not reduce protein concentration if synthesis rates compensate by preferentially increasing in inverse proportion to degradation rates[4]. Although significant compensation has been demonstrated, it is not yet clear to what degree changes in protein synthesis fully counterbalance the shift from degradation-driven clearance to dilution-driven clearance.

Furthermore, the significance of this compensation, or lack there of, depends on the degree of variability in degradation rates. Simply put, if all proteins have relatively similar degradation rates, there would be minimal potential for preferential accumulation of long lived proteins. To measure degradation as well as synthesis rates across the proteome, liquid chromatography tandem mass spectrometry (LC-MS/MS) combined with metabolic pulse labeling has become the tool of choice, both in vitro and in vivo[5]. Such measurements are greatly facilitated by recent advances, such as plexDIA[6,7]. Metabolic pulse studies have reported that protein specific degradation rates can span three orders of magnitude[1,3,8–11]. This suggests that changes in cell growth require protein synthesis to compensate substantially if protein concentrations are to be maintained.

In addition to variation between different proteins, studies have reported significant differences in degradation rates for the same protein in different cell types[9] and tissues[11]. This variation has been shown to shape the abundances of specific complexes and functional groups[11]. However, the extent to which degradation or dilution plays a significant role in shaping concentrations proteome

wide remains unclear.

In this study, we establish that clearance rates substantially determine protein concentrations across the proteome and across cell states and tissues. We first explore the influence of protein degradation within proliferative and several non-proliferative cell types. In non-proliferative cell types, we find that between 28 and 52% of protein concentrations are determined by degradation, substantially larger than previous reports. These results confirm that counteracting the effects of growth rate changes on protein concentration would require substantial compensation in synthesis rates. We calculate these compensation rates in resting versus activated B-cells, finding primarily that synthesis does not adequately compensate. We show that degradation rate specific protein dilution explains 37 % of the variance in relative protein concentrations between activated and resting B-cells.

We then establish the relevance of this phenomenon in shaping tissue specific proteomes via in vivo metabolic pulse labeling. We again find that the average rate of cell growth within tissue strongly predicts the amount of variation in protein concentrations explained by clearance. Furthermore, we show that lack of compensation for dilution in rapidly proliferating bone marrow cells explains 34 % of concentration difference between bone marrow and brain tissue. We find that when comparing slowly dividing tissues with little dilution, the brain and pancreas, differences in degradation rate explain as much as 42 % of the variation in protein concentrations. Lastly, we analyzed differences in protein concentration across all tissues and find that on average, clearance explains approximately 25 % of the variation in concentration across all tissues. Further, we identify many individual proteins and pathways that are primarily controlled by protein clearance.

Results

The impact of degradation on protein concentrations varies across cell types

We first sought to quantify the contribution of degradation to setting protein concentration across the proteome in variety of cell types, including several that were not undergoing significant growth. To estimate the influence of degradation, we must first define its relationship to protein . In non-

growing cells, the steady-state concentration for the *i*th protein (p_i) equals the ratio of the rates of protein synthesis $k_{s,i}$ and protein degradation or secretion $k_{d,i}$, Fig. 1a. In growing cells, an additional term is needed to account for dilution of proteins due to the increase in volume that occurs as cells grow, k_g[4,12], Fig. 1a. Thus, protein concentration is inversely proportional to the overall amount of protein clearance $k_{r,i}$, the combined influence of degradation and dilution, Fig. 1a. As synthesis and degradation rates are well approximated by the log-normal distribution[13], Supplemental Fig. S1a, we define the relative contribution of degradation as the sum of squared residuals explained by K_r in the following linear regression: $\log(P) = \log(K_s) - \log(K_r)$. This approach is more robust than the square of the Pearson correlation, which is agnostic of the regression coefficient and thus can overestimate when the rates of synthesis and degradation are correlated Supplemental Fig. S1b.

Applying this approach to metabolic pulse labeling LC-MS/MS data, we estimated R^2 to quantify the relative contributions of degradation and protein synthesis in proliferative cells from Schwanhausser *et al.*[1] and non-proliferative cell types from Mathieson *et al.*[9] Remarkably, R^2 estimates for the fraction of variation explained by degradation span a wide range, varying from 7 % to 62 %. It is imperative to consider and account for the influence of measurement noise on these estimates. Systemic biases, such as ionization efficiency and digestion efficiency, cannot be accounted for with replicate measurements. These factors predominantly affect the estimation of protein abundance and synthesis rates. They affect similarly light and heavy peptides and thus cancel out in the heavy to light ratios used for estimating degradation rates. Therefore, these systemic biases are not shared between the estimates of protein degradation and abundance and unlikely to inflate their correlations. For a more in depth discussion of handling measurement noise, see methods.

Degradation contributes most significantly to protein concentrations in non-growing natural killer cells, even exceeds the influence of protein specific synthesis rates, Fig. 1b. In rapidly growing cells, degradation only accounted for 7% of the protein concentration variation while protein synthesis explains 93 % Fig. 1c, aligning with previous estimates made from the same data and two additional proliferative cell lines[1]. Conversely, in non-proliferative cells, degradation explained over 28 % of the variance in all cases, Fig. 1c. Within non growing cell types, the R^2 varied

from 28% to 62%, highlighting the influence of additional factors, such as the overall variation of synthesis and degradation rates across different proteins.

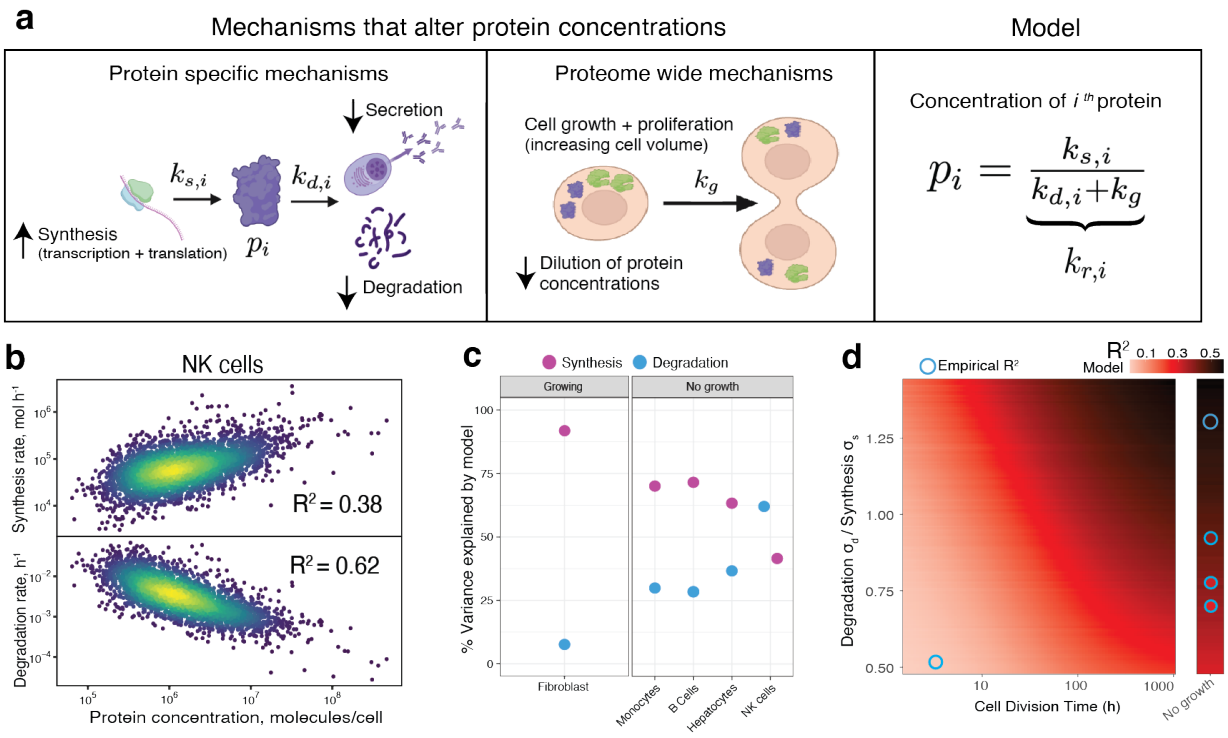


Figure 1 | Modeling the impact of degradation on protein abundance in growing and non growing human cells. a, Protein concentration (p_i) is determined by protein specific factors such as synthesis ($k_{s,i}$) and degradation / secretion rates ($k_{d,i}$) as well as global factors such as the rate of dilution caused by cell growth (k_g). **b**, Data from non-proliferative natural killer collected by Mathieson *et al.*[9] show strong dependence of protein concentration on rates of degradation, $R^2 = 0.62$, suggesting that protein degradation account for 62 % of the variation of protein concentrations. Protein synthesis rates due to transcription and translation account for 38 % of the variation in protein concentrations in natural killer cells. **c**, The % variance explained by degradation and synthesis varies for growing cells, fibroblasts collected by Schwanhauser *et al.*[1], and additional non-growing cell types collected by Mathieson *et al.*[9]. **d**, Model predictions for the variance of protein concentrations explained by degradation as a function of cell division time and the ratio of standard deviations for the log normally distributed rates of protein degradation (k_s) and synthesis (k_d). Predictions align with empirical estimates derived via regression.

To investigate additional factors and evaluate if growth can account for the magnitude of difference between growing and non-growing cells, we sought to model the variation of synthesis and clearance rates across conditions. Specifically, we modeled the synthesis and degradation rates from a bi-variate log-normal distribution. The extent of co-variation between synthesis and degradation was modest in the samples we analyzed, Supplemental Fig. S1c. We speculate that the negative correlation present only for proliferative fibroblasts may reflect the compensation in

protein synthesis rates for long lived proteins, as increased synthesis is needed to make up for the decreased accumulation of proteins with slower degradation rates. However, because co-variation was modest, we chose to set the co-variation to 0 falling between estimates for growth and non-growth.

Using this model, protein concentrations were then computed from the ratio of simulated synthesis and clearance rates, where clearance rates equal the degradation rate plus the growth rate dilution constant, Fig. 1a. We simulated data through a physiologically relevant range of growth rates and variances for the distribution of synthesis and degradation rates, Supplemental Fig. S1d. The resulting R^2 estimates closely match empirical estimates across a wide range of parameters (Supplemental Fig. S1d), thus suggesting explanations to both questions raised from the results in Fig. 1c. First, we observed that degradation rates have cell-type specific degrees of variability. As a result, the influence of degradation differed across non-growing cell types, Fig. 1d and Supplemental Fig. S1d. Secondly, the results confirmed that the level of growth in the proliferative fibroblasts could indeed quantitatively explain the differing amounts of variance explained by degradation in growth and non-growth conditions Fig. 1d.

Degradation rates exert strong influence on protein concentrations at slow growth. This observation implies that longer-lived proteins may have higher abundance under slow growth conditions. However, this effect may be counteracted if synthesis rates increase proportionally for long-lived proteins at high growth. Indeed, the inverse correlation between synthesis and clearance rate present only in the fast growth condition suggests that this compensation occurs to some degree. This possibility is exemplified by the increase in histone synthesis during S-phase of the cell division cycle, thus allowing cells to maintain histone concentrations. However, it is less clear whether such compensatory changes in protein synthesis rates fully counteract reduced accumulation across the proteome.

Incomplete compensation for protein dilution upon B-cell activation

Thus, we sought to determine the extent to which growth rate related changes in protein dilution are compensated by changes in protein synthesis. To this end, we first focused on growth rate changes within a cell type, namely primary B-cells in the resting state (non proliferative) and activated

state (highly proliferative). We utilized concentration measurements made by Rieckmann *et al.*[14] and degradation rate measurements made in resting primary B-cells by Mathieson *et al.*[9]. This system was ideal for several reasons. First, without stimulation, B cells can persist for months in vivo without dividing[15]. However, upon stimulation, B-cells can divide as rapidly as every 16 hours, providing a large growth rate differential[16], Fig. 2a. Additionally, degradation rates are likely to be more similar within as opposed to across cell-type. Thus, measuring changes within a cell type provides a relatively controlled comparison. Additionally, replicate measurements of protein abundance and protein clearance allowed us to estimate and account for random noise in the estimation of protein fold changes and protein clearance, see methods.

As growth rate increases, the production of all proteins must increase. We define this increase uniformly for each protein as a constant δ. Deviations in synthesis from the expected scalar for a given protein *i* are defined with an additional term μ_i. Thus the change in synthesis upon an increase in growth rate can be defined as $\Delta k_{s,i} = \delta\mu_i$. If δ explains all increases in protein synthesis, $\mu_i = 1$ for all proteins, changes in clearance caused by dilution will alter protein concentrations. To demonstrate why, in non-growing B cells, protein clearance equals protein degradation rate and spans a dynamic range from 10^{-4} to 10^{-1}, Fig. 2b. As uniform protein dilution outpaces degradation in activated B cells, clearance rates span only 10-fold dynamic range, Fig. 2b. Thus, with only a uniform scalar change in synthesis rate, protein concentrations will decrease upon activation proportionally to the change in clearance, Fig. 2b. This scenario represents an absence of compensation.

Note that our minimal model for clearance changes upon B-cell activation focuses only on changes caused by growth increase. Because dilution dominates degradation at high growth, changes in degradation rates associated with activation will not substantially affect estimated changes in clearance, except potentially for the least stable proteins. Further, any substantial changes in degradation would reduce the explanatory power of our model, which assumes that degradation rates remain constant between conditions. Thus, our model will report conservative estimates for the impact of growth induced clearance changes. Further, note that measurements come from different labs and variation between the B cells (if any) will result in misaligned degradation rates and protein fold change estimates. Again, this will lead to an underestimation of

percent variance explained by our model as these differences would attenuate the trend.

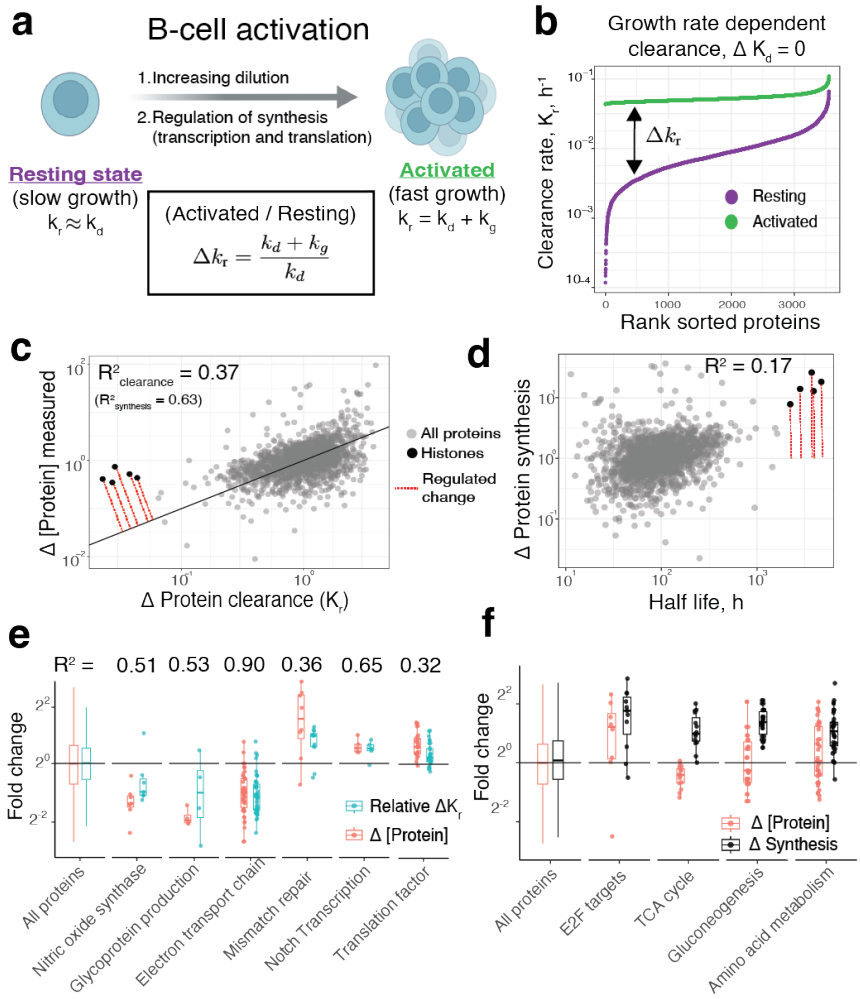


Figure 2 | Proteome differences between resting and activated B-cells originate from altered protein dilution. **a**, When B cells are exposed to antigen, they transition from a resting state where there is minimal cell growth, to an activated state where cells divide every 16 hours[4]. **b**, Thus, clarence rates increase for all proteins with k_g, resulting in larger relative changes for more stable proteins. **c**, The relative changes in clearance rates between the activated and resting state explains 37 % of the variance in protein concentration changes between these two states. **d**, Functional protein groups driven by dilution include the electron transport chain where 90 % of protein fold changes are caused by increased protein clearance due to growth. **e**, However, not all proteins change concentration proportional to the shift in clearance. Many proteins such as histones are actively regulated with increased protein synthesis tied to the rate of cell division. **f**, Proteins involved in processes such as gluconeogenesis are actively regulated to increase their abundance to counteract the increased rate of clearance.

When comparing concentration changes and expected clearance changes, we find that change in clearance caused by protein dilution alone explains 37 % of the variation in protein concentrations between the activated and resting state, Fig. 2c. This demonstrates that compensation incompletely accounts for a significant proportion of changes in protein clearance. However, pro-

teins such as histones notably defy this trend and maintain their concentration. In these cases, synthesis rates are regulated to compensate and prevent significant concentration decrease. We interpret these deviations from the expected growth induced dilution as regulated changes in synthesis, μ_i. Regulated changes correlate more weakly with protein half life explaining only 17 % of the variation in regulated changes. This suggests that the majority of changes in synthesis do not specifically work to counteract dilution, Fig. 2d.

Because dilution induced concentration changes are largely not compensated for, this raises the question of whether this dilution contributes to defining the phenotype of the activated state. Towards this direction, proteins from a number of concerted pathways change their concentration substantially between the activated and resting state. Many of these pathways can attribute over half the variance or more to the changes in protein clearances caused by the change in dilution, Fig. 2e. For example, proteins participating in the electron transport chain decrease their concentration by two fold and are predominantly driven by dilution. This may reflect the decrease in metabolism from oxidative phosphorylation to glycolysis during activation[17]. Other proteins weakly affected by dilution such as translation factors and proteins participating in notch transcription increase in concentration, Fig. 2e. This may reflects the increased demand for protein synthesis.

Conversely, examining the extent of active regulation reveals several pathways for which changes in synthesis rate exceed changes in protein concentration. These pathways include the E2F transcription pathway as well as proteins participating in gluconeogenisis, Fig. 2f. Regulation of metabolic enzymes participating in glucose metabolism may be required to meet the energy demands of increased proliferation.

Growth and degradation rate differences shapes tissue-specific proteomes

To test whether differences in compensation for protein dilution plays a role in additional contexts, we collected in vivo pulse labeling samples from four tissues with varying growth rates, Fig. 3a. These samples included the brain, one of the tissues with the lowest reported cellular turnover, as well as the bone marrow, one of the fastest renewing systems in the body[18]. Samples were collected from 10 different mice that span both genders and two different age ranges, 4 and 20 months.

In order to draw accurate conclusions from the data, we sought to account for measurement noise. To this end, we estimated concentration and clearance rates from samples digested by either LysC only, or LysC plus trypsin. Samples digested by different proteases generate different peptides and thus share less systematic bias from factors such as ionization and digestion efficiency compared to technical replicates[19]. In both digests, we were able to quantify concentrations and degradation rates for over 5,000 proteins via data independent acquisition (DIA) mass spectrometry, Supplemental Fig. S2a. LysC measurements from the pancreas were not considered due to the large number of tryptic peptides detected, which we attributed to trypsin expressed in the tissue.

The quantitative accuracy of the measurements is reflected in the agreement between digest conditions. For each tissue, we compared protein concentrations from digest conditions and observed strong agreement in line with expectations from LysC and Trypsin proteases[19], Supplemental Fig. S2b. For clearance rate measurements, to further mitigate the influence of noise, we filtered estimates within each digest and tissue for proteins with a coefficient of variation (CV) below 20 %. This set contained roughly 2,000 proteins for each tissue. We then compared the consistency between digests for the intersected subset of low CV observations Supplemental Fig. S2c. We observed good consistency for brain and lung measurements, $\rho = 0.82$ and $\rho = 0.78$ respectively. The consistency was notably lowest for bone marrow. This is primarily due to the small dynamic range in clearance rates which are primarily driven by dilution. Lastly, doubling rates for each tissue were estimated under the assumption that the histone degradation rates were negligible[20,21]. Therefore, we defined k_g for each sample as the degradation rates of histone H3 and H4. We observed high consistency between estimates made from each digest condition, Supplemental Fig. S2d. The average population doubling times ranged from approximately 1 day for bone marrow to 200 days for some brain samples.

Having observed a broad range of doubling times, we wanted to examine how doubling times affected the relationship between degradation and protein concentration. Consistent with our earlier results, clearance rates weakly contribute to protein concentrations in rapidly proliferating cells such as the bone marrow. Conversely, clearance contributed up to 50 % of concentration variation in slowly dividing tissues such as the brain. These estimates were corrected via the spearman correction based on the consistency of the replicates from different digests. When looking across all

tissues, we observed a continuous trend with a strong correlation, $\rho = 0.89$, Fig. 3b. To directly attribute this to growth, we computed the average variability in degradation rates and clearance rates for each tissue, Fig. 3c. We observed that while degradation rates are similarly variable for brain and bone marrow, the clearance rates are significantly less variable for the bone marrow. This difference can only be driven by the significant dilution in the bone marrow, reducing protein to protein differences in clearance. In brain tissues, the variability in clearance and degradation rates are about the same, as degradation is the primary mechanism of clearance.

We then set out to determine how these growth rate differences shape clearance and concentration differences across tissues. First, we plotted clearance rates for the brain against the bone marrow, Fig. 3d. We observed that while clearance rates were correlated, $\rho = 0.40$. The slope estimated by total least squares is shallow, 0.13, indicating reduced variability in clearance rates for the bone marrow. This reduced variability is consistent with the faster cell growth in bone marrow, and thus larger contribution of the constant dilution term to the clearance rates.

Despite the shallow slope, the correlation of the clearance rates defines a clear axis of similarity along the line passing through the data points. This axis captures the shared variance in clearance rates, which we termed a common component. The tissue-specific differences in degradation rates are reflected by the variance orthogonal to the common component, and we term it specific component. Mathematically, the common component and specific component correspond to the principal components of the data, PC1 and PC2. The common component reflecting shared clearance variation explains the majority of the variance, 84% on PC1. When we compare the brain to the lung, we see that the slope is less shallow, 0.80, reflecting the smaller impact of dilution for the lung Fig. 3e. Again, the common component explains the majority of the variance in clearance rates.

Although the common component explains the majority of variance in both tissue comparisons, its relationship to clearance and concentration differences between the tissues differs. In tissues with a large growth rate differential such as the brain and bone marrow, the common component explains 76 % of the clearance differences Fig. 3f. As these differences reflected the axis of shared variation in degradation rates, they are directly attributable to dilution differences. Similar to observations made in resting and activated B-cells, we found that the common component explains a substantial fraction of concentration differences between the brain and bone marrow, $R^2 = 0.34$,

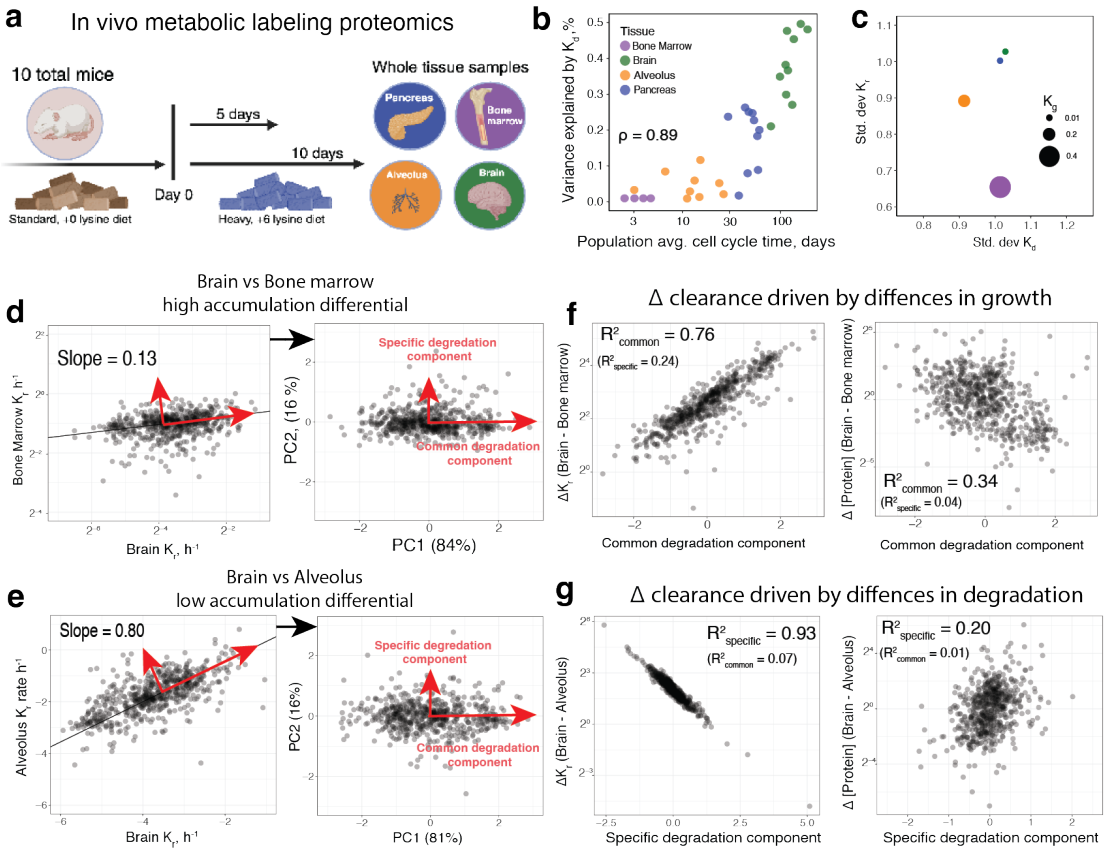


Figure 3 | Tissue-specific protein dilution and degradation rates explain protein abundance variation across tissues a, Protein clearance and synthesis rates were measured from murine tissue samples using in vivo metabolic pulse labeling with heavy K. **b**, The lower the growth rate of a issue, the more of its protein abundance variation is explained by degradation rate differences across the proteome. **c**, Variability in degradation rates and clearance rates for each tissue. Growth rates are reflected by size of points and explain differences between degradation and clearance variability for alveoli and bone marrow. **d**, When plotting clearance rates for brain against bone marrow, the shallow slope indicates reduced variance in bone marrow clearance rates due to high dilution in bone marrow. We then decompose clearance rate trends into a common component reflecting the shared axis of variation in protein degradation rates and a specific component reflecting differences in degradation. **e**, These trends are also shown in comparison between lung and bone marrow where the growth differences are less significant and thus the slope is less shallow. **f**, When growth differences are substantial, the common degradation component drives clearance differences. This in turn contributes to concentration differences between the tissues. **g**, When growth differences are small, the specific degradation component drives clearance and concentration differences between tissues.

Fig. 3f. This again suggests that cells do not fully compensate for the reduction in protein accumulation in fast growth states. However, when tissues have similar clearance rate variability, the specific degradation component explains the majority of clearance rate differences, Fig. 3g. Here, the specific component driven by degradation rate differences explains 20 % of protein concentration differences between the brain and lung Fig. 3g.

Differences in clearance driven by degradation are most pronounced when comparing rates for the two tissues with the lowest growth rate, brain and pancreas. Here, we observed a modest correlation between clearance rates, $\rho = 0.38$, demonstrating significant degradation rate differences,Fig. 4a. Protein fold changes between the pancrease and brain were strongly predicted by clearance differences, $\rho = 0.65$, suggesting that clearance explains 42 % of the tissue-specific protein differences between these tissues, Fig. 3b. This trend was the strongest for any pairwise comparison, Fig. 3e.

We next sought to examine the extent to which variations in the concentrations of individual proteins across samples could be explained by degradation variation. We identified many proteins whose abundance could be mostly explained by degradation, $R^2 > 0.70$, Fig. 3f, supplemental table 2. When examining the distribution across all proteins, we found an average correlation of 0.31. When using the spearman correction factor to account for the influence of measurement noise via the consistency of LysC and trypsin replicates, the average correlation increased to 0.50. To elucidate this trend, it was important to control for average differences in clearance rates. Clearance rates for all proteins will be higher in bone marrow compared to brain, but it is the differences relative to the average clearance that will impart specificity to protein concentrations.

Additionally, we observed that protein variation for several pathways is substantially controlled by protein degradation. Increased abundance in the brain of proteins from pathways such as Neurotrophilin signaling and WNT5A signaling can be almost entirely explained by their slower degradation rates, Fig. 3f. Several other pathways such as serotonin and insulin recyptor recycling are also up regulated in the brain. Other pathways with significant control by degradation such as Toll like receptor 9 cascade and interleukin 17 signaling were upregulated in the bone marrow. Overall we identified 141 pathways whose abundance is controlled by degradation at more than 20%, supplemental table 2.

Discussion

Our analysis and data indicate that protein degradation have comparable impact to protein synthesis at slow growth. This impact stems from the fact that proteins accumulate inversely proportional

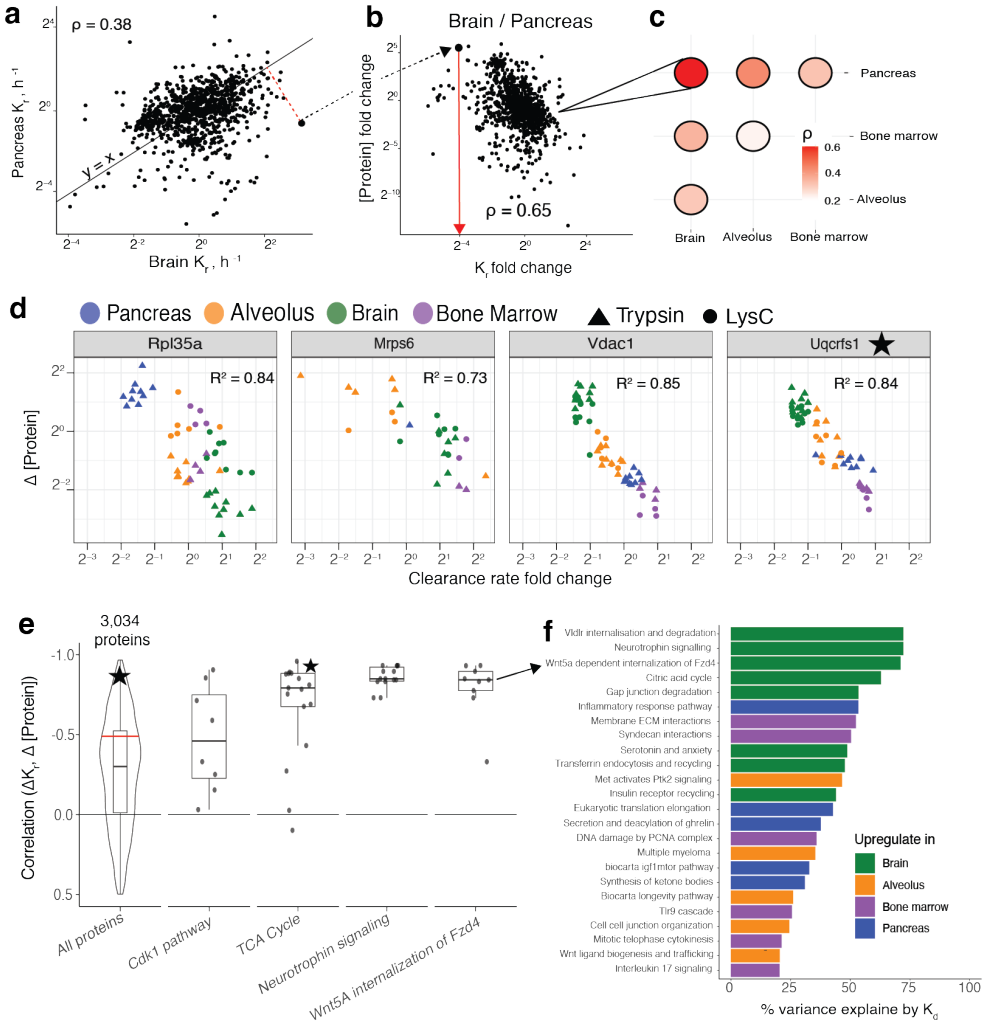


Figure 4 | Tissue specific concentrations of many proteins are primarily determined by their tissue specific clearance rates. a, Protein clearance rates for pancreas and brain plotted against each other show modest correlation. **b**, Differences in clearance rates between brain and pancreas explain differential protein concentrations, Pearson correlation $\rho = 0.65$. **c**, Similarly, differences in clearance rates explain differences in concentrations between other tissue types to varying degrees. **d**, For some proteins, abundance across tissues is almost entirely explained by variation in their clearance rates. **e**, On average, the correlation between clearance rates and correlation profiles is 0.35. When correcting the correlation for measurement for consistency of lysC and trypsin digests via the spearman correction, the average is 0.50 depicted by the red line. The star depicts the correlation of Uqcrfs1 shown in panel **d**. This protein participates in the TCA cycle, whose enzymes are primarily regulated by degradation. **f**, GO terms with significant contribution of degradation rates are summarized along with what tissue the pathway proteins are more abundant.

to their rate of degradation. In growing cells, however, protein accumulation is reduced as dilution becomes a major source of protein clearance. To prevent the reduction in accumulation from altering protein concentrations at fast growth, synthesis rates must compensate and increase proportionally to the reduction in accumulation. In proteins like histones, such compensation almost

fully offsets the reduction in accumulation Fig. 2c,d. For other proteins we observe only partial compensation, consistent with previous reports emphasizing the role of compensation[4]. However, we find that this partial compensation does not offset the majority of changes in clearance rates caused by increased dilution. Our measurements and analysis quantify this effect for the first time for thousands of proteins in physiologically relevant conditions. Thus, we establish the role of clearance differences caused by dilution in shaping cell type and tissue specific proteomes.

We find that this phenomenon explains a substantial amount of variation in protein concentrations. It contributes 37% of concentration change between resting and activated B-cells (Fig. 2) and 34 % of the concentration changes between brain and bone marrow tissues (Fig. 3). This raises the interesting possibility that long lived proteins are generally less needed or inhibitory for high growth states and more important for low growth states. While it has been posited that certain proteins such as those involved in signaling require high turnover to allow cells to adapt quickly to environmental stimulus[5], our observation provides potential explanation in the opposite trend. Proteins crucial for slow growth such as the proteins in the electron transport chain may have preferentially evolved to have longer half lives.

We also report large and context-dependent roles of protein degradation both across cellular states and across different proteins. While the possibility for such extensive regulation has been suggested by protein variation that is unexplained by mRNA variation[22–25], our analysis focused on explaining protein variation by measured protein degradation rates. The results emphasize a larger than previously described role for protein degradation in determining protein abundance variation. They also explains the smaller roles reported previously in fast growing cells in number[1] or in cell size[3]. Our results did not include shrinking cells, such as erythroblasts transforming into erythrocytes[26], but they predict that protein degradation is likely to play an even more dominant role. Our data demonstrate that variation in protein degradation rates is an important global mechanism for shaping protein concentration, Fig. 4. This large contribution helps explain the high abundance of stable proteins resulting from alternate RNA decoding[27]. Together, our results motivate future work on understanding the biological impacts of protein degradation across diverse contexts.

Our tissue analysis faces limitations, such as averaging across the diverse cells making up the analyzed tissues. For example, tissues like the brain comprise non-proliferative differentiated cells

and proliferative cells. This complicates interpretations as it is challenging to disentangle the exact amount of dilution each quantified protein faces. Single-cell measurements of degradation, growth, and concentration rates in vivo will add additional context to understanding when synthesis rates must compensate for protein dilution.

Resource availability:

Lead contact: Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Andrew Leduc (leduc.an@northeastern.edu).

Materials availability: This study did not generate new unique reagents.

Data availability: All raw data and search engine outputs can be found on MassIVE with ID: MSV000097050. Additional processed data needed for reproducing the analysis can be found on Zenodo under DOI 10.5281/zenodo.14827609.

Supplemental data table 1

Supplemental data table 2

Code availability: The code used for data analysis and figures is at: github.com/Slavovlab/Protein-Clearance

Acknowledgments: We thank Yanxin Xu and Bin Zhang for help with mouse tissue preparation, Zhixun Dou for help procuring mice, and Alexander Franks for advice on statistical analysis. The work was funded by a Bits to Bytes award from MLSC to N.S., an NIGMS award R01GM144967 to N.S., and a MIRA award from the NIGMS of the NIH (R35GM148218) to N.S, and a UH3CA268117 award from NIH to N.S.

Competing Interests: N.S. is a founding director and CEO of Parallel Squared Technology Institute, which is a nonprofit research institute. The authors declare that they have no other competing interests.

Correspondence: Correspondence and requests should be addressed to nslavov@northeastern.edu, nslavov@alum.mit.edu

Author Contributions

Experimental design: A.L. and N.S.

Sample preparation: A.L

Raising funding & supervision: N.S.

Data analysis: A.L. and N.S.

Writing & editing: A.L. and N.S.

Methods

Mouse model and handling

All mice experiments were performed in compliance with the Institutional Animal Care and Use Committee at Massachusetts General Hospital. Both male and female C57BL/6 mice, either 24-month-old or 4-month-old, were ordered from the NIA. Mice were euthanized with CO2 followed by cervical dislocation. The mouse used was male. Tissues were harvested post-euthanasia and perfusion with PBS.

Metabolic pulse labeling and tissue collection

Mouse Express L-LYSINE [13C6, 99%] MOUSE FEED kit was purchased from Cambridge Isotopes Laboratories. Mice were first acclimated to the Lys+0 feed for a period of 10 days before switching to the Lys+6 diet for either 5 or 10 days. Mice were fed 5g per day for either 5 days or 10 days before euthanisation. Whole tissues including the pancreas, brain, aveolus, and bone marrow were harvested a total of 10 mice. Descriptions of the age, gender and feeding time for each mice can be found in the meta data file on Zenodo. After harvesting, tissues were stored at -80 until prepared for LC-MS/MS analysis.

Sample preparation for MS

Samples were prepared based on an optimized protocol[28]. Briefly, whole tissues were thawed from -80 degree freezer at room temperature in 1 mL of 1x PBS. Samples were then ground using 1 mm diameter glass beads and a tissue homogenizer (BioSpec Mini-Beadbeater). Samples were then brought to 1% SDS and 1x benzonase to break apart DNA. Samples were then brought to 5% SDS and boiled for 10 minutes. 10 micrograms of protein from each sample were then added to SP3 beads for clean up and overnight digestion at 37 degrees C in 20 ng/μl LysC. LysC digest was then split into one fraction for LysC analysis and another fraction for a second overnight digest with

Trypsin at 37 degrees C in 10 ng/μl Promega Trypsin Gold. Samples were then dried down and resuspended at a concentration of 750 ng/μl in 0.1 % formic acid for injection to the LC-MS/MS.

LC-MS/MS data acquisition

All samples were were separated via online nLC on a Vanquish Neo UHPLC using a 25cm x 75 μl IonOpticks Aurora Series UHPLC column (AUR2-25075C18A). Samples were resuspended at a concentration of 500 ng/μl and 1 μl volumes were injected out of glass inserts (ThermoFisher 60180-509) sealed by rubber caps (ThermoFisher C5000-51B).

All samples were run on a 95 minute total length method (53 minutes active gradients). 0.1 % formic acid in water was used for Buffer A and a 20% solution of 0.1% formic acid mixed in 80 % acetonitrile was used for buffer B. A constant flow rate of 200nl/min was used throughout sample loading and separation. Samples were loaded onto the column for 20 minutes at 1% B buffer, then ramped to 5 B buffer over two minutes. The active gradient then ramped from 5% B buffer to 25% B buffer over 53 minutes. The gradient was then ramped to 95% B buffer over 2 minutes and stayed at that level for 3 minutes to remove excess peptide and protein residue from column. The gradient then dropped to 1% B buffer over 0.1 minutes and stayed at that level for 4.9 minutes for the re-equilibration.

All samples were analyzed by a Thermo Scientific Exploris 480 mass spectrometer. Electrospray voltage was set to 1.8 V, applied at the end of the analytical column. The temperature of ion transfer tube was 250 ^oC and the S-lens RF level was set to 80. MS1 scans had the following parameters: 140k resolution, 3e6 AGC target, and a scan range from 450 to 1258Th. Two sets of DIA windows were used: 21 20Th-wide windows (spanning the space from 450Th to 860Th) and 8 50Th-wide windows (spanning the space from 859.5Th to 1256Th). The DIA windows had the following characteristics: 35k resolution, AGC target of 5e5, maximum injection time determined automatically, fixed first mass of 200Th, NCE of 27, and a default charge state of 2. DIA windows spanned the space from 450Th to 1256Th and included a 0.5Th overlap.

Interpreting raw mass spectra

To reanalyze of data dependent acquisition MS data from Mathieson et al.[9], proteomics data from SILAC-labeled human cell lines was downloaded from the publication data repository (B cells: PXD008511, hepatocytes: PXD008512, monocytes: PXD008513, NK cells: PXD008515). Samples were searched by MSFragger via the SILAC3 workflow and MSBooster workflow[29] against the human protein sequence fasta downloaded from Uniprot in September 2024.

Label free DIA runs of cell cycle fractions were searched with DIA-NN v1.9.0[30] using an in silico fasta generated library enabled by deep learning. Raw files were searched together with match between runs (MBR). DIA-NN search settings: Library Generation was set to "IDs, RT, & IM Profiling", Quantification Strategy was set to "Peak height", scan window = 1, Mass accuracy = 10 ppm, and MS1 accuracy = 5 ppm, "Remove likely interferences", "Use isotopologues", and "MBR" were enabled. SILAC was set as a fixed modification on lysine residues using the plexDIA module, and two channels were defined, one at 0 da and one at 6.051984 da.

We also performed a variable modification search in DIA-NN with the variable SILAC modification specified at 6.051984 da to identify partially labeled peptides that contained multiple lysines due to a missed cleavage. All other settings were the same other than the use of the fixed SILAC modification and channel command.

Downstream data analysis

Analysis of in vitro external data sets

Protein clearance rate data for fibriblasts, B-cells, NK cells, Monocytes, and Hepatocytes was downloaded from supplemental files of Schwanhausser et al.[1] and Mathieson et al.[9]. Estimates of protein clearance rate are proportional to the ratio of heavy over light peptide intensities, and thus accurate estimation of protein clearance do not depend on accurate knowledge of cell growth rate. For fibroblasts, protein copy numbers were also downloaded from the supplement of Schwanhausser et al. For the other cell types, we computed protein intensity from the fragpipe searched data by first summing the intensities of light and heavy SILAC precursors. We then estimated pro-

tein abundance as the median of the top 3 most abundant precursors for a protein. Random noise in the measurements was accounted for using the Spearman correction and the reliability was estimated by correlating replicate measurements as previously described[22]. We did not explicitly correct for sources of systematic bias such as ionization efficiency and digestion efficiency, though our use of both trypsin and lysC digests affords many peptides and naturally diminishes the impact of such biases. These biases are shared for estimates of protein abundance and synthesis rates and less shared for estimation of the heavy to light ratios used to compute clearance rates. This is because ionization and digestion biases are likely to effect both heavy and light ions equally. Thus, our estimation of the relative influence of synthesis and degradation conservatively estimates the influence of degradation.

Estimating clearance and cell division rates in vivo

To estimate the change in the fraction of light lysine in the unbound lysine pool over time, $\gamma(t)$, we first calculated the fraction of heavy lysine in the free pool at our two time points. To do this, we utilized the average ratio of partially labeled to fully labeled peptides as done by Jovanovic et al.[3]. We then fit an equation of the form $\gamma(t) = 1 - 0.5e^{-at} - 0.5e^{-bt}$ to estimate the change in available light lysine over time as previously developed and used in ref.[10].

Having solved for the change in the free lysine pool over time, we then formulated the change in any given light peptide over time $\frac{dL_i}{dt}$ via the equation:

$$\frac{dL_i}{dt} = k_{s,i}\gamma(t) - k_{r,i}L(t)$$

We then integrate this equation to solve for $L_i(t)$ for the initial condition $L(0) = L_0$ where L_0 is defined as the sum of heavy and light intensities under the steady state assumption that overall protein concentration is not changing over time. Integrating and further substituting $k_{s,i}$ via the steady state equality $L_0 = \frac{k_{s,i}}{k_{r,i}}$This gives the following equation:

$$L(t) = k_{r,i}L_0\left(\frac{0.5}{k_{r,i}-a}e^{-at} + \frac{0.5}{k_{r,i}-b}e^{-bt}\right) + \left(L_0 - k_{r,i}L_0\left(\frac{0.5}{k_{r,i}-a} + \frac{0.5}{k_{r,i}-b}\right)\right)e^{-k_{r,i}t}$$

The arithmetic was performed in Mathematica to obviate manual integration errors. Lastly, this equation was then solved for via a non-linear solver to find $k_{d,i}$ for each peptide from each sample.

Math for model of clearance upon growth

For the following derivations, are more comprehensively derived in Baum *et al.*[4]. Briefly, in cells that are not growing, under the steady state assumption protein concentrations can be formulated as a balance of rate of synthesis and degradation.

$$p_i = \frac{k_{s,i}}{k_{d_i}}$$

In cells that are growing, the steady state model can be maintained under the assumption that protein concentrations remain constant as cells double the amount of all proteins at an equal rate. However, under this regime, we need to introduce an additional term to reflect protein clearance due not only to degradation, but to dilution as well.

$$p_i = \frac{k_{s,i}}{k_{d,i} + k_g}$$

The dilution rate k_g is constant for all proteins and is inversely proportional to the rate of cell division, T_{CC}, $k_g = log(2)/T_{CC}$. For both models, we define the overall amount of protein clearance from degradation and dilution is $k_{r_i} = k_{d_i} + k_g$, with $k_g = 0$ in the absence of growth.

When comparing a rapid change from slow to high growth, we have to define the change in both clearance and synthesis rates that contribute to protein concentrations. The difference in clearance rates for all proteins ΔK_r is defined as:

$$\Delta k_{r,i} = \frac{k_{d,i} + k_g}{k_{d,i}}$$

This assumes no significant changes in degradation rates as a cell increases its growth rate. While this assumption may not hold, because the influence of degradation is diminished in the fast growth regimes, the ability of degradation rate changes to compensate for dilution is limited.

The change in synthesis rate has two components. First, there is the overall increase in synthesis

that is applied to all proteins, δ uniformly to account for the increase in average clearance rate from dilution. Secondly, there is the protein specific change in synthesis, μ_i, that accounts for deviations from the scalar increase, and may compensate for the decreased accumulation of long lived proteins. The final equation describing the difference in clearance

$$\Delta k_{s,i} = \frac{\delta \mu_i k_{s,i}}{k_{s,i}}$$

Relative contributions of synthesis and degradation to protein levels

To determine the relative contributions of synthesis rates (K_S) and degradation rates (K_d) to protein levels (P), we log-transformed the equation $P = \frac{K_s}{K_r}$ to obtain $\log(P) = \log(K_s) - \log(K_d)$. The variables $\log(K_s)$, $\log(K_r)$, and $\log(P)$ were then mean-centered to remove any intercepts and isolate the variability. We then calculated the relative contribution of either term as the sum of squared residuals from a leave-one-out regression. Further, we corrected R^2 values for the reliability of the data using the Spearman correction as previously described[22]. Briefly, the fraction of variance explained was divided by the reliabilities of each measurement, concentration and degradation rate. The reliabilities were estimated as the correlations between different measurements, such as estimates derived from biological replicates. In the case of the mouse tissues, digestions by lysC and trypsin were also used for reliability estimation.

Simulation Description

We simulate the steady-state protein concentrations $[P]$ using the following equation:

$$p_{\mathrm{i}} = \frac{k_{\mathrm{s,i}}}{k_{\mathrm{d,i}} + k_{\mathrm{g}}}$$

where:

- $k_{\mathrm{s,i}}$ is the protein-specific synthesis rate,
- $k_{\mathrm{d,i}}$ is the protein-specific degradation rate,

- k_{g} is the dilution rate due to cell growth, defined as $k_{\mathrm{g}} = \ln(2)/T_{\mathrm{cc}}$, with T_{cc} being the cell doubling time.

To simulate $k_{\mathrm{s,i}}$ and $k_{\mathrm{d,i}}$, we assume that their logarithms follow a bivariate normal distribution. Let K_s and K_d be the random variable defined as:

$$K_s = \log(k_{s,i}), \quad K_d = \log(k_{d,i}), \quad i \in \{1, 2, \ldots, n\}.$$

where $k_{s,i}$ represents the individual translation rates for each protein i in the population. Values for K_s and K_d were then simulated from a bi-variate normal distribution:

$$\begin{bmatrix} K_s \\ K_r \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{\mathrm{s}} \\ \mu_{\mathrm{d}} \end{bmatrix}, \begin{bmatrix} \sigma_{\mathrm{s}}^2 & \mathrm{cov}(K_s, K_d) \\ \mathrm{cov}(K_s, K_d) & \sigma_{\mathrm{d}}^2 \end{bmatrix} \right),$$

where:

- μ_{s} and μ_{d} are the means of K_s and K_d, respectively,
- σ_s andσ_d are the standard deviations of $\log(k_{\mathrm{s,i}})$ and $\log(k_{\mathrm{d,i}})$,
- ρ is the correlation between K_s and K_d.

Pathway enrichment analysis

Pathway enrichment analysis was performed on both protein fold changes, regulated protein synthesis changes, and correlations between concentrations and clearance rates. In all cases, a t-test was performed between values from a particular KEGG gene ontology pathway and the null distribution of comparisons for all other proteins. Comparisons were only made if 5 or more proteins were present in a pathway. The distribution of p-values generated from the comparisons was then corrected for multiple hypothesis testing via the Benjamini-Hochberg false discovery rate (FDR) correction. Pathways below 1 % FDR were considered significant.

References

1. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–342 (2011).

2. Kristensen, A. R., Gsponer, J. & Foster, L. J. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Molecular systems biology* **9,** 689 (2013).

3. Jovanovic, M. *et al.* Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347,** 1259038 (2015).

4. Baum, K., Schuchhardt, J., Wolf, J. & Busse, D. Of gene expression and cell division time: a mathematical framework for advanced differential gene expression and data analysis. *Cell Systems* **9,** 569–579 (2019).

5. Ross, A. B., Langer, J. D. & Jovanovic, M. Proteome turnover in the spotlight: approaches, applications, and perspectives. *Molecular & Cellular Proteomics* **20** (2021).

6. Derks, J. *et al.* Increasing the throughput of sensitive proteomics by plexDIA. *Nature Biotechnology.* https://doi.org/10.1038/s41587-022-01389-w (2022).

7. Borteçen, T., Müller, T. & Krijgsveld, J. An integrated workflow for quantitative analysis of the newly synthesized proteome. *Nature Communications* **14,** 8237 (2023).

8. Eichelbaum, K. & Krijgsveld, J. Rapid temporal dynamics of transcription, protein synthesis, and secretion during macrophage activation. *Molecular & Cellular Proteomics* **13,** 792–810 (2014).

9. Mathieson, T. *et al.* Systematic analysis of protein turnover in primary cells. *Nature communications* **9,** 689 (2018).

10. Fornasiero, E. F. *et al.* Precisely measured protein lifetimes in the mouse brain reveal differences across tissues and subcellular fractions. *Nature communications* **9,** 4230 (2018).

11. Li, W. *et al.* An Extensive Atlas of Proteome and Phosphoproteome Turnover Across Mouse Tissues and Brain Regions. *bioRxiv,* 2024–10 (2024).

12. Eden, E. *et al.* Proteome half-life dynamics in living human cells. *Science* **331,** 764–768 (2011).

13. Lu, C. & King, R. D. An investigation into the population abundance distribution of mRNAs, proteins, and metabolites in biological systems. *Bioinformatics* **25,** 2020–2027 (2009).

14. Rieckmann, J. C. *et al.* Social network architecture of human immune cells unveiled by quantitative proteomics. *Nature immunology* **18,** 583–593 (2017).

15. Macallan, D. C. *et al.* B-cell kinetics in humans: rapid turnover of peripheral blood memory cells. *Blood* **105,** 3633–3640 (2005).

16. Milo, R. & Phillips, R. *Cell biology by the numbers* (Garland Science, 2015).

17. Michalek, R. D. *et al.* Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. *The Journal of immunology* **186,** 3299–3303 (2011).

18. Banjac, I., Maimets, M. & Jensen, K. B. Maintenance of high-turnover tissues during and beyond homeostasis. *Cell Stem Cell* **30,** 348–361 (2023).

19. Peng, M. *et al.* Protease bias in absolute protein quantitation. *Nature methods* **9,** 524–525 (2012).

20. Toyama, B. H. *et al.* Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* **154,** 971–982 (2013).

21. Dörrbaum, A. R., Kochen, L., Langer, J. D. & Schuman, E. M. Local and global influences on protein turnover in neurons and glia. *Elife* **7,** e34202 (2018).

22. Franks, A., Airoldi, E. & Slavov, N. Post-transcriptional regulation across human tissues. *PLoS computational biology* **13,** e1005535. https://doi.org/10.1371/journal.pcbi.1007082 (2017).

23. Jiang, L. *et al.* A quantitative proteome map of the human body. *Cell* **183,** 269–283 (2020).

24. Van den Berg, P. R., Bérenger-Currias, N. M. L. P., Budnik, B., Slavov, N. & Semrau, S. Integration of a multi-omics stem cell differentiation dataset using a dynamical model. *PLOS Genetics* **19,** 1–23. https://doi.org/10.1371/journal.pgen.1010744 (May 2023).

25. Khan, S. *et al.* Inferring post-transcriptional regulation within and across cell types in human testis. *bioRxiv*. https://doi.org/10.1101/2024.10.08.617313 (2024).

26. Nguyen, A. T. *et al.* UBE2O remodels the proteome during terminal erythroid differentiation. *Science* **357,** eaan0218 (2017).

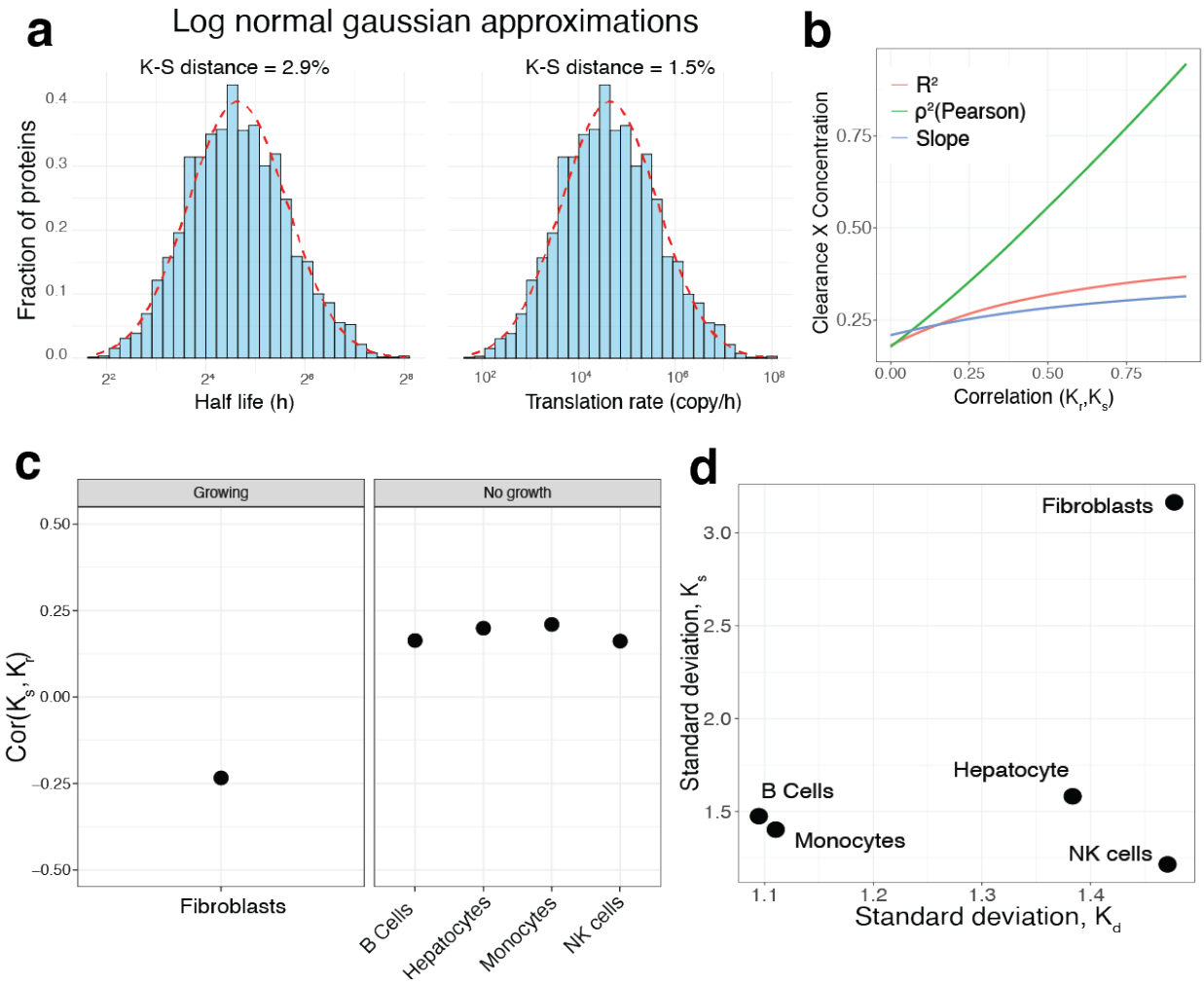
27. Tsour, S., Machne, R., Leduc, A., Widmer, S. & Slavov, N. Alternate RNA decoding results in stable and abundant proteins in mammals. *bioRxiv*. https://doi.org/10.1101/2024.08.26.609665 (2024).

28. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature protocols* **14,** 68–85 (2019).

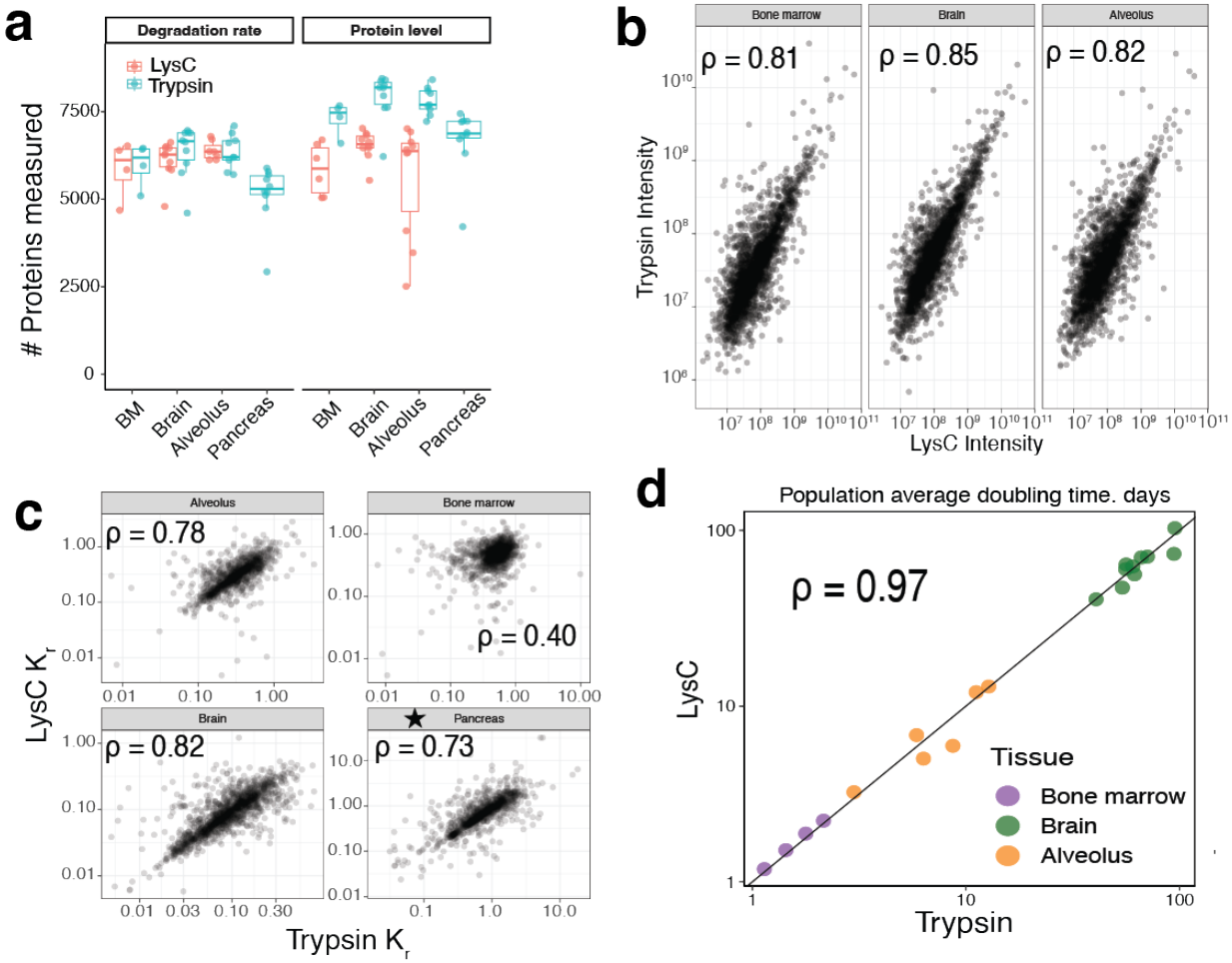
29. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nature methods* **14,** 513–520 (2017).

30. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nature methods* **17,** 41–44 (2020).

Supplemental Figures



Supplemental Fig. S1 | Evaluating characteristics of degradation and synthesis distributions (**a**), The K-S distances of an approximated log normal fit to the half life and synthesis rate distributions show close concordance differing at most by 3%. (**b**) The R^2 computed as the sum of squared residuals from regressing $-log(K_r) = log(P)$ does not overestimate the percent variance with increased correlation between growth and synthesis. This can be observed as the slope computed between $-log(K_r)$ and $log(P)$ remains modest and does not approach 1. (**c**) Correlation between synthesis and degradation rates plotted for each cell type. (**d**) Standard deviations of the $log(K_r)$ and $log(K_s)$ distributions for each cell type.



Supplemental Fig. S2 | Quantitative accuracy and reproducibility of protein degradation and clearence rates for mouse tissues (**a**) Number of proteins with degradation rates and abundances quantified for LysC and Trypsin digests. (**b**) Reproducibility for protein abundance measurements for trypsin and LysC digests. (**c**) Reproducibility for clearance rate measurements for trypsin and LysC digests. Because the pancreas did not have a LysC condition due to the large amount of trypsin expressed in the tissue, technical replicates are plotted. (**d**) Reproducibility for measurements of cell doubling time as defined by average clearance rate of histone H3 and H4.