11-2016

Interspecies Comparison of Peptide Substrate Reporter Metabolism using Compartment-Based Modeling [post-print]

Allison Tierney
Trinity College, Hartford Connecticut, allison.tierney@trincoll.edu

Nhat Pham
Trinity College, Hartford Connecticut, nhat.pham@trincoll.edu

Kunwei Yang
Trinity College, Hartford Connecticut, kunwei.yang@trincoll.edu

Brooks Emerick
Trinity College, Hartford Connecticut, brooks.emerick@trincoll.edu

Michelle Kovarik
Trinity College, Hartford Connecticut, michelle.kovarik@trincoll.edu

Interspecies Comparison of Peptide Substrate Reporter Metabolism using Compartment Based Modeling

Journal:	*Analytical and Bioanalytical Chemistry*
Manuscript ID	Draft
Type of Paper:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Tierney, Allison; Trinity College, Chemistry Pham, Nhat; Trinity College, Mathematics Yang, Kunwei; Trinity College, Chemistry Emerick, Brooks; Trinity College, Mathematics Kovarik, Michelle; Trinity College, Chemistry
Keywords:	Amino acids / Peptides, Capillary electrophoresis / Electrophoresis, Modeling, Bioanalytical methods, Biological samples, Enzymes

Interspecies Comparison of Peptide Substrate Reporter Metabolism using Compartment-Based Modeling

Allison J. Tierney,[a] Nhat Pham,[b] Kunwei Yang,[a] Brooks K. Emerick,[b] and Michelle L. Kovarik[a,*]

[a]Department of Chemistry, Trinity College, 300 Summit Street, Hartford, CT 06106

[b]Department of Mathematics, Trinity College, 300 Summit Street, Hartford, CT 06106

*Corresponding Author. Email: michelle.kovarik@trincoll.edu. Phone: 860-297-5275. Fax: 860-297-5129

Abstract

Peptide substrate reporters are fluorescently labeled peptides that can be acted upon by one or more enzymes of interest. Peptide substrates are readily synthesized and more easily separated than full-length protein substrates; however, they are often more rapidly degraded by peptidases. As a result, peptide reporters must be made resistant to proteolysis in order to study enzymes in intact cells and lysates. This is typically achieved by optimizing the reporter sequence in a single cell type or model organism, but studies of reporter stability in a variety of organisms are needed to establish the robustness and broader utility of these molecular tools. We measured peptidase activity toward a peptide substrate reporter for protein kinase B (Akt) in *E. coli*, *D. discoideum*, and *S. cerevisiae* using capillary electrophoresis with laser-induced fluorescence (CE-LIF). Using compartment-based modeling, we determined individual rate constants for all potential peptidase reactions and explored how these rate constants differed between species. We found the reporter to be stable in *D. discoideum* ($t_{1/2}$ = 82-103 min) and *S. cerevisiae* ($t_{1/2}$ = 279-314 min), but less stable in *E. coli* ($t_{1/2}$ = 21-44 min). These data suggest that the reporter is sufficiently stable to be used for kinase assays in eukaryotic cell types while also demonstrating the potential utility of compartment-based models in peptide substrate reporter design.

Introduction

Peptide substrate reporters are short peptides, typically 3-20 amino acids long, that can be acted on by one or more enzymes of interest [1, 2]. Peptide substrates are commonly used for activity assays in place of full-length endogenous substrates because they are easier to synthesize and separate than full-length endogenous protein substrates. For detection, peptide substrates may be tagged with fluorogenic moieties [3–5], a fluorescent label [6–10], or a FRET pair [11–13]. Reporters have been developed for many enzymes, particularly kinases [2, 4, 6, 8, 10, 12, 14–16], but also proteases [11, 13, 17–19], phosphatases [9], and others [7]. Reporter development typically starts with a peptide library, often based on the consensus sequence for an enzyme's known endogenous substrates, and proceeds through optimization of the amino acid sequence for rapid kinetics, high specificity, and stability in cells and lysates [1, 10, 16].

Stability of exogenous peptides, including substrate reporters, is an active area of research because cells express a number of cytosolic peptidases that degrade peptides into amino acids for recycling into new proteins [20]. Degradation by peptidases is not a problem for assays performed with purified kinase (or other enzyme), but resistance to degradation is a key parameter for substrates for measuring the activities of enzymes when proteases are present, such as in intact cells or cell lysates. In cells or lysates, the kinetics of reporter degradation must be appreciably slower than the kinetics of reporter reaction with the enzyme-of-interest or meaningful measurements cannot be made. Practically speaking, this usually means the half-life of the reporter in the cytoplasm or cell lysate should be at least 30 minutes or longer. In general, peptides are rapidly metabolized in the cytoplasm with half-lives ranging from < 1 min to 20 min [20]. However, degradation rates depend on peptide sequence, and a variety of design principles

have been used to render exogenous peptides resistant to degradation. Strategies include the incorporation of non-native amino acids (e.g., D-amino acids or N-methylated amino acids [16, 21]) and protection of N-termini by bulky modifications [22, 23]. The choice of modifications used to stabilize a given substrate is constrained by the substrate preferences of the enzyme-of-interest but is also largely empirical. As a result, peptide substrate reporters are optimized by a rather time-consuming, iterative design process, in which modifications to improve stability are screened for their effect on kinetics and specificity toward the enzyme-of-interest.

To date, reporters have been developed and tested almost exclusively in mammalian cells, and many reporters have been validated in only one specific human cell line or a few closely-related lines. Only a few reporters are tested in non-mammalian cell types [14], and fewer still in non-vertebrate or single-celled model organisms [5]. General design principles for peptide substrate reporters are lacking, and few reporters have been tested across species. Ideally, a reporter optimized in one organism would be widely applicable to other organisms with minimal re-optimization. This is particularly important as the need for studies in non-traditional model organisms becomes more apparent [24]. However, to date, there has been minimal research on the transferability of peptide substrate reporters between species. Such testing should examine how differences between organisms affect reporter kinetics, specificity, and stability. A comparison of stability across species is prerequisite to a comparison of kinetics and specificity across species because meaningful kinetic data cannot be collected if reporters are not stable in a range of cell types.

Variation in peptidase types and activities between organisms will certainly result in differing

stability of reporters between species. While all species universally express a core set of sixteen peptidase families, other peptidases and peptidase families are expressed exclusively in a specific kingdom. For example, the peptidase family involved in signal peptide processing is a core family found in the genomes of all living organisms checked to date. In contrast, *E. coli* expresses a peptidase family specifically involved in bacterial interactions with surfaces, while animal cells produce peptidases required for remodeling of the extracellular matrices around tissue [25]. Even within a single species, peptidase activity varies by cell type; for example, peptidase activity in cancer cells and cells from rheumatoid arthritis patients differs from that of healthy cells [26, 27]. As a result of this variability in peptidase expression and activity, systematic comparisons of peptide half-lives and a framework for interpreting peptide degradation data are needed to inform future peptide substrate reporter design and applications.

In this work, we compare the degradation of a peptide substrate reporter (VI-B) for protein kinase B across four species from four different kingdoms, *Escherichia coli* (Bacteria), *Dictyostelium discoideum* (Protozoa), *Saccharomyces cerevisiae* (Fungi), and *Homo sapiens* (Animalia). These organisms are evolutionarily divergent and express widely varying peptidases. These differences are reflected in varying peptidase activity toward the reporter and can be quantified using compartment-based models which reveal what steps in the degradation process are most important in each organism. Modeling of breakdown kinetics reveals which amino acid residues and fragments are targeted by peptidases; consequently, modeling results should prove useful in future optimization of peptide substrate reporters. Additionally, a more thorough understanding of interspecies variation in peptidase activity is relevant to future applications of other peptide-based indicators, hormones, and pro-drugs.

Materials and Methods

Cell culture and lysis conditions. Overnight cultures of *Escherichia coli* K12 were grown in LB (10 g/L Bacto-tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.5) at 37 °C with shaking. These cultures were diluted 1:100 and grown to mid-log phase (OD 0.5) before lysis. *Dictyostelium discoideum* was obtained from the Dicty Stock Center [28] and cultured at room temperature in HL-5 pH 6.4-6.7, (14 g/L proteose peptone 3, 7 g/L yeast extract, 3.5 mM dibasic sodium phosphate, and 11 mM monobasic potassium phosphate) with shaking at 180 rpm [29]. Cells were used at a density of 2-4×10^6/mL. For *D. discoideum* social development experiments, cells were washed and resuspended in development buffer (5 mM sodium phosphate dibasic and 5 mM potassium phosphate monobasic, pH 6.5 with calcium chloride and magnesium chloride added to final concentrations of 10 mM and 20 mM respectively before use) at a density of 10^7/mL for 2-6 h prior to lysis. *Saccharomyces cerevisiae* (ATC 4040002) was cultured at 30 °C. Shaken, liquid cultures were started in YPD (20 g/L Bacto-peptone, 10 g/L yeast extract, 20 g/L dextrose) from single colonies grown on YPD plates. These overnight cultures were diluted to OD 0.2 and grown to OD 0.4 to reach mid-log phase before lysis.

Prior to lysis, all cell types were pelleted, washed, and resuspended in phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). For *E. coli*, 200 mL of OD 0.5 culture was washed and resuspended in 1 mL of ice-cold PBS and lysed by sonicating 3× for 10 s at power 3-4 with 1.5-2 min on ice between cycles. For *D. discoideum* cultures, 1.5 × 10^7 cells were lysed by three freeze-thaw cycles using liquid nitrogen. For *S. cerevisiae*, 50 mL of OD 0.4 culture was washed and resuspended in 1.0 mL of ice-cold PBS and mixed with 1 mL

of 0.5 mm glass beads. Cells were lysed by bead beating 5× for 10 s/cycle with 1 min on ice between cycles. All of the resulting lysates were centrifuged for 5-15 min at 15,000 ×g, and the supernatant was removed and stored at -80 °C for up to 10 days before use. Biological replicates were prepared as described above from three separate overnight cultures of each cell type.

Degradation assays. The total protein concentration of each lysate was determined using the fluorescamine assay with a bovine serum albumin calibration curve [30]. Lysates were diluted 1:1000 in 30 mM borate buffer (pH 9), mixed in a 3:1 ratio with 3 mg/mL fluorescamine in acetone, and incubated in the dark for 2 min at room temperature. Fluorescence was excited at 390 nm and measured at 475 nm. Based on this assay, the lysate was diluted to a final concentration of 3 mg/mL in phosphate buffered saline. The reaction was started by addition of the peptide substrate reporter, 6FAM-GRP-NMeArg-AFTF-NMeAla-NH_2,[16] to a final concentration of 1 µM. Reactions were run at the normal growth temperature of each cell type: 37 °C for *E. coli*, room temperature (25 °C) for *D. discoideum*, and 30 °C for *S. cerevisiae.* Aliquots were removed at 15, 30, 60, 90, 120, 180, 240, 300, and 360 min after the start of the reaction and heated to 95 °C for 5 min to stop the reaction. Data for HeLa and LNCap cells were published previously and graciously provided by the authors [16].

Capillary electrophoresis. Samples from the lysate assays were analyzed using a PA-800 Plus capillary electrophoresis instrument (Beckman Coulter). The run buffer was 100 mM borate, 15 mM SDS, pH 11.4 [27]. The capillary was 50 µm diameter bare silica, 21 cm effective length with an applied potential of 393 V/cm. Using purified standards, we confirmed that the parent peptide substrate reporter and all fluorescent N-terminal fragments were separated under these

conditions. Peaks were identified by their migration times and confirmed by addition of purified standards. Peak integration was performed using the 32 Karat Software (Beckman Coulter).

Data analysis and compartment-based modeling. Peak areas were converted to percent of total peak area and then to concentration based on an initial peptide concentration of 1 μM. To fit our compartment-based model to the data, we first solved Equations (1)-(5) (see below and the Electronic Supplemental Material). We then used least-squares regression to solve for the rate constants by fitting the explicit solutions to Equations (1)-(5) to the kinetic data. We fit the solutions to Equations (1)-(5) sequentially, starting with the solution to Equation (1), which could be linearized. The solutions to Equations (2)-(5) cannot be linearized, so we took a nonlinear least-squares approach for these fits using the built-in Matlab function called lsqnonlin. This built-in function uses an iterative approach that searches for a minimum in the relevant parameter space using the method of steepest descent [31]. Since this is an iterative method, an initial guess must be provided with an upper and lower bound for each rate constant. We assumed that rate constants were non-negative with an appropriate upper bound (as discussed in the Supplemental Material); however, we had no initial insight into the actual rate constant values, so we attempted fitting with many different initial values and chose the parameter set that yielded the smallest residual error. To determine a 90% confidence interval for the half-lives and initial rates, we performed bootstrapping using a Monte Carlo simulation to generate 1000 synthetic data sets based on the mean and standard deviation of triplicate measurements. The method of fitting the data was selected after extensive optimization. Further details of all mathematical analyses are in the Electronic Supplemental Material.

Results and Discussion

Identification of the products of reporter degradation. In this work, we studied the degradation of a peptide substrate reporter for protein kinase B (PKB or Akt) in cell lysates from *E. coli*, *D. discoideum*, and *S. cerevisiae*. This reporter, called VI-B, was previously optimized using HeLa and LNCaP cell lysates [16]. PKB is a serine/threonine kinase that plays a key role in cell proliferation, stress, response, and apoptosis [32]. Previous studies have used the reporter to measure PKB activity in individual human cells, including primary tissue samples from patients diagnosed with pancreatic cancer and rheumatoid arthritis [16, 27, 33]. Both *S. cerevisiae* and *D. discoideum* express homologs of human PKB that are important in stress response [34, 35]. This reporter could be useful in probing PKB activity in these organisms, but first it is necessary to determine whether the peptide is phosphorylated by the homologous enzymes and sufficiently resistant to degradation in these systems. In this work, we assessed the degradation resistance of the peptide in lysates from *S. cerevisiae* and *D. discoideum.* Although *E. coli* does not express a PKB/Akt homolog, we also tested the stability of the reporter in *E. coli* lysates as a more general exploration of peptidase activity toward reporter molecules in range of evolutionarily-divergent organisms.

Capillary electrophoresis with laser-induced fluorescence detection was used to separate and detect all fluorescent products of the reporter degradation reactions in cell lysates (Fig. 1 a-c). In our discussion, we call the full-length peptide reporter *R* and refer to the detectable fragments by the number of amino acid residues still attached to the N-terminal fluorescent tag; for example, F_1 refers to the one amino acid fragment 6FAM-G. The optimized run buffer was able to separate all ten peaks in under twenty minutes (Fig. 1d), although peaks for F_3, F_4, and F_5 were

not baseline resolved from one another. Cell lysate samples for *E. coli* included F_4, F_5, F_6, and F_7 (Fig. 1a) while lysates from *D. discoideum* and *S. cerevisiae* contained F_5, F_6, F_7, and F_8. Lysates from all three model organisms also included the full-length, unmodified reporter, which degraded over time due to peptidase activity (Fig. 1e). Phosphorylation of the reporter was not expected to occur under the experimental conditions and was not observed. For quantitation, we determined the percent peak area for each peptide species, which was presumed to be proportionate to the relative concentration. We assessed degradation kinetics by tracking these relative concentrations across ten time points from 0 min to 360 min.

The relative concentrations of reporter and fragments, measured as fraction of total peak area, were tracked as a function of reaction time. As expected, the concentration of unmodified reporter decreases over time as it is cleaved by peptidases in the lysates to form shorter fragments (Fig. 2 a-c), but the rates of reporter degradation differed between the three organisms. Degradation was most rapid in *E. coli* (Fig. 2a), while *S. cerevisiae* lysates showed the slowest degradation (Fig. 2c). As expected for first-order kinetics, the semi-log plot of reporter concentration versus time is linear with more rapid degradation corresponding to a steeper slope (Fig. 2d). As the parent was degraded, the relative concentrations of the fluorescent fragments increased. In both *E. coli* and *D. discoideum*, F_7 is the major fragment that forms; in *S. cerevisiae*, F_5 is the main fragment. At longer time points, the relative concentrations of some fragments also begin to decrease, suggesting that these fragments were further degraded by peptidases. To elucidate the relative kinetics of these degradation reactions, a compartment-based model was employed.

Compartment-based modeling. A compartment-based model or multi-compartment model is a mathematical model used to describe the transmission of materials or concentrations among different components of a physical system [36]. This method is commonly used in pharmacokinetic studies and was recently applied to studies of protein turnover [37]. The value of a compartment-based model lies in its simplicity as well as the fact that it represents an underlying physical process (in this case the chemistry of peptide metabolism). Essentially, a function of interest is generated from the physical process itself rather than general trends in the data. Compartment-based modeling works well with peptide degradation since the peptide reporter breaks down into smaller fragments whose concentrations are measured. In this work, we use the model to describe the concentration of each peptide over time (Fig. 3). The model includes several assumptions: (1) that there is no source of peptide fragments except for the initial full-length reporter; (2) that there is no sink, i.e., the parent peptide and fragments do not leave the system, but are simply converted to smaller fragments; (3) that any larger peptide could be cleaved to form any smaller peptide, but that only the N-terminal fragments, which retain the 6-FAM label, will be detected and (4) that the system follows first-order kinetics and all rate constants are non-negative. (A ten-fold increase in substrate concentration, tested in *Dictyostelium* lysates, resulted in a ten-fold increase in rate and comparable half-life, validating the choice of a first-order kinetic model.) These assumptions describe a closed linear system whose solutions are a linear combination of exponential functions. Similar methods have been used previously to model enzymatic reactions [38].

The compartment-based model can be translated into a series of differential equations (Equations 1-5). For example, consider Equation (3), which describes the concentration of the seven amino

acid long fragment (F_7) as a function of time (t). F_7 can be formed either from the parent reporter (R) or from the 8 amino acid long fragment (F_8), resulting in the terms $+k_7R$ and $+k_{87}F_8$, respectively, where k_7 is the rate constant for formation of F_7 from the full-length reporter, R is the concentration of the reporter, k_{87} is the rate constant for the formation of F_7 from F_8, and F_8 is the concentration of the eight amino acid fragment. These two positive terms account for the formation of F_7. The negative term, $-(k_{76} + k_{75})F_7$, accounts for degradation of the seven amino acid fragment to form the six amino acid fragment (F_6) and the five amino acid fragment (F_5). The rate constants for these two degradation reactions can be combined into a rate constant reflecting the overall disappearance of F_7, called $k_{\text{net, 7}}$. Thus Equation (3) describe the entirety of possible reactions involving F_7 that are compatible with the data (in which F_5 was the smallest fragment detected).

$$\frac{dR}{dt} = -(k_8 + k_7 + k_6 + k_5)R = -k_{net,R}R \quad (1)$$

$$\frac{dF_8}{dt} = -(k_{87} + k_{86} + k_{85})F_8 + k_8R = -k_{net,8}F_8 + k_8R \quad (2)$$

$$\frac{dF_7}{dt} = -(k_{76} + k_{75})F_7 + k_{87}F_8 + k_7R = -k_{net,7}F_7 + k_{87}F_8 + k_8R \quad (3)$$

$$\frac{dF_6}{dt} = -k_{65}F_6 + k_{76}F_7 + k_{86} + k_6R \quad (4)$$

$$\frac{dF_5}{dt} = k_{65}F_6 + k_{75}F_7 + k_{85}F_8 + k_5R \quad (5)$$

Fig. 3 and Equations (1)-(5) describe the fragments and corresponding reactions observed for *D. discoideum*, *S. cerevisiae*, and the human cell lines. Qualitatively the fragmentation patterns for these cells types, which are all eukaryotic, were similar; the same fragments were observed,

albeit in different quantities, in each cell type. Many peptidases unique to multicellular organisms, including humans, are extracellular peptidases that may not be retained in cytoplasmic lysates. This may explain the qualitative similarity of the fragmentation patterns observed in the human cell lines and other eukaryotic cell types [25]. The *E. coli* lysates produced slightly different fragments: F_8 was not observed, but F_4 was. However, the corresponding compartment-based model (Fig. S1) has the same mathematical form as that shown in Fig. 3. This type of model is deterministic, meaning the output of the model is completely determined by the model parameters. Rate constants for each reaction were obtained by solving the differential equations and then sequentially fitting the solutions to the data using least-squares regression (Table 1). For the eukaryotic cell lysate data, minimum residuals for the fitted rate constants ranged from 10^{-6}-10^{-1} with typical values around 10^{-4}; minimum residuals were somewhat higher for the *E. coli* lysate data, ranging from 10^{-3} to 10^{-1} (Table S1).

Rate constants for the degradation reactions varied widely between reactions and organisms. At one end, some very low k values (10^{-13}-10^{-14}) suggested that certain degradation steps occur so slowly as to be negligible contributors to metabolism of the reporter. For example, formation of F_6 from F_8 in the human cells lines would be kinetically unfavorable. At the other extreme, the maximum reported rate constant of 1 was constrained by the parameters used to fit the data in MatLab. This value was only reached for one reaction, the formation of F_5 from F_6 in LNCaP cells. This occurred because F_6 (the six amino acid fragment) was not observed in LNCaP lysates. This observation could be explained either of two ways: F_6 formed but was rapidly degraded to F_5 or F_6 was never formed. Because $k_6 > 0$ for the LNCaP data, the model suggests that F_6 did form from the full-length reporter but was rapidly degraded to F_5. To further test this

hypothesis, we fit the LNCaP data to a simplified model that assumed F_6 never formed ($k_6 = k_{86} = k_{76} = 0$; Fig. S2). Based on qualitative assessment and minimum residuals, the more complex model (that included formation of F_6) was a better fit to the data for F_7 and a comparable fit for R, F_8, and F_5 when compared to the simplified model. The model results also suggest a further test to confirm that F_6 was formed and rapidly degraded. The value for k_{85} is eight orders of magnitude higher in the complex model than in the simplified model, so incubation of F_8 in LNCaP lysates could provide further support for one model over the other. Excluding the proposed rapid destruction of F_6 to form F_5, the highest rate constant observed was 10^{-1} for the formation of F_5 from F_7 in LNCaP lysates. In LNCaP lysates, F_5 was the dominant fragment observed, and rate constants for its formation were $\geq 10^{-2}$ for all possible starting peptides.

Interestingly, some reactions indicative of carboxypeptidase activity had non-negligible rate constants. Carboxypeptidases are peptidases that cleave the final or penultimate C-terminal amino acid residues. Many carboxypeptidases function in the extracellular space, and cytosolic carboxypeptidases (CCPs) are commonly involved in removal of glutamate residues [39]. Perhaps for this reason, previous research reported negligible carboxypeptidase activity toward substrates in the cytosol [22]; however, recent studies have confirmed carboxypeptidase processing of peptides in the cytosol and endoplasmic reticulum [40]. In agreement with these recent findings, all eukaryotic cell lysates generated F_8 by removal of the C-terminal N-methylalanine residue. Additionally, rate constants obtained by compartment-based modeling also suggested carboxypeptidase activity in *S. cerevisiae* and human cells toward shorter fragments (Table 1). In contrast, except for the formation of F_8 from the full-length reporter (R), the rate constants for removal of the C-terminal residue in *D. discoideum* lysates are all

exceedingly low. This may reflect the fact that the *D. discoideum* genome contains only four genes expected to code for carboxypeptidases, and all are expected to be membrane-bound or extracellular [28]. Like the *D. discoideum* data, data from *E. coli* lysates showed little evidence of removal of the C-terminal amino acid; instead, rate constants for *E. coli* reflected a preference for removal of two C-terminal amino acids in a single step. For example, F_8 was not observed in *E. coli*, but formation of F_7 from the full-length (9 amino acid) reporter had a relatively high rate constant (k_7). Similarly, k_{75} and k_{64} were large while single amino acid removal steps (represented by k_{76}, k_{65}, and k_{54}) made negligible contributions to degradation. The result is that production of F_4 leveled off after ~180 min when F_6 was depleted even though larger fragments were still present (Fig. 2A). The mathematical modeling agreed with and explained this particular qualitative observation and clarified in general which reactions contributed to peptide degradation.

In elucidating the kinetics of individual degradation reactions, the results of the compartment-based modeling are useful in optimizing the reporter for application to specific organisms. For example, the major fragment formed during degradation of the reporter in *D. discoideum* lysates was the seven-amino acid fragment, F_7. One strategy for further stabilizing the reporter would be to replace either residue 7 or residue 8 with a non-native amino acid to reduce peptidase activity at this bond; however, residue 7 is the threonine residue that is phosphorylated by protein kinase B, and modifications at this site are likely to affect phosphorylation rates adversely. Compartment-based modeling revealed that the formation of this fragment from the full length peptide (R) was kinetically more favorable than formation of F_7 from F_8 (i.e., $k_7 >> k_{87}$). This suggests that the eight-amino acid fragment, F_8, would be more resistant to this degradation step

than the full-length reporter *R*. In this case, it would be useful to evaluate the activity of the *D. discoideum* PKB homolog toward F_8 to determine whether the fragment could be more stable substitute for the full-length peptide. In yeast and human cells, the magnitudes of the rate constants are reversed (i.e., $k_7 < k_{87}$), so the same strategy would be ineffective.

Interspecies variation in reporter metabolism. To assess the overall degradation resistance of the reporter, the value of $k_{net,R}$ was used to calculate half-life of the reporter and the initial rate of its destruction in each cell type. These values varied widely between cell types, as did more qualitative measures of peptidase activity, such as the major fragment observed during the first hour of degradation (Table 2). The range of initial degradation rates reported here spans two orders of magnitude and is generally comparable to previous reports, which have identified degradation rates of 0.02-3 pmol mg^{-1} s^{-1} for this peptide reporter and others in lysates and intact cells [23, 27, 33].

We considered whether the differences in rates corresponded to each cell type's demand for amino acids to generate new proteins. Cytosolic peptidase activity is required for recycling of amino acids into new proteins; however, differences in degradation rate between cell types were not correlated to cell proliferation rates or proliferation rates normalized to cell volume. For example, *D. discoideum* and HeLa cells are similarly sized, but *D. discoideum* cells double in cell density every 8-12 h [29], while HeLa cells double every 15-30 h [41]. Despite the more rapid proliferation of *D. discoideum* compared to HeLa cells, the two cell types showed remarkably similar overall half-lives for the reporter (Table 2). *E. coli* and *S. cerevisiae* are smaller than these cells, but also proliferate more rapidly. Based on doubling times and cell

volumes, we estimate that *E. coli* cells generate 0.01-0.25 μm^3 of cytoplasm per minute, compared to 0.1-2.3 μm^3/min for *S. cerevisiae*, and 0.3-5.6 μm^3/min for HeLa cells. These rates are wide, overlapping, and uncorrelated with the half-lives observed for the reporter, suggesting that for this particular peptide sequence degradation rates do not reflect general rates of amino acid recycling. This possibility is further supported by our observation that resuspending *D. discoideum* cultures in nutrient-free phosphate buffer (to initiate the organism's social life cycle) did not enhance reporter degradation in these lysates (Table S3). Instead, the degradation rate is likely constrained by the activity of specific peptidases in each organism that find the amino acid sequence of the reporter or fragment to be a suitable substrate. This suggests that certain peptide stabilization strategies may be more effective in certain cell types, depending on the specific peptidases responsible for processing the reporter.

Qualitative differences in degradation showed similar interspecies variation. For the eukaryotic cells, the terminal fragment (i.e., the smallest one observed) was the five amino acid fragment, F_5. Formation of the smaller F_4 fragment would have required cleavage of the peptide bond at the non-native N-methylarginine. In general, peptidase preferences skew toward small, aliphatic residues at the scissile bond with particularly strong preference for the amino acid at the N-terminal side of the scissile bond [42]. As a result, the basic, non-native N-methylarginine residue is likely to stabilize the F_5 fragment. Only *E. coli* lysates showed activity toward the N-methylarginine-alanine bond and formation of F_4. *E. coli* cells express peptidases from 19 families that are not found in the other three cell types tested [25]. Of these 19 families, most are characterized by substrate preferences that do not match that of the bond cleaved to form F_4. However, the endopeptidase omptin prefers a basic residue on the N-terminal side of the scissile

bond and a non-basic residue on the C-terminal side [43]. This peptidase is expressed by *E. coli* but not *D. discoideum, S. cerevisiae*, or human cells. These observations suggest that N-methylarginine substitution may be broadly protective of the peptide bonds in eukaryotic organisms, but further studies with a variety of peptide sequences are needed to confirm this.

Conclusions

The genomes of these organisms vary widely in their capacity to express peptidases. Human cells encode for a total of 745 peptidases from 77 peptidase families, compared to 92 peptidases and 44 families in *S. cerevisiae,* 166 peptidases and 59 families in *D. discoideum*, and 89 peptidases and 48 families in *E. coli*. Hierarchical cluster analysis of the peptidase genes from these species and many other organisms demonstrates that there exists a core group of peptidases expressed by all cells, as well as clusters of peptidases specific to prokaryotes, eukaryotes, fungi, and metazoans [25]. The cell types investigated here correspond to one organism from each of these groups. As a result, we expect that these data should be representative of how the reporter would be degraded in a wide range of organisms. Although degradation in LNCaP lysates is very rapid for this reporter, in general, the degradation resistance of VI-B suggests that it is sufficiently stable for application across the eukaryotic kingdom. This is important since the PI3K-PKB pathway is well-conserved throughout eukaryotic cell types; however, further research is needed to determine whether the reporter is a suitable substrate for phosphorylation by PKB homologs in other species. While *E. coli* cells do not express a PKB homolog, prokaryotic organisms do use their own serine/threonine kinase signaling pathways [44], and these results could inform design of reporters for bacterial kinases. Additionally, bacterial pathogens often disrupt eukaryotic signaling, including the PI3K-PKB pathway, during infection

[45]. Consequently, this reporter could be applied to samples that exhibit some prokaryotic peptidase activity. More generally, these results suggest the utility of compartment-based models for optimizing substrate reporters for degradation resistance. Compartment-based modeling of degradation of several substrates with disparate amino acid sequences in many cell types will yield generalizable design principles that should speed future reporter design. Additionally, a more complete understanding of peptide metabolism in cells will be useful in the design of other peptide-based tools, including hormones and pro-drugs.

Electronic Supplemental Material. Additional details about the compartment-based modeling, including the model used for the *E. coli* data and minimum residuals for all rate constants, may be found in the electronic supplemental material.

Acknowledgments. The authors thank the Allbritton Laboratory at the University of North Carolina for generously providing advice and peptide standards, particularly Angela Proctor and Emilie Mainz for helpful discussions and their collaborator Qunzhao Wang for synthesis of the peptides. We also thank Jeremiah Marden and the Graf Laboratory at the University of Connecticut for assistance with the *S. cerevisiae* and *E. coli* cultures and lysis. This work was supported by Trinity College.

Conflict of Interest. The authors declare that they have no conflict of interest.

References

1. Hardie DG (2000) Peptide Assay of Protein Kinases and Use of Variant Peptides to Determine Recognition Motifs. In: Walker J, Keyse S (eds) Stress Response. Humana Press, pp 191–201
2. Wu D, Sylvester JE, Parker LL, Zhou G, Kron SJ (2010) Peptide reporters of kinase activity in whole cell lysates. Biopolymers 94:475–486. doi: 10.1002/bip.21401
3. Yaron A, Carmel A, Katchalski-Katzir E (1979) Intramolecularly quenched fluorogenic substrates for hydrolytic enzymes. Analytical Biochemistry 95:228–235. doi: 10.1016/0003-2697(79)90210-0
4. Shults MD, Imperiali B (2003) Versatile Fluorescence Probes of Protein Kinase Activity. J Am Chem Soc 125:14248–14249. doi: 10.1021/ja0380502
5. Kraft M, Radke D, Wieland GD, Zipfel PF, Horn U (2007) A fluorogenic substrate as quantitative in vivo reporter to determine protein expression and folding of tobacco etch virus protease in Escherichia coli. Protein Expr Purif 52:478–484. doi: 10.1016/j.pep.2006.10.019
6. Chen C-A, Yeh R-H, Lawrence DS (2002) Design and Synthesis of a Fluorescent Reporter of Protein Kinase Activity. J Am Chem Soc 124:3840–3841. doi: 10.1021/ja017530v
7. Arkhipov SN, Berezovski M, Jitkova J, Krylov SN (2005) Chemical cytometry for monitoring metabolism of a Ras-mimicking substrate in single cells. Cytometry 63A:41–47. doi: 10.1002/cyto.a.20100
8. Wang Q, Cahill SM, Blumenstein M, Lawrence DS (2006) Self-Reporting Fluorescent Substrates of Protein Tyrosine Kinases. J Am Chem Soc 128:1808–1809. doi: 10.1021/ja0577692

9. Phillips RM, Bair E, Lawrence DS, Sims CE, Allbritton NL (2013) Measurement of protein tyrosine phosphatase activity in single cells by capillary electrophoresis. Anal Chem 85:6136–6142. doi: 10.1021/ac401106e

10. Turner AH, Lebhar MS, Proctor A, Wang Q, Lawrence DS, Allbritton NL (2016) Rational Design of a Dephosphorylation-Resistant Reporter Enables Single-Cell Measurement of Tyrosine Kinase Activity. ACS Chem Biol 11:355–362. doi: 10.1021/acschembio.5b00667

11. Rehm M, Dussmann H, Janicke RU, Tavare JM, Kogel D, Prehn JHM (2002) Single-cell fluorescence resonance energy transfer analysis demonstrates that caspase activation during apoptosis is a rapid process. Role of caspase-3. J Biol Chem 277:24506–24514. doi: 10.1074/jbc.M110789200

12. Ni Q, Titov DV, Zhang J (2006) Analyzing protein kinase dynamics in living cells with FRET reporters. Methods 40:279–286. doi: 10.1016/j.ymeth.2006.06.013

13. Ng EX, Miller MA, Jing T, Chen C-H (2016) Single cell multiplexed assay for proteolytic activity using droplet microfluidics. Biosensors and Bioelectronics 81:408–414. doi: 10.1016/j.bios.2016.03.002

14. Lee CL, Linton J, Soughayer JS, Sims CE, Allbritton NL (1999) Localized measurement of kinase activation in oocytes of Xenopus laevis. Nat Biotechnol 17:759–762. doi: 10.1038/11691

15. Bozinovski S, Cristiano BE, Marmy-Conus N, Pearson RB (2002) The Synthetic Peptide RPRAATF Allows Specific Assay of Akt Activity in Cell Lysates. Analytical Biochemistry 305:32–39. doi: 10.1006/abio.2002.5659

16. Proctor A, Wang Q, Lawrence DS, Allbritton NL (2012) Development of a peptidase-resistant substrate for single-cell measurement of protein kinase B activation. Anal Chem 84:7195–7202. doi: 10.1021/ac301489d

17. Tung CH, Mahmood U, Bredow S, Weissleder R (2000) In vivo imaging of proteolytic enzyme activity using a novel molecular reporter. Cancer Res 60:4953–4958.

18. Dragulescu-Andrasi A, Liang G, Rao J (2009) In vivo bioluminescence imaging of furin activity in breast cancer cells using bioluminogenic substrates. Bioconjug Chem 20:1660–1666. doi: 10.1021/bc9002508

19. Brown RB, Hewel JA, Emili A, Audet J (2010) Single amino acid resolution of proteolytic fragments generated in individual cells. Cytometry A 77:347–355. doi: 10.1002/cyto.a.20880

20. Yewdell JW, Reits E, Neefjes J (2003) Making sense of mass destruction: quantitating MHC class I antigen presentation. Nat Rev Immunol 3:952–961. doi: 10.1038/nri1250

21. Proctor A, Wang Q, Lawrence DS, Allbritton NL (2012) Metabolism of Peptide Reporters in Cell Lysates and Single Cells. Analyst 137:3028–3038. doi: 10.1039/c2an16162a

22. Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout JW, Neefjes J (2003) Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. Immunity 18:97–108.

23. Yang S, Proctor A, Cline LL, Houston KM, Waters ML, Allbritton NL (2013) β-Turn sequences promote stability of peptide substrates for kinases within the cytosolic environment. The Analyst 138:4305. doi: 10.1039/c3an00874f

24. Gladfelter AS (2015) How nontraditional model systems can save us. Mol Biol Cell 26:3687–3689. doi: 10.1091/mbc.E15-06-0429

25. Page MJ, Di Cera E (2008) Evolution of Peptidase Diversity. Journal of Biological Chemistry 283:30010–30014. doi: 10.1074/jbc.M804650200

26. Tamura Y, Niinobe M, Arima T, Okuda H, Fujii S (1975) Aminopeptidases and arylamidases in normal and cancer tissues in humans. Cancer Res 35:1030–1034.

27. Mainz ER, Serafin DS, Nguyen TT, Tarrant TK, Sims CE, Allbritton NL (2016) Single Cell Chemical Cytometry of Akt Activity in Rheumatoid Arthritis and Normal Fibroblast-like Synoviocytes in Response to Tumor Necrosis Factor α. Analytical Chemistry. doi: 10.1021/acs.analchem.6b01801

28. Fey P, Dodson RJ, Basu S, Chisholm RL (2013) One stop shop for everything Dictyostelium: dictyBase and the Dicty Stock Center in 2012. In: Eichinger L, Rivero F (eds) Dictyostelium discoideum Protocols. Humana Press, pp 59–92

29. Fey P, Kowal AS, Gaudet P, Pilcher KE, Chisholm RL (2007) Protocols for growth and development of Dictyostelium discoideum. Nat Protoc 2:1307–1316. doi: 10.1038/nprot.2007.178

30. De Bernardo S, Weigele M, Toome V, Manhart K, Leimgruber W, Böhlen P, Stein S, Udenfriend S (1974) Studies on the reaction of fluorescamine with primary amines. Arch Biochem Biophys 163:390–399.

31. Kelley, CT (1999) Iterative Methods for Optimization. SIAM, Philadelphia

32. Alessi DR, Cohen P (1998) Mechanism of activation and function of protein kinase B. Current Opinion in Genetics & Development 8:55–62. doi: 10.1016/S0959-437X(98)80062-2

33. Proctor A, Herrera-Loeza SG, Wang Q, Lawrence DS, Yeh JJ, Allbritton NL (2014) Measurement of protein kinase B activity in single primary human pancreatic cancer cells. Anal Chem 86:4573–4580. doi: 10.1021/ac500616q

34. Fabrizio P, Pozza F, Pletcher SD, Gendron CM, Longo VD (2001) Regulation of longevity and stress resistance by Sch9 in yeast. Science 292:288–290. doi: 10.1126/science.1059497

35. Meili R, Ellsworth C, Lee S, Reddy TB, Ma H, Firtel RA (1999) Chemoattractant-mediated transient activation and membrane localization of Akt/PKB is required for efficient chemotaxis to cAMP in Dictyostelium. EMBO J 18:2092–2105. doi: 10.1093/emboj/18.8.2092

36. Anderson DH (2013) Compartmental Modeling and Tracer Kinetics. Springer Science & Business Media

37. Guan S, Price JC, Ghaemmaghami S, Prusiner SB, Burlingame AL (2012) Compartment Modeling for Mammalian Protein Turnover Studies by Stable Isotope Metabolic Labeling. Anal Chem 84:4014–4021. doi: 10.1021/ac203330z

38. Kuzmic P (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal Biochem 237:260–273. doi: 10.1006/abio.1996.0238

39. Petrera A, Lai ZW, Schilling O (2014) Carboxyterminal Protein Processing in Health and Disease: Key Actors and Emerging Technologies. Journal of Proteome Research 13:4497–4504. doi: 10.1021/pr5005746

40. Weimershaus M, Evnouchidou I, Saveanu L, van Endert P (2013) Peptidases trimming MHC class I ligands. Curr Opin Immunol 25:90–96. doi: 10.1016/j.coi.2012.10.001

41. Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Res 38:D750–D753. doi: 10.1093/nar/gkp889

42. Rawlings ND (2009) A large and accurate collection of peptidase cleavages in the MEROPS database. Database (Oxford) 2009:bap015. doi: 10.1093/database/bap015

43. Rawlings ND, Barrett AJ, Finn R (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 44:D343-350. doi: 10.1093/nar/gkv1118

44. Cousin C, Derouiche A, Shi L, Pagot Y, Poncet S, Mijakovic I (2013) Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. FEMS Microbiology Letters 346:11–19. doi: 10.1111/1574-6968.12189

45. Krachler AM, Woolery AR, Orth K (2011) Manipulation of kinase signaling by bacterial pathogens. J Cell Biol 195:1083–1092. doi: 10.1083/jcb.201107132

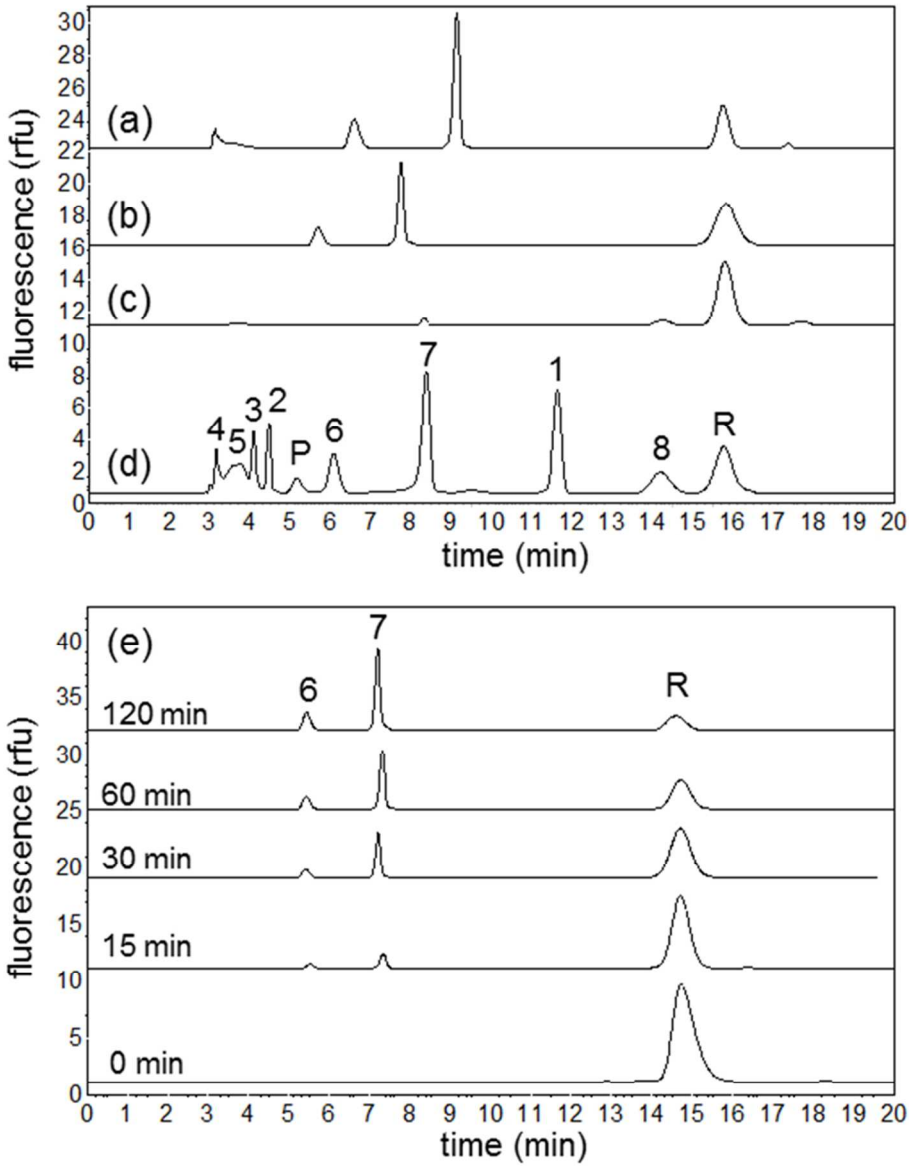


Fig. 1 Capillary electropherograms of peptide samples incubated for 60 min in cell lysates prepared from (a) *E. coli*, (b) *D. discoideum*, and (c) *S. cerevisiae*, and (d) of peptide standards. *R* represents the unmodified, full-length reporter, *P* is the phosphorylated reporter, and 1-8 represent N-terminal fluorescent fragments of the reporter (referred to as F_1-F_8 in the text). For display only, the time axes of all electropherograms were normalized to the migration times for the parent reporter (R) and the five amino acid fragment to facilitate comparisons. (e) Capillary electropherograms for reporter incubated in *D. discoideum* cell lysate for 0 min to 120 min showing degradation of the reporter and formation of fragments over time.

82x106mm (300 x 300 DPI)

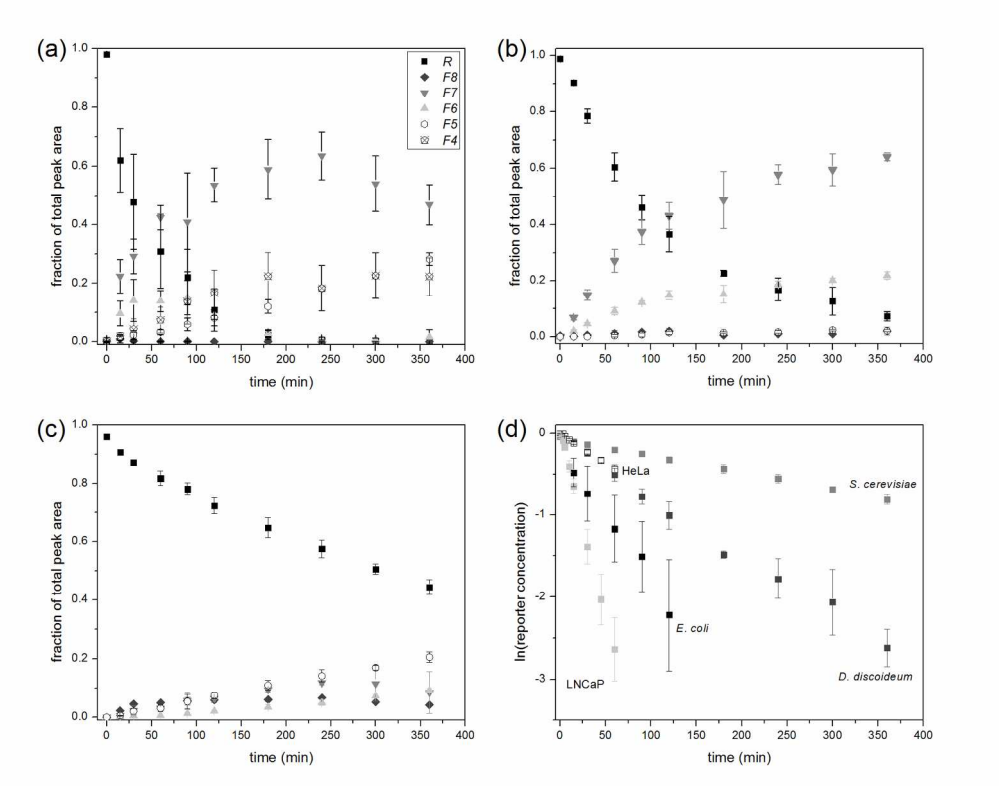


Fig. 2 Peptide abundances as a function of time for metabolism of the reporter in lysates from (a) *E. coli*, (b) *D. discoideum*, and (c) *S. cerevisiae*. Data points are average values for n = 3 biological replicates; error bars show the standard deviation. (d) Linearized semi-log plots of the abundance of full-length reporter (R) as a function of time. The slope of each line gives the first-order rate constant for overall degradation of the reporter ($k_{net,\,R}$). Data for HeLa and LNCaP cells are from ref. [16].

173x137mm (300 x 300 DPI)

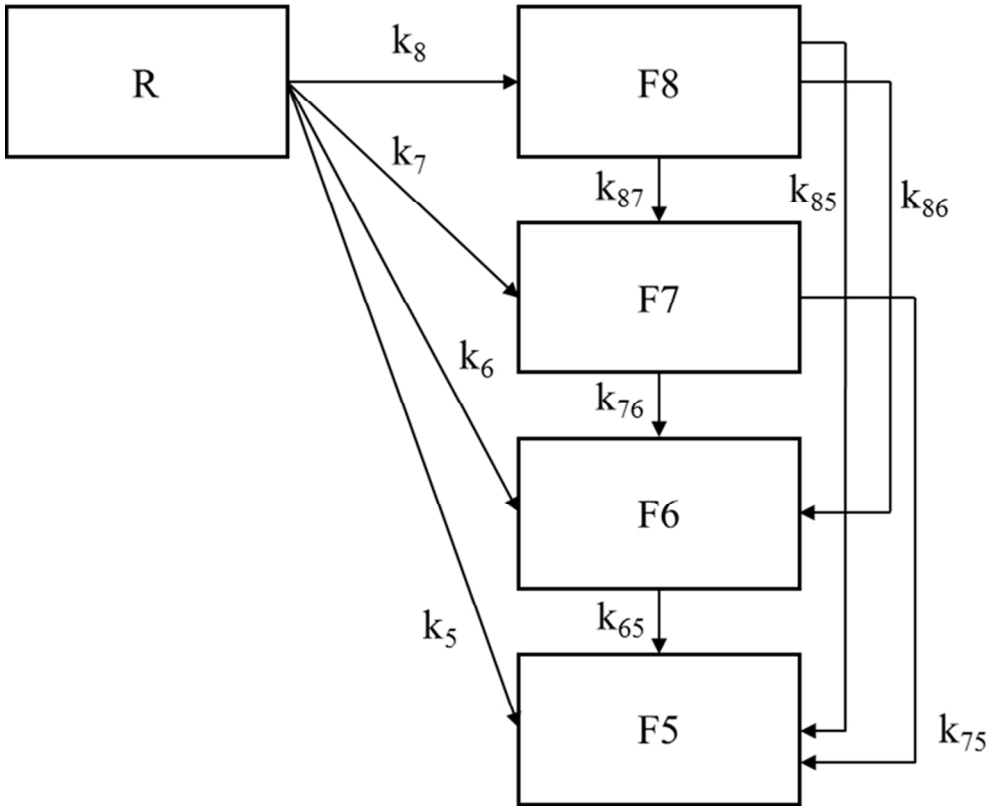


Fig. 3 Compartment-based model of reporter (*R*) metabolism into shorter N-terminal fragments (F_X) with *X* remaining amino acid residues after the N-terminal fluorescent label. The kinetics of each reaction are described as a rate constant, *k*.

82x66mm (300 x 300 DPI)

Table 1 Rate constants for peptidase reactions of the reporter in each species, as determined by compartment-based models. Unreported values (--) indicate that one of the fragments involved in the reaction was not observed experimentally and therefore not included in the model. For example, R_4 was not observed in any experiments except those conducted with *E. coli*.

Rate Constant (min^{-1})	*E. coli*	*D. discoideum*	*S. cerevisiae*	HeLa	LNCap
k_8	--	2×10^{-4}	2×10^{-3}	3×10^{-3}	5×10^{-3}
k_7	1×10^{-2}	5×10^{-3}	2×10^{-4}	2×10^{-4}	5×10^{-4}
k_6	9×10^{-3}	2×10^{-3}	2×10^{-14}	2×10^{-4}	2×10^{-2}
k_5	8×10^{-8}	2×10^{-4}	6×10^{-4}	4×10^{-3}	2×10^{-2}
k_4	2×10^{-14}	--	--	--	--
$k_{net,R}$	2×10^{-2}	7×10^{-3}	2×10^{-3}	7×10^{-3}	4×10^{-2}
k_{87}	--	1×10^{-13}	2×10^{-2}	2×10^{-2}	5×10^{-2}
k_{86}	--	4×10^{-3}	4×10^{-8}	1×10^{-11}	1×10^{-14}
k_{85}	--	2×10^{-14}	2×10^{-14}	2×10^{-2}	1×10^{-2}
k_{84}	--	--	--	--	--
$k_{net,8}$	--	4×10^{-3}	2×10^{-2}	4×10^{-2}	7×10^{-2}
k_{76}	3×10^{-6}	2×10^{-11}	5×10^{-3}	3×10^{-2}	2×10^{-14}
k_{75}	4×10^{-4}	4×10^{-4}	4×10^{-3}	1×10^{-3}	1×10^{-1}
k_{74}	2×10^{-14}	--	--	--	--
$k_{net,7}$	4×10^{-4}	4×10^{-4}	1×10^{-2}	3×10^{-2}	1×10^{-1}
k_{65}	2×10^{-14}	3×10^{-12}	2×10^{-9}	1×10^{-2}	1 *
k_{64}	2×10^{-2}	--	--	--	--
$k_{net,6}$	2×10^{-2}	3×10^{-12}	2×10^{-9}	1×10^{-2}	1 *
$k_{54} = k_{net,5}$	5×10^{-13}	--	--	--	--

*Upper bound of allowed values.

Table 2 Summary of kinetic data for reporter metabolism in each cell type. Ranges represent the 90% confidence interval for each value as determined by a bootstrapping method. The major fragments for HeLa and LNCaP lysates were determined in reference [16].

Cell Type	Half-Life (min)	Initial Rate (pmol mg^{-1} s^{-1})	Major Fragment during First Hour
E. coli	21-44	0.09-0.18	6FAM-GRP(nR)AFT
D. discoideum	82-103	0.038-0.047	6FAM-GRP(nR)AFT
S. cerevisiae	279-314	0.012-0.014	6FAM-GRP(nR)AFTF
HeLa (human cervical cancer)	86-105	0.037-0.045	6FAM-GRP(nR)A
LNCaP (human prostate cancer)	13-18	0.22-0.30	6FAM-GRP(nR)A

Electronic Supplemental Material

Interspecies Comparison of Peptide Substrate Reporter Metabolism using Compartment-Based Modeling

Allison J. Tierney,[a] Nhat Pham,[b] Kunwei Yang,[a] Brooks K. Emerick,[b] and Michelle L. Kovarik[a,*]

[a]Department of Chemistry, Trinity College, 300 Summit Street, Hartford, CT 06106

[b]Department of Mathematics, Trinity College, 300 Summit Street, Hartford, CT 06106

*Corresponding Author. Email: michelle.kovarik@trincoll.edu. Phone: 860-297-5275. Fax: 860-297-5129

Compartment-Based Model for *E. coli*

In lysates from eukaryotic cells (*D. discoideum*, *S. cerevisiae*, HeLa, and LNCaP) we detected fragments F_8 through F_5. In *E. coli* lysates, we observed fragments F_7 to F_4. As a result, a slightly different compartment-based model was required to describe this system (Fig. S1). This model is mathematically equivalent to the model in Fig. 3 of the article (used to describe data for the eukaryotic cells), and the same assumptions were applied. Namely, (1) there is no source of peptide fragments except for the initial reporter; (2) the parent peptide or any smaller fragments do not leave the system, but are simply converted to smaller fragments; (3) any larger peptide could be cleaved to form any smaller peptide, but only the N-terminal fragments, which retain the 6-FAM label, will be detected and (4) a linear system of equations is created, given that the system follows first-order kinetics and all rate constants are non-negative.

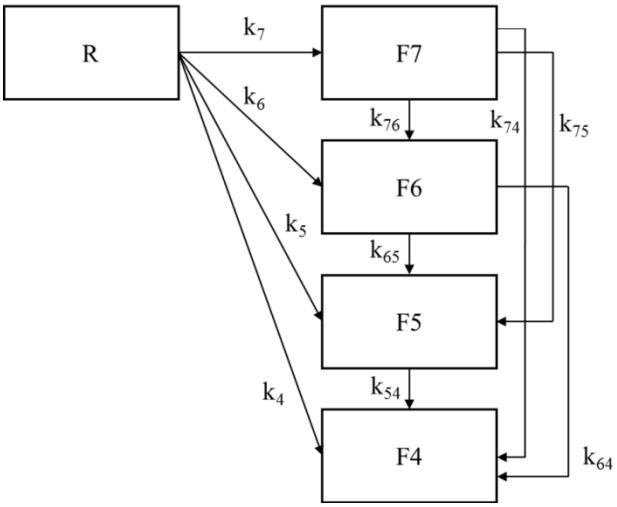


Fig. S1. Compartment-based model of reporter (*R*) metabolism into shorter N-terminal fragments (F_X) with *X* remaining amino acid residues after the N-terminal fluorescent label for *E. coli*.

Simplified Model for LNCaP

F_6 (the six amino acid fragment) was not observed in LNCaP lysates. This observation could be explained either of two ways: F_6 formed but was rapidly degraded to F_5 or F_6 was never formed. Results from the original model for eukaryotic cells (Fig. 3) suggested that F_6 formed but was rapidly degraded. To further test this hypothesis, we constructed an alternate, simplified model (Fig. S2) in which F_6 never formed.

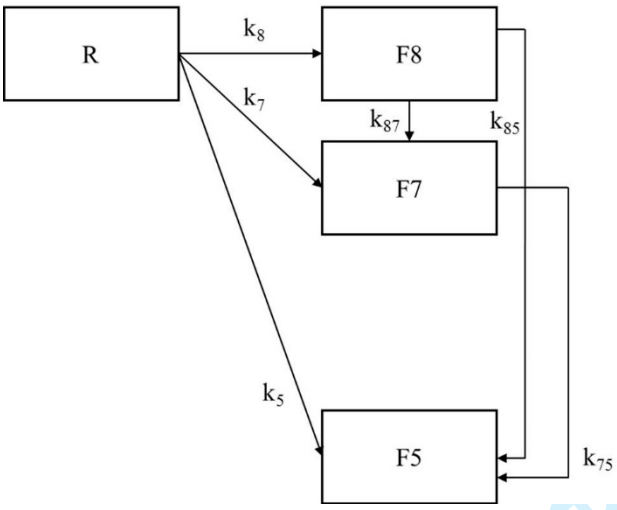


Fig. S2. Simplified compartment-based model of reporter (*R*) metabolism in LNCaP lysates, assuming that F_6 was never detected because it never formed.

Data Fitting

To fit our model to the data, we first solved the linear system that describes the compartment based model. Equations (1)-(5) found in the article and reproduced below represent the model in Fig. 3. An analogous series of equations was written and solved for the *E. coli* model in Fig. S1.

$$\frac{dR}{dt} = -(k_8 + k_7 + k_6 + k_5)R = -k_{net,R}R \quad (1)$$

$$\frac{dF_8}{dt} = -(k_{87} + k_{86} + k_{85})F_8 + k_8R = -k_{net,8}F_8 + k_8R \quad (2)$$

$$\frac{dF_7}{dt} = -(k_{76} + k_{75})F_7 + k_{87}F_8 + k_7R = -k_{net,7}F_7 + k_{87}F_8 + k_8R \quad (3)$$

$$\frac{dF_6}{dt} = -k_{65}F_6 + k_{76}F_7 + k_{86} + k_6R \quad (4)$$

$$\frac{dF_5}{dt} = k_{65}F_6 + k_{75}F_7 + k_{85}F_8 + k_5R \quad (5)$$

These solutions to these equations are found using basic techniques for solving differential equations and are shown in Equations (6)-(10):

$$R(t) = e^{-k_{net,R}t} \tag{6}$$

$$F_8(t) = \frac{k_8}{k_{net,R}-k_8}\left(e^{-k_{net,8}t} - e^{-k_{net,R}t}\right) \tag{7}$$

$$F_7(t) = \frac{k_{87}k_8(k_{net,7}-k_{net,R})-(k_{87}k_8+k_7k_{net,8}-k_7k_{net,R})(k_{net,7}-k_{net,8})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,8})((k_{net,7}-(k_{net,R})} e^{-k_{net,7}t} + \frac{k_{87}k_8}{(k_{net,R}-k_{net,8})(k_{net,7}-k_{net,8})} + \frac{k_{87}k_8+k_7(k_{net,8}-k_{net,R})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})} \tag{8}$$

$$F_6(t) = -\left[\frac{k_{76}k_{87}k_8(k_{net,7}-k_{net,R})+k_{76}(k_7k_8-k_{87}k_8-k_7k_{net,R})(k_{net,7}-k_{net,8})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})(k_{net,7}-k_{net,8})(k_{net,6}-k_{net,7})} + \frac{k_{76}k_{87}k_8+k_{86}k_8(k_{net,7}-k_{net,8})}{(k_{net,6}-k_{net,8})(k_{net,7}-k_{net,8})(k_{net,R}-k_{net,8})} + \frac{k_{76}k_{87}k_8+k_{76}k_7(k_{net,8}-k_{net,R})+k_6k_8(k_{net,7}-k_{net,R})+k_6(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})(k_{net,6}-k_{net,R})}\right] e^{-k_{net,6}t} + \frac{k_{76}k_{87}k_8(k_{net,7}-k_{net,R})+k_{76}(k_7k_8-k_{87}k_8-k_7k_{net,R})(k_{net,7}-k_{net,8})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})(k_{net,7}-k_{net,8})(k_{net,6}-k_{net,7})} e^{-k_{net,7}t} + \frac{k_{76}k_{87}k_8+k_{86}k_8(k_{net,7}-k_{net,8})}{(k_{net,6}-k_{net,8})(k_{net,7}-k_{net,8})(k_{net,R}-k_{net,8})} e^{-k_{net,8}t} + \frac{k_{76}k_{87}k_8+k_{76}k_7(k_{net,8}-k_{net,R})+k_6k_8(k_{net,7}-k_{net,R})+k_6(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})}{(k_{net,8}-k_{net,R})(k_{net,7}-k_{net,R})(k_{net,6}-k_{net,R})} e^{-k_{net,R}t} \tag{9}$$

$$F_5(t) = 1 - R(t) - F_8(t) - F_7(t) - F_6(t) \tag{10}$$

We note that the function $F_5(t)$ is completely determined by the previous four solutions; thus, the parameters k_{65}, k_{75}, k_{85} and k_5 are determined after finding best fit values for $k_{net,6}$, $k_{net,7}$, $k_{net,8}$, and $k_{net,R}$. This leaves ten model parameters that we wish to fit to the experimental data points. We note here that the solution functions are written to show what parameters they depend on as well as the independent variable, t. For example, the function $R(t)$ depends on the parameters $k_{net,R}$, so we write $R(t, k_{net,R})$.

The experiments yielded time-series data collected over the course of one or three hours (for human and non-human cell types, respectively) for the reporter and each N-terminal fragment of the reporter. The measurements were taken at ten different time points, and there were three measurements at each time point. We used the average of the three values for our data-fitting purposes and considered the standard deviation of these measurements for our Monte-Carlo simulations (see below). We found the best fit parameters for each segment by matching Equations (6)-(10) (i.e., the explicit solutions to Equations (1)-(5)) to the given time-series data in the least-squares sense. Because we quantified each fragment individually, we first found the best fit parameter for the parent peptide, $R(t)$, then found the best fit parameters for $F_8(t)$, and so on. Our method for fitting all ten parameters is summarized below.

Fitting data for the reporter (*R*)

Given the data for the degradation of the reporter (*R*), we found the parameter $k_{net,R}$ using the least-squares data fitting approach with the function $R(t; k_{net,R})$. The function $R(t; k_{net,R}) = e^{-k_{net,R}t}$ only has one parameter, namely $k_{net,R}$, the degradation rate of the reporter. We linearized the function by taking the natural logarithm, $\ln(R) = -k_{net,6}t$, and matched it to the logarithm of the given data. Using the least-squares regression line, we minimized the residual error and found the best fit for the parameter $k_{net,R}$.

Fitting data for F_8

We substituted this best fit value for $k_{net,R}$ into the solution for F_8, which means F_8 is dependent on $k_{net,8}$ and k_8. We fit the function $F_8(t; k_{net,8}, k_8)$ to the data for F_8. Since the function $F_8(t; k_{net,8}, k_8)$ cannot be linearized like the function $R(t; k_{net,R})$, we used a nonlinear least-squares approach. Assuming the error in measurements is taken from the standard normal curve, we defined the function

$$f_8(k_{net,8}, k_8) = \| F_8([\text{Time Data}]; k_{net,8}, k_8) - [F_8 \text{ Data}]\| \qquad (11)$$

as the error function to be minimized. Here, [Time Data] represents the ten time points, [F_8 Data] is the average measurements of F_8 concentration at each of ten time points, and $\| \cdot \|$ is the Euclidean norm. Using the MatLab lsqnonlin function, we ran the iterative method for each trial until a minimum was found in ($k_{net,8}$; k_8) parameter space and reported the best fit parameters, corresponding to the minimum residual error, in Table 1 in the article. We assume that the degradation constant, $k_{net,8}$ is non-negative and bounded above by one. Further, to ensure that other parameters are non-negative, we bound the parameter k_8 above by $k_{net,R}$. The reason for this is because we know $k_{net,R} - k_8 = k_7 + k_6 + k_5 \geq 0$, which implies $k_{net,R} - k_8 \geq 0$ and hence $k_8 \leq k_{net,R}$. A similar argument can be made for the remainder of the parameters below.

Fitting data for F_7

We substituted the best fit parameter values for $k_{net,8}$ and k_8 into Equation (8) for F_7 to form the function $F_7(t; k_{net,7}, k_{87}, k_7)$. We defined the error function as

$$f_7(k_{net,7}, k_{87}, k_7) = \left\| F_7([\text{Time Data}];\ k_{net,7}, k_{87}, k_7) - [F_7 \text{ Data}] \right\| \qquad (12)$$

We used the nonlinear least-squares method to fit the data for F_7 using 27,000 trials. To ensure all parameter values were non-negative, we let $k_{net,7}$ be bounded above by 1, k_{87} was bounded above by $k_{net,8}$, and k_7 was bounded above by $k_{net,R} - k_8$. All values were bounded below by 0.

Fitting data for F_6 and F_5

We substituted the best fit parameter values for $k_{net,7}$, k_{87}, and k_7 into F_6 and F_5 to form the functions $F_6(t; k_{net,6}, k_{86}, k_{76}, k_6)$ and $F_5(t; k_{net,6}, k_{86}, k_{76}, k_6)$. Therefore, to find the best parameters for $k_{net,6}$, k_{86}, k_{76} and k_6, we minimized the error between the functions F_6 and F_5 with the data given for the two fragments, respectively. We defined the following error function

$$f_{65}(k_{net,6}, k_{86}, k_{76}, k_7, k_6) = \left\| \begin{bmatrix} F_6(\text{Time Data}; k_{net,6}, k_{86}, k_{76}, k_6) \\ F_5(\text{Time Data}; k_{net,6}, k_{86}, k_{76}, k_6) \end{bmatrix} - \begin{bmatrix} F_6 \text{ Data} \\ F_5 \text{ Data} \end{bmatrix} \right\| \qquad (13)$$

We used the nonlinear least-squares method to fit F_6 and F_5 using 50,625 trials. To ensure every parameter is non-negative, we constrain the parameters for F_6 to a lower bound of zero and the following upper bounds: $k_{net,6}$ is bounded above by 1; k_{86} is bounded above by $k_{net,8}$ - k_{87}; k_{76} is bounded above by $k_{net,7}$; and k_6 is bounded above by $k_{net,R}$ - k_8 - k_7.

In each step of the data-fitting procedure above, we minimized the residual norm between our functions evaluated at the time data and the measured data of each peptide concentration. The minimum residuals for each of the parameters are given in Table S1.

Table S1. Minimum residuals from the fitting of Equations (6)-(10) to the experimental data for each peptide and species.

<table>
<tr><th></th><th>E. coli</th><th>D. discoideum</th><th>S. cerevisiae</th><th>HeLa</th><th>LNCap</th><th>LNCaP simplified</th></tr>
<tr><td>R</td><td>1.47×10^{-1}</td><td>9.32×10^{-2}</td><td>1.16×10^{-1}</td><td>2.10×10^{-2}</td><td>7.94×10^{-2}</td><td>7.94×10^{-2}</td></tr>
<tr><td>F_8</td><td>--</td><td>2.77×10^{-4}</td><td>3.02×10^{-4}</td><td>9.75×10^{-5}</td><td>3.21×10^{-5}</td><td>3.21×10^{-5}</td></tr>
<tr><td>F_7</td><td>2.36×10^{-2}</td><td>2.30×10^{-3}</td><td>6.90×10^{-4}</td><td>6.53×10^{-6}</td><td>1.13×10^{-5}</td><td>4.36×10^{-3}</td></tr>
<tr><td>F_6</td><td>1.29×10^{-3}</td><td rowspan="2">4.24×10^{-3}</td><td rowspan="2">1.69×10^{-2}</td><td rowspan="2">2.24×10^{-4}</td><td rowspan="2">4.63×10^{-3}</td><td>--</td></tr>
<tr><td>F_5</td><td rowspan="2">1.98×10^{-1}</td><td>4.36×10^{-3}</td></tr>
<tr><td>F_4</td><td>--</td><td>--</td><td>--</td><td>--</td><td>--</td></tr>
</table>

Given only ten data points for each peptide, we found it more economical to fit the data in a sequential order rather than all at the same time. In this way, we searched a much smaller parameter space in each step of the process. We compared both methods, and the sequential method described above yielded smaller residual norms. The nonlinear least-squares method requires an initial condition, and we found that different initial conditions sometimes yielded different best-fit parameters. To deal with this, we ran the same simulation with thousands of different initial conditions and chose the best-fit parameter to be the parameter set that yielded the minimum residual error among all trials. These values are given for each organism in Table 1 in the article. We compared the best-fit value to the most frequent value and concluded that generally they are very close in value within at least in order of magnitude. In cases where the best-fit value differs from the most frequent value, the error function likely has a broad local minimum that corresponds to the most frequent value and a sharper, deeper global minimum corresponding to the best-fit value. In this situation, most initial conditions will terminate to the broad, local minimum even though a better fit is obtained at the deeper global minimum. For this reason, we report values of k (Table 1) that correspond to the lowest minimum residuals (Table S1), even when they were not among the most frequently found parameter values.

Bootstrapping

We considered the error in each measurement to create a confidence interval for the degradation constant for the full-length reporter, $k_{net,R}$, for each organism. Each assay was repeated in triplicate, so for each time point, t, we had an average reporter concentration ($\bar{R}$) and standard

deviation (s) based on the three measurements. We assumed each measurement was a random variable that followed a Normal distribution with mean, $\bar{R}$ and standard deviation, s. Using a Monte Carlo simulation, we generated a new, hypothetical data point called $\hat{R}$, at each time value, t, from the Normal distribution $N(\bar{R}, t)$. We constrained the hypothetical data points such that values below the limit of quantitation were assigned a value of zero and the maximum value was 1 since peptide concentrations were reported as fraction of total peak area. We generated 1000 data sets based on these criteria and fit the function $R(t) = e^{k_{net,R}t}$ to each data set as described above. From this, we obtained histograms for the best fit parameter $k_{net,R}$ from which we gain a confidence interval on our best fit value for $k_{net,R}$ (Table S2). This essentially allows us to provide an idea of how the error in our measurements effects the confidence in our best fit values for the reporter degradation constant.

Table S2. Bootstrapped 90% confidence intervals for $k_{net,R}$.

	$k_{net,R}$ (min^{-1})		
E. coli	1.57×10^{-2}	-	3.33×10^{-2}
D. discoideum	6.76×10^{-3}	-	8.49×10^{-3}
S. cerevisiae	2.20×10^{-3}	-	2.48×10^{-3}
HeLa	6.60×10^{-3}	-	8.07×10^{-3}
LNCaP	3.94×10^{-2}	-	5.38×10^{-2}

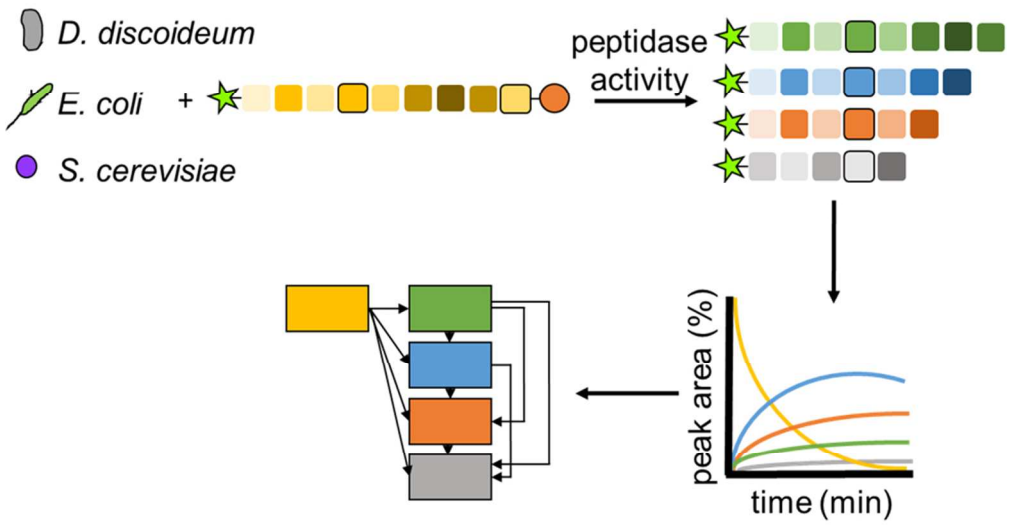
Reporter Degradation and *D. discoideum* Social Development

To determine whether the reporter was metabolized differently in *D. discoideum* cells undergoing social development, we prepared lysates using the procedure described in the main article from cells that had been resuspended in development buffer (DB; 5 mM Na_2HPO_4, 5 mM KH_2PO_4, 1mM $CaCl_2$ and 2mM $MgCl_2$) at a density of 10^7 cells/mL for 2, 4, or 6 h. In this cell-dense, nutrient-free environment, *D. discoideum* cells begin gene expression for the initial stages of their social life cycle.

Table S3. Summary of kinetic data for VI-B metabolism in each cell type. Ranges represent the 90% confidence interval for each value as determined by the bootstrapping method described above.

Development Time (h)	Half-Life (min)	Initial Rate (pmol mg^{-1} s^{-1})	Major Fragment during First Hour
0	82-103	0.038-0.047	6FAM-GRP(nR)AFT
2	79-89	0.043-0.049	6FAM-GRP(nR)AFT
4	83-92	0.042-0.047	6FAM-GRP(nR)AFT
6	94-110	0.035-0.041	6FAM-GRP(nR)AFT

Minimum residuals for $k_{net,R}$ for the 2, 4, and 6 h development times were 5.3×10^{-2}, 8.6×10^{-2}, and 1.4×10^{-1}, respectively.



Graphical Abstract

82x44mm (300 x 300 DPI)