Published in final edited form as:

Methods Enzymol. 2019 ; 622: 221–248. doi:10.1016/bs.mie.2019.02.023.

Selection and Optimization of Enzyme Reporters for Chemical Cytometry

Angela Proctor^a, Qunzhao Wang^b, David S. Lawrence^{a,b}, Nancy L. Allbritton^{*,a,c}

^aDepartment of Chemistry, University of North Carolina, Chapel Hill, NC, 27599, USA

^bDepartment of Chemical Biology and Medicinal Chemistry, School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA

^cJoint Department of Biomedical Engineering, University of North Carolina, Chapel Hill, NC, 27599, USA and North Carolina State University, Raleigh, NC, 27695, USA

Abstract

Chemical cytometry, sensitive analytical measurements of single cells, reveals inherent heterogeneity of cells within a population which is masked or averaged out when using bulk analysis techniques. A particular challenge of chemical cytometry is the development of a suitable reporter or probe for the desired measurement. These reporters must be sufficiently specific for measuring the desired process; possess a lifetime long enough to accomplish the measurement; and have the ability to be loaded into single cells. This chapter details our approach to rationally design and improve peptide substrates as reporters of enzyme activity utilizing chemical cytometry. This method details the iterative approach used to design, characterize, and identify a peptidase-resistant peptide reporter which acts as a kinase substrate within intact cells. Small-scale, rationally designed peptide libraries are generated to rapidly and economically screen candidate reporter peptides for substrate suitability and peptidase resistance. Also detailed are strategies to characterize and validate the designed reporters by determining kinetic parameters, intracellular substrate specificity, resistance to degradation by intracellular peptidases, and behavior within lysates and intact cells.

Keywords

chemical cytometry; peptide reporter development; enzyme activity measurement; kinase substrate; peptide libraries; single cell; enzyme activity

1. Introduction

It has become increasingly apparent that seemingly identical cells in a population can actually be a heterogeneous mixture differing at genome, transcriptome, and/or proteome levels. Analysis of this heterogeneity amongst cells within a population is critical to provide deeper understanding of cell behavior in response to intrinsic differences and extrinsic inputs in both normal and diseased states. (Keating et al., 2018; Mincarelli et al., 2018; Wang and

*Corresponding Author: nallbri@unc.edu (N.L. Allbritton).

Navin, 2015) For example, lineage tracing of somatic mutations in the genome can provide detailed information about mosaicism in normal tissue.(Biesecker and Spinner, 2013; D’Gama and Walsh, 2018; Mincarelli et al., 2018) Furthermore, the ramifications of clonal evolution in the development of diseased states can be realized as mutations are traced to differences in single cells; for example, in formation of both solid and blood cancers, immunological responses, and neurodegenerative diseases.(D’Gama and Walsh, 2018; Keating et al., 2018; Mincarelli et al., 2018; Ortega et al., 2017; Verheijen et al., 2018) The clinical relevance of cellular heterogeneity is increasingly realized as patient-specific therapies are targeted toward distinct mutations in individualized therapy.(Keating et al., 2018; Ortega et al., 2017) Although the importance of single-cell analysis is clear, significant challenges must be overcome to obtain relevant and reliable results. These challenges stem from the small size of the single cell, with most mammalian cells ranging from 10 – 50 μm in diameter and 1 – 100 pL in volume.(Alberts, 2008; Wang and Navin, 2015) Packed into this miniature environment is a dense assortment of nucleic acids, lipids, carbohydrates, proteins, and metabolites which can confound sample collection and assay readout (Table 1; (Alberts, 2008; Ortega et al., 2017; Wang and Navin, 2015; Zeng et al., 2018)). Many targeted analytes such as key signaling proteins are present within this complex milieu at vanishingly small copy numbers (10 – 1,000 molecules).(Huang et al., 2007) Additionally, the small size of the cell necessitates low detection limits and often prevents multiple sampling from the same cell, making it challenging to parse out technical variability of the assay method from true biological variability between cells.(Keating et al., 2018; Macaulay et al., 2017) The ability to reliably characterize single cells is of utmost importance in basic biological understanding, disease etiology, and development of personalized medicine.

Multiple technologies exist for analyzing the contents of single cells, including single-cell DNA sequencing, RNA expression analysis, and protein level and activity measurements. (Hu et al., 2018; Narrandes and Xu, 2018; Ortega et al., 2017; Wang and Navin, 2015) Single-cell DNA sequencing is generally accomplished by next generation sequencing after whole genome amplification.(Zeng et al., 2018; Zhu et al., 2018) Though this method can provide detailed genetic information, challenges in these techniques arise due to the low amount of DNA in a single cell (2 copies of each gene), necessitating amplification prior to sequencing. This can lead to non-uniform coverage of the genome and bias in the results as early technical errors are amplified, yielding overestimation of certain alleles coupled with complete loss of others.(Mincarelli et al., 2018) Single-cell RNA sequencing provides transcriptome information, yet also relies on amplification of the starting material and falls prey to some of the same sampling biases that can occur in DNA sequencing.(Zeng et al., 2018) Further bias occurs when higher abundance mRNA (~100,000 copy numbers) are more likely to be amplified than lower abundance mRNA (~10 copy numbers), skewing interpretation of true variability in the cells.(Mincarelli et al., 2018; Wagner et al., 2016) Protein amount can be measured at the single-cell level via antibody-based techniques including FACS (fluorescence activated cell sorting), phosphoflow cytometry, microscopic imaging, and single-cell Western blotting; but these are limited by antibody specificity and purity, a non-trivial issue.(Baker, 2015; Bougen-Zhukov et al., 2017; Su et al., 2017) Protein activity in single cells can be measured with energy transfer probes or fluorescent substrates,

yet face the challenge of loading probes into viable cells. (Goryashchenko et al., 2018; Kovarik and Allbritton, 2011; Rowland et al., 2015) Proximity ligation assays have also been used to assess enzyme activity, but this method relies on genetic transformation of cells, a difficulty when utilizing primary samples obtained directly from a patient. (Li et al., 2017) In summary, while several methods have been described for analyzing single cells, multiple challenges exist: antibody reliance can limit practical application; the small size of cells makes sampling requirements and detection challenging; and clinically obtained samples often contain a small number of cells (hundreds to thousands) of mixed origin (including fibroblasts, stroma, etc. in addition to the diseased cells of interest), compounding existing challenges with low sample sizes. (Keating et al., 2018) Methods to overcome these difficulties would be of high utility for single-cell analysis.

Chemical cytometry utilizes sensitive analytical techniques such as microelectrophoresis or mass spectrometry to analyze and characterize the contents of single cells. (Dovich, 2010; Dovich and Hu, 2003) Microelectrophoretic chemical cytometry is well-suited to address many of the challenges of single-cell analysis by virtue of its low sample volume requirements (pL to nL), superb resolving power (100s of analytes), and extremely low detection limits (10^{-21} moles), enabling separation of a large number of analytes from single cells with sub-pM detection limits. (Vickerman et al., 2018) These attributes, in combination with the absence of the need for cell genetic engineering, make this technique ideal for analyzing single cells from small mixed populations such as that from a primary clinical sample. Furthermore, chemical cytometry can provide a direct readout of enzyme activity, irrespective of the DNA, RNA, or protein levels in the cells, yielding valuable information about active cellular processes that cannot be obtained from genetic or expression information alone. However, a key challenge of this technique is design of a suitable probe that meets the strict requirements for reporting enzyme activity in single cells. These requirements include a probe or substrate with sufficient specificity to reliably report the activity of an enzyme or group of enzymes, the ability to load the probe into single cells without perturbing signaling pathways, and an intracellular lifetime sufficiently long to measure the desired process. Herein, we focus on the design of a reporter for chemical cytometry, specifically on the selection, optimization, and validation of intracellular reporters (typically short peptides) for protein kinase activity measurements by chemical cytometry.

2. Target Enzyme Attributes and Substrate Peptide Selection

Protein kinases catalyze the phosphorylation of serine, threonine, and tyrosine residues in both proteins and peptides using ATP as the phosphoryl donor. The human kinome is comprised of 518 protein kinases and 40 lipid kinases. The vast majority (478) of the former contain the so-called “eukaryotic protein kinase” (EPK) domain, a stretch of approximately 250 residues that encompass the catalytic region. (Duong-Ly and Peterson, 2013; Kostich et al., 2002) The substrate specificity of the large EPK family is controlled by three key determinants:

(1) The **active site substrate specificity** defines the ability of a protein kinase to phosphorylate one or more alcohol-bearing residues in the active site. Indeed, the EPKs are

divided into three distinct groups based on their active site specificity: the tyrosine protein kinases (TPKs), the serine/threonine protein kinases (SPKs), and the so-called dual specificity protein kinases that catalyze the phosphorylation all three types of alcohol-bearing residues.(Miller and Turk, 2018) However, the active site specificity of protein kinases is not limited to the three genetically encoded alcohol containing amino acids. A wide variety of unnatural residues are phosphorylated by protein kinases, including a structurally constrained tyrosine residue [(7-(S)-hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Htc)].(Kwon et al., 1994; Kwon et al., 1993; Lee et al., 1995; Prorok et al., 1989; Turner et al., 2016) Substrates containing Htc are particularly useful as probes of TPK activity since the corresponding phosphorylated product is resistant to dephosphorylation by intracellular protein phosphatases.(Turner et al., 2016) One of the key advantages of using peptides as protein kinase substrates is that unnatural residues are readily introduced during peptide synthesis. Non-naturally occurring residues have been used to endow peptide-based substrates with useful properties, including enhanced selectivity for specific protein kinases, resistance to intracellular proteolysis (*vide infra*), and photo-transformation from inactive to active substrates.

(2) The **sequence substrate specificity** of a protein kinase defines its preference for the amino acid sequence encompassing the phosphorylatable residue. This is also referred to as the consensus sequence of a protein kinase. This form of substrate specificity was initially derived from sequencing studies on intact phosphorylated protein substrates. However, starting in the 1990s, peptide-based libraries have been used to define the sequence specificity of individual protein kinases. Although these studies have generally been limited to the use of the standard genetically encoded amino acids, the sequence specificity of protein kinases extends well beyond the conventional 20 amino acids. For example, we have used a combination of D- and N-(Me) amino acids at various sites along the peptide chain to create effective protein kinase substrates that are resistant to proteolysis.(Mainz et al., 2016a; Mainz et al., 2016b; Proctor et al., 2012a, b; Proctor et al., 2016)

(3) **Multisite substrate specificity** describes interactions between a protein kinase and its substrate that occur outside of the kinase catalytic cleft. These ancillary protein-protein interactions are used to dock the protein substrate near the appropriate protein kinase. For example, the substrate specificity of different members of the mitogen-activated protein kinases is determined by their ability to recognize so-called D-site or DEF-site motifs on their substrates, an interaction that occurs distal from the active site region.(Miller and Turk, 2018) Ancillary non-active site interactions can be recapitulated in peptides by appending the secondary recognition motif to the consensus sequence via a linker region.(Profit et al., 2001)

3. Iterative Design Approach Overview

An iterative design strategy is used to identify a peptidase-resistant reporter that acts as an efficient substrate for kinases in cell lysates as well as intact cells (Figure 1). In this process, small-scale, rationally designed libraries are used to screen peptides and identify peptide bonds that require stabilization in the face of cytosolic proteases while maintaining similar or improved kinetic properties for the targeted kinase. Critical to the method's success is that

library members are simultaneously screened for resistance to all cytosolic peptidases and proteases at their relative cellular concentrations. This feature enables a rational redesign of the peptide to improve its lifetime within the harsh intracellular environment in which peptides must be efficiently degraded for recycling of their amino acids.(Cullen and Steinberg, 2018; Goldberg and Dice, 1974; Goldberg and St. John, 1976; Wolf and Menssen, 2018) In brief, an initial peptide is selected from the literature and is typically derived from protein regions known to be phosphorylated *in vivo* or from a known consensus sequence of a kinase (Section 2). The initial peptide is assessed for its time to 50% phosphorylation by the kinase, degradation half-life within a cell lysate, and location of the initial or dominant peptidase cleavage site. A small-scale, rationally designed peptide library is then constructed by replacing the amino acids surrounding the susceptible peptide bond or by modifying the bond itself to enhance peptidase resistance. Care is taken to use replacement residues or bonds that are tolerated by the kinase. The library is screened for phosphorylation by the kinase, degradation in a lysate, and its initial peptidase cleavage site as described above. The process is repeated, with each sequential round of redesign yielding a peptide reporter with increased proteolytic resistance and often with significantly improved kinetic constants for the kinase. This strategy has yielded peptides with as much as a 15X greater lifetime in cell lysates (>90 min) while sustaining as much as a 7X increase in k_{cat} with screening of no more than 18 peptides.(Proctor et al., 2012a, b; Proctor et al., 2016; Turner et al., 2016) As such, the method yields an efficient and cost-effective strategy to rapidly develop peptidase-resistant kinase substrates for use in both lysate-based assays as well as single-cell assays of kinase activity using chemical cytometry.

4. Design and Synthesis of a Small-Scale Library

Substrate peptides are synthesized containing an amidated C-terminus using the TGR resin and standard Fmoc solid phase peptide synthesis (SPPS). All peptides contain an N-terminal fluorescein moiety (carboxyfluorescein; *FAM*), which is used to enable detection of both the nonphosphorylated and phosphorylated forms of the peptide by capillary zone electrophoresis (CZE). As briefly noted above, the literature serves as a guide for the starting sequence of the initial peptide. For example, *FAM*-GGAYAAPFKKKA-amide, an Abl TPK substrate, was based on a related sequence acquired from a combinatorial peptide library. (Songyang et al., 1995) The sequence, *FAM*-EDDEYEEV-amide, was utilized for the epidermal growth factor receptor (EGFR) TPK (Guyer et al., 1994) whereas the peptide, *FAM*-GRPRAATFAEG-amide, serves as a substrate for Akt and was identified from a peptide library (Alessi et al., 1996). The iterative design strategy outlined in the previous section was then applied to improve specific properties. For example, protease-resistance is an essential property for peptides that will be used to probe intracellular protein kinase activity. Incubation of the parent peptides with cell lysates and analysis by CZE revealed the formation of fragments and thus, by inference, those sites that are particularly sensitive to proteolysis (Section 6.3). Derivatives of the starting peptide were subsequently prepared containing alternative amino acids at the protease susceptible sites. These peptides were subsequently analyzed for their ability to serve as substrates for their respective protein kinases as well as for their resistance to proteolysis.

5. Substrate Characterization of Candidate Peptide Reporters in a Library

5.1 Measurement of Substrate Phosphorylation

The ability of the designed peptide reporters to act as substrates for the target kinase is assessed using *in vitro* reactions of the candidate reporters with purified enzyme. Phosphorylation of each reporter is measured over time and the time to reach 50% phosphorylation ($t_{50\% \text{ P}}$) is determined for each candidate reporter. This enables rapid assessment of each reporter as a substrate with selection of the substrate with the lowest $t_{50\% \text{ P}}$ generally selected as the lead peptide sequence for further optimization. The poorest substrates are rapidly rejected without extensive reagent costs and time needed to measure kinetic constants.

5.2 Substrate Phosphorylation ($t_{50\% \text{ P}}$) Assay

1. Prepare and store enzyme stock according to the manufacturer. To minimize impact of lot to lot variability in enzyme activity, purchase a sufficient quantity of enzyme for all experiments, mix together, aliquot, and freeze.
2. Prepare and store additional assay components at 10X the final concentration, aliquot, and freeze to minimize assay to assay variability throughout the iterative process.
3. Use a known peptide substrate as a control alongside all $t_{50\% \text{ P}}$ assay trials. Begin with the recommended amount of enzyme provided by the manufacturer and use the substrate concentration at the K_M , which will give measurable changes in phosphorylation over time. Tip: If needed, adjust enzyme or substrate concentration to alter reaction rate. Once determined, utilize the same concentrations throughout the entire candidate reporter screening process.
4. To minimize variability between reactions run at the same time, mix all assay components together except for the ATP and the reporter and distribute this mixture into reaction vessels.
5. Add the appropriate reporter to each vessel.
6. Initiate the assay by adding ATP to start the reaction and incubate at 30 °C.
7. Collect sample aliquots from the reaction vessel over time and terminate the reaction in the aliquot by inactivating the enzyme (*e.g.* denature the enzyme via heating at 90 °C for 4 min.)
8. Freeze denatured samples at –20 °C until analysis (Section 7.3.1.)

6. Measurement of Resistance to Intracellular Proteases

6.1 Degradation Resistance Determination

To determine if candidate reporters are resistant to proteolytic degradation, they are incubated with active proteases and peptidases from cytosolic lysates. Resistance to degradation is quantified in terms of half-life ($t_{1/2}$), the time it takes for half of the intact reporter to degrade in a cell lysate without added protease and peptidase inhibitors.

Candidate peptides with longer $t_{1/2}$ values are more resistant to breakdown and will be further evaluated as likely substrates while those with short $t_{1/2}$ values will be rejected.

6.2 Lysate Stock Preparation

1. Use a human cell line relevant to the enzyme being studied (*e.g.* select a cell line with upregulated Akt activity, such as PANC-1, to study Akt.) Suggested cell lines include HeLa, PANC-1, LNCaP, Caco-2, K-562, or SKBr3.
2. Freeze a cell pellet suspended in a minimal volume of phosphate buffered saline (PBS) in liquid nitrogen and rapidly thaw at 37 °C for a total of three cycles.
3. Centrifuge at high speed (14,000*g*) to pellet cell debris and collect the supernatant containing the majority of the proteases and peptidases. Discard the cell debris pellet.
4. Quantify the total protein concentration of the lysate using whichever method is preferred.(Noble and Bailey, 2009)
5. Dilute the lysate to a final protein concentration of 3.33 mg mL⁻¹ in PBS, aliquot, and store at -20 °C for 6 months.

6.3 Reporter Degradation ($t_{1/2}$) Assay

1. Distribute lysate stock into reaction vessels so the final concentration of total protein will be 3 mg mL⁻¹.
2. Initiate the reaction by adding the reporter to the lysate and incubate at 37 °C.
3. Collect sample aliquots from the reaction vessel over time and terminate the reaction in the aliquot by inactivating the enzymes in the lysate (*e.g.* by adding an equal volume of 200 mM HCl to denature the enzymes.)
4. Freeze denatured samples at -20 °C until analysis (Section 7.3.2.)

7. Use of Capillary Zone Electrophoresis for Assay Readout

7.1 Generating Electrophoretic Controls for CZE

CZE separates analytes based predominately on shape and charge, exploiting differing electrophoretic mobilities to separate components in a mixture and coupling CZE to laser induced fluorescence detection (CZE-LIF) yields excellent detection limits (pM for most commercially available instruments).(Weinberger, 1993) The phosphorylated product of the reporter possesses two additional negative charges and a negligible change in mass compared to the substrate reporter, enabling facile separation and quantification of the two species via CZE. Additionally, peptide reporters cleaved by peptidases possess a lower molecular weight, a carboxylated terminus, and often a distinct net charge when compared to the intact reporter, enabling resolution between all possible species. To establish acceptable separations, appropriate standards and controls must be generated.

7.1.1 Solid Phase Peptide Synthesis

1. Use standard SPPS techniques to generate simple and/or short reporter fragment and phosphorylation standards.(Amblard et al., 2006)
2. Purify peptide standards with HPLC and verify molecular weight with mass spectrometry.
3. Aliquot and store peptide standards in 100 mM Tris buffer or water at -20°C at concentrations 100 μM .
4. Verify that standards contain only a single peak with CZE-LIF.

7.1.2 Pronase E Digestion—Pronase E, a commercially available enzymatic mixture of the extracellular fluid of *Streptomyces gresius*, contains numerous proteases and peptidases capable of hydrolyzing most peptide bonds.(Bermejo-Barrera et al., 1999; Vosbeck et al., 1975) This mixture of proteases (including five serine-type proteases, two Zn^{2+} -endopeptidases, two Zn^{2+} -leucine aminopeptidases, and one Zn^{2+} -carboxypeptidase) will cleave peptide bonds from both the N- and C-terminus and within the peptide to theoretically yield a mixture of all possible fragments.(Bermejo-Barrera et al., 1999) For long peptides or those fluorescently labeled on an interior amino acid, Pronase E can be a rapid and cost-effective method for generating a fragment mixture for use in identifying peptide bonds susceptible to degradation and can be used to ensure that the intact peptide reporter does not co-migrate with any of the fragments during CZE.(Mainz et al., 2015)

1. Incubate the reporter with multiple concentrations of Pronase E for varying times at 37°C .
2. Terminate the reaction by inactivating the enzymes (e.g. denature via heating at 90°C for 4 min.)
3. Confirm fragment generation by mass spectrometry and CZE-LIF. Note: It is likely that multiple samples with different incubation times will need to be combined to generate a mixture containing all possible fragments (e.g. a low Pronase E concentration for short times will yield fragments from the most labile bonds while higher concentrations for longer times may be required to yield fragments from bonds that are more stable.)
4. Freeze fragment mixtures at -20°C between use.

7.1.3 Alternative Methods—In the case that all peptide fragments are not generated from incubation with Pronase E, alternative enzymatic or chemical methods such as those used in proteomics can be utilized to generate additional fragments. Enzymatic methods include incubation with trypsin or chymotrypsin to cleave the carboxy side of specific residues—lysine or arginine (trypsin) or phenylalanine, tyrosine, and tryptophan (chymotrypsin).(Olsen et al., 2004; Switzar et al., 2013) Chemical methods can also be used to cleave specific residues C-terminal of methionine (cyanogen bromide) or cysteine (2-nitro-5-thiocyanobenzoic acid) or between an aspartic acid-proline bond (formic acid) or an asparagine-glycine bond (hydroxylamine). (Switzar et al., 2013) Combinations of these

enzymatic and chemical cleavage methods can be used to supplement SPPS and Pronase E degradation to obtain all fragments.

7.2 Separation of Peptides by CZE

The largest hurdle in CZE separations is the determination of an electrophoretic buffer capable of separating all species in a mixture, as a single CZE buffer will not be universally compatible. Peptide sequence and any secondary structure can dramatically alter electrophoretic mobilities of analytes and will affect separation efficiency and quality, so a buffer capable of separating the substrate and any potential products must be identified. A screen through various buffer conditions will be necessary to determine an acceptable buffer for the optimal separation.(Landers, 1996; Weinberger, 1993)

1. For the screen, use a 1:1 mixture of substrate:product standards to evaluate peak shape and resolution. Baseline resolution is a minimum requirement. Use software with the CZE instrument or alternative peak finding software (*e.g.* OriginLab) to analyze the electropherograms.
2. Start the screen with analysis of various buffer salts at a pH near the pKa so they have the greatest buffering capacity. Suggestions of common buffer salts are shown in Table 2. Select the buffer salt which provides the greatest resolution between the substrate and product.
3. Once a salt has been selected, screen the salt at multiple concentrations (suggested 50 mM increments) while holding the pH constant and select the buffer which provides the greatest resolution between the substrate and product.
4. Continue the screen by holding the salt and salt concentration constant and modify the pH of the buffer (suggested in 0.5 unit increments) to achieve the greatest resolution.
5. Lastly, buffer additives such as organic modifiers, bile salts, or surfactants can be screened to determine if they enhance the separation (Table 2.) Additives change the mechanism of separation (*e.g.* surfactants above the critical micellar concentration separate via MEKC, micellar electrokinetic chromatography, rather than CZE).(Landers, 1996; Weinberger, 1993) Screen through additives by holding the buffer salt, salt concentration, and pH constant while varying the identity of the additive to select that which yields the greatest resolution.
6. Finally, additive concentration can be varied to achieve the best separation, determined by the greatest resolution between all peaks and verification that each species is represented by a single peak with a unique migration time.

7.3 Interpretation of Separation Electropherograms

7.3.1 Substrate Phosphorylation ($t_{50\% \text{ P}}$) Assay

1. Electrophorese all sample aliquots from the $t_{50\% \text{ P}}$ assay using the determined electrophoretic conditions (Section 7.2.)

2. Obtain the corrected area (CA; the integrated area under the peak multiplied by the velocity of the analyte) for each of the peaks corresponding to substrate and phosphorylated product (Figure 2A) at each timepoint.
3. For each electropherogram, calculate the percentage of phosphorylated reporter by dividing the CA of the product peak by the sum of the CAs for the substrate and product peaks (Equation 1):

$$\%P = \frac{CA_{product}}{CA_{product} + CA_{substrate}} \times 100$$

%P = Phosphorylation percentage

CA_{product} = Corrected area under the product peak

CA_{substrate} = Corrected area under the substrate peak

4. For each candidate reporter being tested, plot the assay incubation time on the x-axis and the corresponding %P on the y-axis (Figure 2B).
5. Assumptions for the enzyme assays are the same assumptions defined by Michaelis-Menten kinetics (a single substrate is converted to product, the enzyme concentration does not change, and the enzyme/substrate complex amount is negligible.) (Segel, 1976) Based on these assumptions, fit the data with a logarithmic function to obtain a best fit equation with the following parameters (Equation 2):

$$y = A \ln(x) + B$$

y = Phosphorylated product (%)

x = Assay time

A and B are constants determined from the best fit of the data

6. To obtain a t_{50% P} value, solve for the assay time (x) when the substrate is 50% phosphorylated (Equation 3):

$$t_{50\% P} = e^{\frac{50\% - B}{A}}$$

7. Compare t_{50% P} values for all candidate reporters to each other and to control peptides. Lower t_{50% P} values indicate a faster phosphorylation rate and identify a more suitable substrate than one with a higher t_{50% P} value.

7.3.2 Reporter Degradation (t_{1/2}) Assay

1. Electrophorese all sample aliquots from the t_{1/2} assay using the determined electrophoretic conditions (Section 7.2.)

2. Obtain the corrected area (CA) for each of the peaks corresponding to substrate and phosphorylated product (Figure 3A) at each timepoint.
3. For each electropherogram, calculate the percentage of intact reporter by dividing the CA of the intact reporter by the sum of the CAs for all of the peaks (Equation 4):

$$\%I = \frac{CA_{\text{substrate}}}{CA_{\text{substrate}} + CA_{\text{products}}} \times 100$$

%I = Percentage of intact (undegraded) reporter

CA_{substrate} = Corrected area under the substrate peak

CA_{products} = Sum of the corrected area under each of the product peaks

4. For each candidate reporter being tested, plot the assay incubation time on the x-axis and the corresponding %I on the y-axis (Figure 3B).
5. Fit the data (excluding the 0 min timepoint) with an exponential decay function to obtain a best-fit equation with the following parameters (Equation 5):

$$y = Ae^{Bx}$$

y = Intact reporter (%)

x = Assay time

A and B are constants determined from the best fit of the data

6. To obtain a $t_{1/2}$ value, solve for the assay time (x) when there is 50% intact peptide remaining (Equation 6):

$$t_{\frac{1}{2}} = \frac{\ln\left(\frac{y}{A}\right)}{B}$$

7. Compare $t_{1/2}$ values for all candidate reporters to each other and to control peptides. Higher $t_{1/2}$ values indicate increased stability of the reporter and identify a more suitable substrate than one with a lower $t_{1/2}$ value.

8. Selection of Lead Peptide from Small Library Screen

8.1 Identification of Bonds Susceptible to Degradation

Amino acid sites targeted for modification are selected based on locations most susceptible to proteolytic degradation, with the thought that a non-native residue adjacent to the labile bond might impart stability to the peptide. CZE enables resolution of peptides with single amino differences for precise identification of peptidase fragmentation locations.

1. Use the lysate aliquots generated in the $t_{1/2}$ assay (section 6.3) and individual fragment standards (section 7.1.)
2. Electrophorese a $t_{1/2}$ assay aliquot.
3. Spike a single peptide fragment standard (Section 7.1.1) into the assay aliquot and electrophorese. Note: Analyze these samples sequentially to minimize any drifts in migration time that can occur in CZE and which make peak identification more difficult.
4. Overlay the two electropherograms (steps 2–3.) The peak corresponding to the peptide fragment will increase in area relative to all other peaks in the sample (Figure 4A.)
5. Continue until all peaks are identified (Figure 4B).
6. Use the samples generated in the $t_{1/2}$ assay (section 6.3) and analyze each electropherogram to obtain the CA for each peak.
7. For each candidate reporter at each timepoint sampled, calculate the percentage of each fragment peptide by dividing the CA of the fragment peptide by the sum of the CAs for all of the peaks (Equation 7):

$$\%F = \frac{CA_{Fragment}}{CA_{Total}} \times 100$$

$\%F$ = percentage of fragment peptide

$CA_{Fragment}$ = Corrected area under the fragment peptide peak

CA_{Total} = Sum of the corrected area under all peaks.

8. Plot assay incubation time on the x-axis and the percentage of fragment peptide on the y-axis to observe how each fragment behaves over time (Figure 4C).
9. Estimate time-averaged initial rates of fragment formation with early timepoints of the $t_{1/2}$ assay (*e.g.* monitoring the change in amount of peptide fragment over the first 5 min for a given cytosolic concentration.) Use for comparisons between different sets of candidate reporters to identify how quickly each fragment forms (Equation 8):

$$Rate = \frac{\Delta n}{\Delta t \times mass}$$

Rate = Time-averaged rate

$n = n_{final} - n_{initial}$, where n = moles of peptide fragment at the final and initial timepoint sampled

$t = t_{final} - t_{initial}$, where t = the time at which the final and initial samples were taken

mass = mass of protein in the lysate

10. In general, the fragment that forms at the fastest initial rate indicates the most labile bond; this site is selected for modification with non-native amino acids.

8.2 Replacement of Protease-Susceptible Residues

1. The first sites targeted for modification are those flanking either side of the bond most susceptible to degradation, indicated by the fragment that forms at the fastest time-averaged rate.
2. Resynthesize new peptide reporters by replacing either the N- or C-terminal amino acid around the most labile bond. Only a single change should be made on any individual peptide to more closely understand how changes affect the suitability of a candidate reporter.
3. Suggested alterations are residues that retain similar properties to the native residue, such as similar charge, hydrophobicity, or shape (Figure 5), to maintain the likelihood of the target enzyme recognizing the consensus sequence of the peptide reporter. These residues include, but are not limited to: D-amino acids, with the opposite chirality at the α -carbon atom; N-methylated amino acids, which contain an additional methyl group on the backbone nitrogen atom; β -amino acids, where the backbone nitrogen atom is moved from the α -carbon to a β -carbon atom; fluorinated R-groups, where fluorine atoms are added to increase the hydrophobicity; and constrained residues, where the backbone nitrogen atom is cyclized to the α -carbon, introducing rigidity into a previously flexible position. (Feng and Xu, 2016; Hanessian et al., 1997; Haviv et al., 1993; Muller, 2018; Yoder and Kumar, 2002)
4. Iterate through synthesis and characterization of substrate suitability and degradation resistance (Sections 5–8) until a superior peptide reporter is synthesized.

9. Characterization of Final Reporter

9.1 Determination of Kinetic Parameters

After the iteration process is complete and a candidate reporter has been selected as the final peptide reporter, further characterization should occur. These characterizations are only completed for the final peptide reporter since they are time-consuming, can incur considerable expense, and are not feasible for every candidate reporter synthesized. The characterizations include a detailed degradation profile (Section 8.1), kinetic analysis of the lead peptide reporter as a substrate for the target enzyme, and reporter behavior within cell lysates and intact cells. Two parameters which describe enzyme-substrate interaction can be obtained from the Michaelis-Menten equation. (Johnson and Goody, 2011; Segel, 1976) $V_{\max}$ is the maximum velocity achieved by the enzyme, the rate at which all enzyme active sites are saturated with substrate. The K_M is the substrate concentration at which the enzyme is operating at half of the $V_{\max}$, and is a measure of how easily a substrate is converted to product, with a lower K_M indicating a better substrate. The additional parameter k_{cat}

(turnover number) of an enzyme-substrate reaction is a measure of how quickly a given amount of enzyme can convert substrate to product in a given unit of time, with high k_{cat} values indicating rapid turnover. In the procedure described below, K_M and k_{cat} are estimated by generating progress curves of substrate phosphorylation by recombinant enzyme and fitting the curves with the time-dependent Michaelis-Menten equation using an analytical approximation of the Lambert function. (Golinski, 2011; Yun and Suelter, 1977) Benefits of this method over the more commonly used initial velocities method include savings in cost, time, and reagents consumed as well as being able to use more data gathered over the entirety of the progress curve, enabling a better fit of the overall data to estimate the kinetic parameters.

9.1.1 Generating Progress Curves

1. Use the same enzyme that was used throughout the screening process. Substrate concentration should be approximately 3X the K_M , which might necessitate some optimization to determine an approximate K_M value (*e.g.* for an approximate K_M of 5 μM , the three final substrate concentrations could be 12 μM , 15 μM , and 18 μM .)
2. Hold the enzyme concentration constant and run three phosphorylation assays with different substrate concentrations to generate the progress curves (section 5.1.3.)
3. Collect approximately 9 sample aliquots over a 60 min reaction time. Note: The shape of the progress curves should be different from each other over the timeframe of the reaction, with lower substrate concentrations showing a larger percentage of product formed compared to higher concentrations. If large amounts of substrate or enzyme are required, two different substrate concentrations at approximately 3X the K_M can be used to manage cost. It is also recommended that a known control peptide be run alongside the leading candidate reporter so a direct comparison can be made and to ensure that the enzyme is fully active.
4. Freeze samples at -20°C until CZE analysis (Section 7.3.2.)

9.1.2 Data Analysis

1. Obtain the CA for the peptide reporter and phosphorylated product peaks at each timepoint for each sample.
2. For each electropherogram, calculate the percentage of phosphorylated product as shown in equation 1.
3. Use this value to calculate the phosphorylated product concentration at each assay timepoint for each sample.
4. For each peptide reporter concentration assayed, plot assay incubation time on the x-axis and the corresponding product concentrations on the y-axis (Figure 6) to obtain the progress curves.

5. Use data fitting software (*e.g.* OriginLab, GraphPad Prism) to fit the progress curves with the time-dependent Michaelis-Menten equation using an analytical approximation of the Lambert function (obtain the detailed equation and derivation from Golini, supplementary data file 1, Page 8.)(Golini, 2011)
6. The fit will give the enzymatic parameters K_M and V_{max} , from which the turnover number can be calculated (Equation 9):

$$k_{cat} = \frac{V_{max}}{[E_T]}$$

k_{cat} = turnover number, in units of s^{-1}

V_{max} = maximum velocity of substrate to product conversion

$[E_T]$ = total enzyme concentration

9.2 Validation of Final Reporter

9.2.1 Loading Peptide Reporter into Cells—Cellular lysates and purified enzymes are used throughout the reporter optimization process because they are easy to use and are a cost-effective way to rapidly screen multiple substrates. However, cellular lysates provide a very different environment than an intact cell, where enzymatic processes are tightly controlled and regulated by various means, such as sequestering proteases and peptidases within membrane-bound organelles. A peptide within the cytosol of an intact cell is likely to be acted upon very differently than in the comparatively homogenous environment of a cellular lysate. Ultimately, the reporters developed with this iterative approach will be used to quantify enzyme activity within a native environment, that of the intact cell. Several considerations must be addressed at this point, such as how to get the reporter into the cells and whether to analyze the cells individually or in bulk. The greatest barrier to cell loading is the plasma membrane, which is quite effective at keeping most peptides from traversing into the cytoplasm. Though many methods exist for loading cargo into cells, such as microinjection, pinocytosis, myristoylation, and others (for thorough coverage of this topic, refer to reference (Stewart et al., 2018)), these methods are both dependent on the type of cell being used as well as the identity of the cargo. Unfortunately, there is not a catch-all method that can be applied for loading any type of cargo into any type of cell and the optimal loading strategy must be determined for each situation.

9.2.2 Analysis of Bulk Lysate

1. Load reporter into cells using the preferred method.(Stewart et al., 2018)
2. Incubate cells in complete growth medium at 37 °C.
3. Collect aliquots over time and terminate reaction by heat-inactivating the cells.
4. Lyse cells with freeze/thaw method (Section 6.2.)
5. Analyze with CE to identify and quantify products (phosphorylation and degradation; Section 7.3.)

9.2.3 Single-Cell Analysis—Single, intact cells loaded with the peptide reporter using any means can be analyzed via CZE to quantify enzyme activity at the single-cell level. Procedures for this are beyond the scope of this work, but the reader is referred to Nelson *et al.* for a detailed protocol (Nelson et al., 2007), as well as more recent publications from our lab using this technique.(Mainz et al., 2016a; Mainz et al., 2016b; Proctor et al., 2014; Proctor et al., 2012a; Proctor et al., 2016; Turner et al., 2016) This protocol is highly amenable to analysis of single cells from primary patient samples, enabling investigations into enzyme activity in individual cells from a specific patient.

9.2.4 Additional Validation—Further validation can be accomplished by modulating the target enzyme and/or closely-related enzymes to quantify how the reporter behaves in alternative environments and especially inside the cell. A combination of validation assays can be used to characterize the reporter. Table 3 lists potential validation assays.(Elbashir et al., 2002; Mainz et al., 2016a; Proctor et al., 2016; Xu et al., 2012) A major goal of these assays is to assess the specificity of the designed reporter for the targeted enzyme or class of enzymes.

10. Additional Modifications of Reporter for Improved Utility

Addition of a photocleavable moiety to the phosphorylation site enables precise control over reaction start time.(Mainz et al., 2016b) Furthermore, this can change the membrane permeability of the reporter, enabling facile distribution of the reporter into cellular cytoplasm.

1. Synthesize the photocleavable residue N-(9-Fluorenylmethyloxycarbonyl)-O-(4,5-dimethoxy-2-nitrobenzyl)-L-threonine (Fmoc-Thr(O-(4,5-dimethoxy-2-nitrobenzyl))-OH, or Fmoc-Thr(DMNB)-OH)) (Figure 7).
2. Couple the photocleavable residue, the remaining residues, and fluorescent label with standard SPPS techniques.
3. Characterize reporter in lysates and intact cells (Section 9).

11. References

- Alberts B, Alexander J; Lewis Julian; Raff Martin; Roberts Keith; Walter Peter, 2008 Cell Chemistry and Biosynthesis, Molecular Biology of the Cell, 5 ed Garland Science, Taylor & Francis Group, LLC, New York, pp. 45–124.
- Alessi DR, Caudwell FB, Andjelkovic M, Hemmings BA, Cohen P, 1996 Molecular basis for the substrate specificity of protein kinase B; comparison with MAPKAP kinase-1 and p70 S6 kinase. FEBS Lett 399, 333–338. [PubMed: 8985174]
- Amblard M, Fehrentz J-A, Martinez J, Subra G, 2006 Methods and protocols of modern solid phase peptide synthesis. Molecular Biotechnology 33, 239–254. [PubMed: 16946453]
- Baker M, 2015 Blame it on the antibodies. Nature 521, 274–276. [PubMed: 25993940]
- Bermejo-Barrera P, Fernandez-Nocelo S, Moreda-Pineiro A, Bermejo-Barrera A, 1999 Usefulness of enzymatic hydrolysis procedures based on the use of pronase E as sample pre-treatment for multi-element determination in biological materials. Journal of Analytical Atomic Spectrometry 14, 1893–1900.
- Biesecker LG, Spinner NB, 2013 A genomic view of mosaicism and human disease. Nature Reviews Genetics 14, 307–320.

- Bougen-Zhukov N, Loh SY, Lee HK, Loo LH, 2017 Large-scale image-based screening and profiling of cellular phenotypes. *Cytometry Part A* 91A, 115–125.
- Cullen PJ, Steinberg F, 2018 To degrade or not to degrade: mechanisms and significance of endocytic recycling. *Nature Reviews Molecular Cell Biology* 19, 679–696. [PubMed: 30194414]
- D’Gama AM, Walsh CA, 2018 Somatic mosaicism and neurodevelopmental disease. *Nature Neuroscience* 21, 1504–1514. [PubMed: 30349109]
- Dovich NJ, 2010 Origin, current status, and future perspectives of chemical cytometry, in: Lu C (Ed.). Wiley-VCH, Weinheim, pp. 1–19.
- Dovich NJ, Hu S, 2003 Chemical cytometry. *Current Opinion in Chemical Biology* 7, 603–608. [PubMed: 14580565]
- Duong-Ly KC, Peterson JR, 2013 The human kinome and kinase inhibition. *Curr Protoc Pharmacol* Chapter 2, Unit2 9.
- Elbashir SM, Harborth J, Weber K, Tuschl T, 2002 Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26, 199–213. [PubMed: 12054897]
- Feng Z, Xu B, 2016 Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomolecular concepts* 7, 179–187. [PubMed: 27159920]
- Goldberg AL, Dice JF, 1974 Intracellular protein degradation in mammalian and bacterial cells. *Annual Review of Biochemistry* 43, 835–869.
- Goldberg AL, St. John AC, 1976 Intracellular protein degradation in mammalian and bacterial cells: part 2. *Annual Review of Biochemistry* 45, 747–803.
- Golnik M, 2011 Explicit analytic approximations for time-dependent solutions of the generalized integrated Michaelis–Menten equation. *Analytical Biochemistry* 411, 303–305. [PubMed: 21241654]
- Goryashchenko AS, Khrenova MG, Savitsky AP, 2018 Detection of protease activity by fluorescent protein FRET sensors: from computer simulation to live cells. *Methods and Applications in Fluorescence* 6.
- Guyer CA, Woltjer RL, Coker KJ, Staros JV, 1994 Peptide substrate recognition by the epidermal growth factor receptor. *Arch Biochem Biophys* 312, 573–578. [PubMed: 8037473]
- Hanessian S, McNaughton-Smith G, Lombart HG, Lubell WD, 1997 Design and synthesis of conformationally constrained amino acids as versatile scaffolds and peptide mimetics. *Tetrahedron* 53, 12789–12854.
- Haviv F, Fitzpatrick TD, Swenson RE, Nichols CJ, Mort NA, Bush EN, Diaz G, Bammert G, Nguyen A, Rhutasel NS, Nellans HN, Hoffman DJ, Johnson ES, Greer J, 1993 Effect of N-methyl substitution of the peptide bonds in leuteinizing hormone-releasing hormone agonists. *Journal of Medicinal Chemistry* 36, 363–369. [PubMed: 8381183]
- Hu Y, An Q, Sheu K, Trejo B, Fan S, Guo Y, 2018 Single Cell Multi-Omics Technology: Methodology and Application. *Frontiers in Cell and Developmental Biology* 6.
- Huang B, Wu HK, Bhaya D, Grossman A, Granier S, Kobilka BK, Zare RN, 2007 Counting low-copy number proteins in a single cell. *Science* 315, 81–84. [PubMed: 17204646]
- Johnson KA, Goody RS, 2011 The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry* 50, 8264–8269. [PubMed: 21888353]
- Keating SM, Taylor DL, Plant AL, Litwack ED, Kuhn P, Greenspan EJ, Hartshorn CM, Sigman CC, Kelloff GJ, Chang DD, Friberg G, Lee JSH, Kuida K, 2018 Opportunities and Challenges in Implementation of Multiparameter Single Cell Analysis Platforms for Clinical Translation. *Cts-Clinical and Translational Science* 11, 267–276.
- Kostich M, English J, Madison V, Gheyas F, Wang L, Qiu P, Greene J, Laz TM, 2002 Human members of the eukaryotic protein kinase family. *Genome Biol* 3, research0043.0041–research0043.0012.
- Kovarik ML, Allbritton NL, 2011 Measuring enzyme activity in single cells. *Trends in Biotechnology* 29, 222–230. [PubMed: 21316781]
- Kwon YG, Mendelow M, Lawrence DS, 1994 The active site substrate specificity of protein kinase C. *J Biol Chem* 269, 4839–4844.

- Kwon YG, Mendelow M, Srinivasan J, Lee TR, Pluskey S, Salerno A, Lawrence DS, 1993 The active site substrate specificity of the cAMP-dependent protein kinase. *J Biol Chem* 268, 10713–10716. [PubMed: 8496138]
- Landers JP, 1996 *Handbook of Capillary Electrophoresis*, 2nd ed CRC Press, Boca Raton, FL.
- Lee TR, Niu J, Lawrence DS, 1995 The extraordinary active site substrate specificity of pp60c-src. A multiple specificity protein kinase. *J Biol Chem* 270, 5375–5380. [PubMed: 7534295]
- Li G, Montgomery JE, Eckert MA, Chang JW, Tienda SM, Lengyel E, Moellering RE, 2017 An activity-dependent proximity ligation platform for spatially resolved quantification of active enzymes in single cells. *Nature Communications* 8.
- Macaulay IC, Ponting CP, Voet T, 2017 Single-Cell Multiomics: Multiple Measurements from Single Cells. *Trends in Genetics* 33, 155–168. [PubMed: 28089370]
- Mainz ER, Dobes NC, Allbritton NL, 2015 Pronase E-based generation of fluorescent peptide fragments: tracking intracellular peptide fate in single cells. *Analytical Chemistry* 87, 7987–7995. [PubMed: 26171808]
- Mainz ER, Serafin DS, Nguyen TT, Tarrant TK, Sims CE, Allbritton NL, 2016a Single cell chemical cytometry of Akt activity in rheumatoid arthritis and normal fibroblast-like synoviocytes in response to tumor necrosis factor α . *Analytical Chemistry* 88, 7786–7792. [PubMed: 27391352]
- Mainz ER, Wang Q, Lawrence DS, Allbritton NL, 2016b An integrated chemical cytometry method: shining a light on Akt activity in single cells. *Angewandte Chemie International Edition* 55, 13095–13098. [PubMed: 27647713]
- Miller CJ, Turk BE, 2018 Homing in: Mechanisms of Substrate Targeting by Protein Kinases. *Trends Biochem Sci* 43, 380–394. [PubMed: 29544874]
- Mincarelli L, Lister A, Lipscombe J, Macaulay IC, 2018 Defining Cell Identity with Single-Cell Omics. *Proteomics* 18.
- Muller MM, 2018 Post-Translational Modifications of Protein Backbones: Unique Functions, Mechanisms, and Challenges. *Biochemistry* 57, 177–185. [PubMed: 29064683]
- Narrandes S, Xu W, 2018 Gene Expression Detection Assay for Cancer Clinical Use. *Journal of Cancer* 9, 2249–2265. [PubMed: 30026820]
- Nelson AR, Allbritton NL, Sims CE, 2007 Rapid sampling for single-cell analysis by capillary electrophoresis. Elsevier Academic Press Inc., San Diego, CA, pp. 709–722.
- Noble JE, Bailey MJA, 2009 Quantitation of Protein, in: Burgess RR, Deutscher MP (Eds.), *Guide to Protein Purification*, Second Edition, pp. 73–95.
- Olsen JV, Ong SE, Mann M, 2004 Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics* 3, 608–614. [PubMed: 15034119]
- Ortega MA, Poirion O, Zhu X, Huang SJ, Wolfgruber TK, Sebra R, Garmire LX, 2017 Using single-cell multiple omics approaches to resolve tumor heterogeneity. *Clinical and Translational Medicine* 6.
- 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. [PubMed: 24716819]
- Proctor A, Wang Q, Lawrence DS, Allbritton NL, 2012a Development of a peptidase-resistant substrate for single-cell measurement of protein kinase B activation. *Analytical Chemistry* 84, 7195–7202. [PubMed: 22881604]
- Proctor A, Wang Q, Lawrence DS, Allbritton NL, 2012b Metabolism of peptide reporters in cell lysates and single cells. *Analyst* 137, 3028–3038. [PubMed: 22314840]
- Proctor A, Zigoneanu IG, Wang Q, Sims CE, Lawrence DS, 2016 Development of a protease-resistant reporter to quantify BCR-ABL activity in intact cells. *Analyst* 141, 6008–6017. [PubMed: 27704073]
- Profit AA, Lee TR, Niu J, Lawrence DS, 2001 Molecular rulers: an assessment of distance and spatial relationships of Src tyrosine kinase Sh2 and active site regions. *J Biol Chem* 276, 9446–9451. [PubMed: 11118446]
- Prorok M, Sukumaran DK, Lawrence DS, 1989 The cyclic AMP-dependent protein kinase from bovine cardiac muscle is a homoserine kinase. *J Biol Chem* 264, 17727–17733. [PubMed: 2553690]

- Rowland CE, Brown CW, Medintz IL, Delehanty JB, 2015 Intracellular FRET-based probes: a review. *Methods and Applications in Fluorescence* 3.
- Segel IH, 1976 *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd ed John Wiley and Sons, United States.
- Songyang Z, Carraway KL 3rd, Eck MJ, Harrison SC, Feldman RA, Mohammadi M, Schlessinger J, Hubbard SR, Smith DP, Eng C, et al., 1995 Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* 373, 536–539. [PubMed: 7845468]
- Stewart MP, Langer R, Jensen KF, 2018 Intracellular Delivery by Membrane Disruption: Mechanisms, Strategies, and Concepts. *Chemical Reviews* 118, 7409–7531. [PubMed: 30052023]
- Su Y, Shi Q, Wei W, 2017 Single cell proteomics in biomedicine: High-dimensional data acquisition, visualization, and analysis. *Proteomics* 17.
- Switzar L, Giera M, Niessen WMA, 2013 Protein Digestion: An Overview of the Available Techniques and Recent Developments. *Journal of Proteome Research* 12, 1067–1077. [PubMed: 23368288]
- Turner AH, Lebharr 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 Chemical Biology* 11, 355–362. [PubMed: 26587880]
- Verheijen BM, Vermulst M, van Leeuwen FW, 2018 Somatic mutations in neurons during aging and neurodegeneration. *Acta Neuropathologica* 135, 811–826. [PubMed: 29705908]
- Vickerman BM, Anttila MM, Petersen BV, Allbritton NL, Lawrence DS, 2018 Design and Application of Sensors for Chemical Cytometry. *Acs Chemical Biology* 13, 1741–1751. [PubMed: 29376326]
- Vosbeck KD, Greenberg BD, Awad WM, 1975 Proteolytic enzymes of K-1 strain of streptomyces griseus obtained from a commercial preparation (pronase). *Journal of Biological Chemistry* 250, 3981–3987. [PubMed: 805135]
- Wagner A, Regev A, Yosef N, 2016 Revealing the vectors of cellular identity with single-cell genomics. *Nature Biotechnology* 34, 1145–1160.
- Wang Y, Navin NE, 2015 Advances and Applications of Single-Cell Sequencing Technologies. *Molecular Cell* 58, 598–609. [PubMed: 26000845]
- Weinberger R, 1993 *Practical Capillary Electrophoresis*, 1st ed Academic Press, London.
- Wolf DH, Menssen R, 2018 Mechanisms of cell regulation - proteolysis, the big surprise. *Febs Letters* 592, 2515–2524. [PubMed: 29790175]
- Xu WC, Allbritton N, Lawrence DS, 2012 Src Kinase Regulation in Progressively Invasive Cancer. *Plos One* 7.
- Yoder NC, Kumar K, 2002 Fluorinated amino acids in protein design and engineering. *Chemical Society Reviews* 31, 335–341. [PubMed: 12491748]
- Yun SL, Suelter CH, 1977 Simple method for calculating KM and Vmax from a single enzyme reaction progress curve. *Biochimica Et Biophysica Acta* 480, 1–13. [PubMed: 831830]
- Zeng ZW, Miao N, Sun T, 2018 Revealing cellular and molecular complexity of the central nervous system using single cell sequencing. *Stem Cell Research & Therapy* 9.
- Zhu ZY, Qiu S, Shao K, Hou Y, 2018 Progress and challenges of sequencing and analyzing circulating tumor cells. *Cell Biology and Toxicology* 34, 405–415. [PubMed: 29168077]

Figure 1:

The workflow of the described method. First, a starting substrate is synthesized and evaluated for substrate suitability ($t_{50\% P}$) and resistance to degradation ($t_{1/2}$). Formation of peptide fragments over time reveal the most labile bonds and a new candidate reporter library is synthesized by replacing the residues flanking the most susceptible bond with non-native amino acids. Candidate peptides are evaluated and the process is repeated until a suitable substrate has been determined. Once selected, the final substrate is further characterized and validated. Adapted with permission from (Proctor et al., 2012a.) Copyright 2012 American Chemical Society.

Figure 2:

Determining substrate suitability of the reporter. (A) Example electropherogram depicting unmodified reporter and the phosphorylated counterpart after reaction with recombinant enzyme. The corrected area under the peak is proportional to the number of moles of analyte in the injected volume. (B) Example graph demonstrating how to calculate a $t_{50\% P}$ value. The percentage of phosphorylated reporter is plotted on the y-axis with the assay time plotted on the x-axis. The data is fit with a logarithmic function and the $t_{50\% P}$ value is the time when 50% of the reporter has been phosphorylated.

Figure 3:

Determining the degradation resistance of the reporter. (A) Example electropherogram showing the intact reporter and peptide fragment peaks. The corrected area under the peak is proportional to the number of moles of analyte in the injected volume. (B) Example graph demonstrating how to calculate a $t_{1/2}$ value. The percentage of intact reporter is plotted on the y-axis as a function of incubation time with the lysate (x-axis.) The data is fit with an exponential decay curve and the $t_{1/2}$ value is the time at which 50% of the intact reporter remains.

Figure 4:

Determining peptide fragments formed when the reporter is incubated in a cell lysate. (A) Identification of the peak corresponding to an individual fragment peptide (Fragment ii, a two-residue fragment.) The bottom black trace is the electropherogram of the degraded reporter and the upper red trace is the same sample with the addition of Fragment ii standard. The single peak which increased in area is the peak that is due to Fragment ii. (B) The electropherogram of the degraded reporter with all peaks identified. Peptide fragment peaks are labeled with Roman numerals which correspond to the number of residues remaining (*e.g.* the fragment peak v corresponds to a five-residue fragment.) (C) Example graph showing fragment formation over time. The percentage of each peptide is plotted on the y-axis as a function of lysate incubation time (x-axis.) Time-averaged initial rates of formation can be determined utilizing the first two or three data points.

Figure 5:
Chemical structures of native amino acids and suggested non-native amino acids to use in synthesis of new candidate reporters.

Figure 6:

An example figure showing the progress curves from recombinant enzyme assays with three peptide reporter concentrations. For each initial reporter concentration (12, 15, and 18 μM), the concentration of phosphorylated reporter is plotted on the y-axis with the assay incubation time on the x-axis. The data from the three progress curves will be fit to the time-dependent Michaelis-Menten equation using an analytical approximation of the Lambert function.

Figure 7:

Reaction scheme for synthesis of the photocleavable residue Fmoc-Thr(DMNB)-OH (N-(9-Fluorenylmethoxycarbonyl)-O-(4,5-dimethoxy-2-nitrobenzyl)-L-threonine.)

Table 1.

Mass of macromolecules in a single cell with an assumed total mass of 1 ng.

Macromolecule	Mass/cell
DNA	6 pg
Total RNA	10 pg
mRNA	0.1 pg
proteins	200 pg
Lipids	50 pg
Carbohydrates	20 pg

Table 2.

Common salts and additives for electrophoretic buffers.

Buffer Salts	Additives
Borate	1-propanol
CAPS	2-propanol
Glycine	Acetonitrile
HEPES	Cetrimonium bromide (CTAB)
MOPS	Cyclodextrin
Sodium Bicarbonate	Dextrin
Sodium Citrate	N,N-dimethylacrylamide
Sodium Phosphate	Dodecyltrimethyl ammonium bromide (DTAB)
Sodium Tetraborate	Methanol
Tricine	Sodium deoxycholate (SDC)
Tris	Sodium dodecylsulfate (SDS)
	Tween 20

Table 3.

Reporter validation assays.

Validation Method	Description	Assay Type
Enzyme inhibitors	Panel of inhibitors specific to enzyme/enzyme family	Lysates; Single cells (if inhibitors are membrane permeant)
Enzyme activators	Panel of activators specific to enzyme/enzyme family	Lysates; Single cells (if inhibitors are membrane permeant)
Pull down	Use antibodies specific to enzyme to remove enzyme from lysate mixture	Lysates
siRNA knockdown	Prevents translation of specific proteins (enzymes)	Lysates; Single cells
Mutant cell lines	Cell lines with over- or underexpression of enzyme	Lysates; Single cells
Genetically engineered cell lines	Engineer cell lines with desired properties	Lysates; Single cells