

**EPIDERMAL GROWTH FACTOR RECEPTOR TYROSINE KINASE ASSAYS IN  
SINGLE INTACT CELLS USING CAPILLARY ELECTROPHORESIS**

Abigail H. Turner

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry in the College of Arts and Sciences.

Chapel Hill  
2015

Approved by:

Nancy L. Allbritton

R. Mark Wightman

Matthew R. Lockett

David S. Lawrence

Channing J. Der

© 2015  
Abigail H. Turner  
ALL RIGHTS RESERVED

## ABSTRACT

Abigail H. Turner: Intracellular Epidermal Growth Factor Receptor Tyrosine Kinase Assays at the Single Cell Level Using Capillary Electrophoresis  
(Under the direction of Nancy L. Allbritton)

Quantification of abnormal Epidermal Growth Factor Receptor (EGFR) tyrosine kinase activity is critical to the clinical success of targeted inhibitors used to treat EGFR-dependent cancers. Current selection criteria based on the presence or absence of activating mutations in the EGFR gene need to be supplemented with complementary assays which quantify EGFR biochemical activity in intact living cells. While *in cellulo* EGFR activity can be measured to great success in tissue cultured cells using genetically encoded reporters, such techniques are incompatible with small, highly heterogeneous tumor samples. A chemical cytometry approach to direct EGFR activity measurements is presented which consists of an *in cellulo* peptide reporter-based kinase assay coupled to ultrasensitive microelectrophoretic analysis. The major challenge to such a measurement is biological assay interference from off-target enzymes. A novel approach to minimizing selectivity-based assay interference and its application in single intact cells is presented.

*For my late grandmother, Vivian Turner, and for my parents, Chris and Mary Turner. You believed that I could do anything, so I believed it myself. Thank you.*

## ACKNOWLEDGEMENTS

I owe thanks to so many friends and colleagues who have supported this work. Dr. Nancy Allbritton has been a wonderful scientific mentor and over the course of this work helped me to learn the value of perseverance and creative thinking. Dr. Angie Proctor has been a constant source of support and advice from my first day in the lab. Her personal and professional ethic set the tone for our team, and her legendary lab notebook keeping reaches a level of excellence to which I can only aspire. My fellow CE team members present and past, Dr. Jazz Dickinson, Dr. Ryan Phillips, Dr. Michelle Kovarik, Emilie Mainz, Greg Woss, Taylor Harris, and Dr. Imola Zigoneanu, have been an invaluable source of ideas and inspiration. My undergraduate trainees, Mike Lebhar and Josh Mu, have contributed so much to these chapters with their hard work and dedication. Dr. Ryan Phillips, Pavak Shah, and Pete Attayek have generously contributed many hours to craft MATLAB codes to support this project. Dr. David Lawrence has been an endless source of creative solutions to scientific quandaries. Dr. Qunzhao Wang, Dr. Liz Cline, and Dr. Kaiulani Houston have patiently guided this Analytical chemist through the sometimes murky waters of peptide synthesis. Nicole Baker and Marissa Cann's assistance in Western blotting has been invaluable. The members past and present

of the Allbritton group, particularly Angie Proctor, Colleen Phillips, Ryan Phillips, Shan Yang, Emilie Mainz, Pavak Shah, Jazz Dickinson, and Greg Woss, have made every day in lab a good day. Finally, I never would have made it through my first semester of graduate school without the support and friendship of Shelly Beard, Kaitie Fague, and Nick Dobes. Thank you to everyone, not only for your help, but also for your friendship.

## TABLE OF CONTENTS

LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiii
LIST OF ABBREVIATIONS AND SYMBOLS.....	xv
CHAPTER 1: IMPORTANCE OF EGFR KINASE ACTIVITY MEASUREMENTS AT THE SINGLE CELL LEVEL .....	1
1. Introduction .....	1
1.1 Enzymes and Their Role in Oncogenic Transformation.....	1
1.2 Oncogenic EGFR Activity .....	2
1.3 Targeting EGFR in the Clinic.....	4
1.4 Overview of Kinase Activity Assays .....	6
1.4.1 Basic Requirement for Kinase Activity Measurements.....	6
1.4.2 Source of Kinase and Sample Matrix.....	7
1.4.3 Kinase Substrate Choice .....	8
1.4.4 Basic Kinetic Theory.....	10
1.5 Assay Platforms for Direct Measurement of EGFR Kinase Activity .....	16
1.5.1 <i>In Vitro</i> Radiometric Assays.....	17
1.5.2 <i>In Vitro</i> Fluorescence/Luminescence-Based Assays.....	19

1.5.3 <i>In Vitro</i> Mobility Shift Assay .....	22
1.5.4 <i>In Cellulo</i> Assay Platforms for Cell Lysates.....	23
1.5.5 <i>In Cellulo</i> Assay Platforms for Fixed Cells .....	26
1.5.3 <i>In Cellulo</i> Assay Platforms for Live Cells .....	28
1.6 Chemical Cytometry with Single Cell Capillary Electrophoresis .....	30
1.6.1 Definition and Applications of Chemical Cytometry .....	30
1.6.2. Theoretical Overview of Capillary Electrophoresis <sup>76,77</sup> .....	31
1.6.3 Single Cell Capillary Electrophoresis Assay of EGFR Kinase Activity .....	38
REFERENCES .....	43
 CHAPTER 2: CONFORMATIONAL CONSTRAINT AS A SOLUTION TO PROTEIN TYROSINE PHOSPHATASE INTERFERENCE IN INTRACELLULAR PEPTIDE-BASED KINASE ASSAYS.....	
2.1 Background .....	59
2.2 Materials and Methods.....	62
2.2.1 Peptide Synthesis and Purification .....	62
2.2.2 Cell Culture .....	64
2.2.3 <i>In Vitro</i> EGFR Assays .....	64
2.2.4 Phosphopeptide Preparation .....	65
2.2.5 <i>In Vitro</i> PTP Assays.....	65
2.2.6 Cell Lysate Preparation.....	65
2.2.7 Lysate Phosphorylation Assays.....	66



2.2.8 Lysate Dephosphorylation Assays.....	66
2.2.9 Lysate Peptidolysis Assays .....	67
2.2.10 Capillary Electrophoresis .....	67
2.2.11 <i>In Vitro</i> Data Analysis .....	68
2.2.12 Pinocytic Loading for Single Cell EGFR Activity Assays.....	69
2.2.12 Single Cell Analysis with Microelectrophoresis .....	70
2.3 Results and Discussion .....	70
2.3.1 Mitigation of Phosphoreporter Dephosphorylation <i>in vitro</i> via Conformational Constraint .....	70
2.3.2 Phosphoreporter Lifetime in Epithelial Tumor Lysates .....	73
2.3.3 Kinetics of Phosphoreporter Dephosphorylation.....	74
2.3.4 Increased Resistance to Peptidolysis in Cell Lysates.....	75
2.3.5 Quantifying Response to EGFR Inhibition .....	76
2.3.6 Single Cell Analysis.....	77
2.4 Conclusions .....	78
REFERENCES .....	80
 CHAPTER 3: TOWARD HIGH THROUGHPUT EGFR ACTIVITY MEASUREMENTS IN SINGLE TUMOR CELLS .....	
3.1 Introduction .....	90
3.2 Methods .....	92
3.2.1 Materials.....	92

3.2.2 2D and 3D A431 Cell Culture and Treatment .....	93
3.2.3 Pinocytic Loading of A431 Cells.....	94
3.2.4 Bulk EGFR Assays in Intact Cells.....	95
3.2.5 Microscopy .....	95
3.2.6 Western Blotting .....	96
3.2.7 Capillary Electrophoresis .....	96
3.2.8 Single Cell Analysis.....	97
3.2.9 Data Analysis .....	97
3.3 Results and Discussion .....	97
3.3.1 Proposed EGFR Assay in 3D Cultured A431 Cells.....	97
3.3.2 Pinocytic Loading of A431 Cells with Htc-tide.....	98
3.3.3 3D Culture of A431 Cells .....	99
3.3.4 Differences in EGFR Activation in 2D and 3D Cultured A431 Cells .....	100
3.3.5 Separation Optimization for 4 cm Effective Length CE-LIF System.....	102
3.3.6 EGFR Activity in Single A431 Cells .....	103
3.4 Conclusions and Future Directions .....	104
REFERENCES .....	105
 CHAPTER 4: A Higher Throughput Site-Directed Combinatorial Approach to Kinase Reporter Optimization.....	 114
4.1 Limitations of Kinase Reporter Selectivity .....	114
4.1.1 Synthetic Approaches to Spatiotemporal Control .....	114

4.1.2 Modification with Unnatural Amino Acids .....	116
4.1.3 Site Directed Combinatorial Method for Substrate Optimization.....	118
4.2 Methods .....	118
4.2.1 General Synthetic Methods .....	119
4.2.2 Cystamine Resin Preparation and Characterization .....	121
4.2.3 Crude Peptide Purification and Characterization .....	122
4.2.5 Library Synthesis and Characterization .....	123
4.2.6 High Yield Synthesis of Asp-Containing Scaffold and Library Hit Compounds .....	126
4.2.7 <i>In Vitro</i> Kinase Assays.....	127
4.2.8 Capillary Electrophoresis .....	129
4.2.9 MALDI-TOF MS .....	129
4.2.10 Homogeneous Luminescence Assays.....	130
4.2.11 Library Screen Data Analysis.....	131
4.3 Results and Discussion .....	133
4.3.1 Selectivity of the Starting EGFR Peptide .....	133
4.3.2 Comparison of Potential Library Screening Platforms .....	133
4.3.2 Luminescence Assay Validation with Synthetic Standards .....	136
4.3.3 Effect of N-Terminal Label on Scaffold Phosphorylation by EGFR.....	137
4.3.4 MALDI-TOF for Phosphorylation Detection and Effect of Ionization Bias .....	138

4.3.5 Lysine Scan to Determine Library Site .....	139
4.3.6 Library Synthesis and Characterization .....	140
4.3.7 Library Screening and Data Analysis .....	141
4.3.8 Library Hit Re-Synthesis and Validation .....	142
4.4 Conclusions and Future Directions .....	145
REFERENCES .....	148
APPENDIX A. PRACTICAL SOLID PHASE PEPTIDE SYNTHESIS .....	168
APPENDIX B. MATLAB CODE .....	184
APPENDIX C. CARBOXYLIC ACID LIBRARY .....	187

## LIST OF TABLES

Table 1. Common Components of Kinase Assays.....	50
Table 2. Assay Platforms for Measuring EGFR Tyrosine Kinase Activity .....	51
Table 3. <i>In vitro</i> Kinetic Parameters.....	83
Table 4 Protein Tyrosine Kinase Consensus Sequences .....	152

## LIST OF FIGURES

Figure 1. Overview of EGFR biology .....	55
Figure 2. Kinetic curve based on the Michaelis-Menten model of enzyme kinetics ..	56
Figure 3. Overview of capillary electrophoresis.....	57
Figure 4. Overview of single cell EGFR assay .....	58
Figure 5. Differential <i>in vitro</i> dephosphorylation of EGFR reporters.....	84
Figure 6. In cellulo dephosphorylation of phosphorylated EGFR reporters .....	85
Figure 7. Vacuum electrostatic potential maps of EGFR kinase domain.....	86
Figure 8. Assessment of reporter proteolysis in cell lysates .....	87
Figure 9. EGFR reporter phosphorylation in A431 cell lysates under perturbation ...	88
Figure 10. Single cell analysis using Tyr-tide and Htc-tide .....	89
Figure 11. Proposed workflow for single cell analysis of EGFR activity in response to lapatinib treatment in 3D cell cultures and solid tumors.....	108
Figure 12. Pinocytic loading of Htc-tide into A431 cells .....	109
Figure 13. Schematic of 3D culture technique.....	110
Figure 14. Differences in EGFR activity in 2D and 3D A431 cultures.....	111
Figure 15. CE separation optimization for high throughput serial analysis.....	112
Figure 16. Single cell analysis on the high throughput system on two separate days.....	113
Figure 17. Library screening workflow.....	153

Figure 18. Summary of library screening platforms. ....	154
Figure 19. Time courses of 6FAM-TyrSub-9 phosphorylation by a panel of tyrosine kinases.....	155
Figure 20. Comparison of phosphorylated Biotin-TyrSub-9 detection by LANCE TR-FRET and AlphaScreen under optimized conditions.....	156
Figure 21. AlphaScreen detection of <i>in vitro</i> phosphorylation of Biotin-TyrSub-9 by EGFR.....	157
Figure 22. MALDI-TOF MS analysis of <i>in vitro</i> EGFR assays with Biotin-TyrSub-9 and 6FAM-TyrSub-9.....	158
Figure 23. Effect of ionization efficiency differential on measurement bias in MALDI-TOF MS .....	159
Figure 24. Site scan to determine optimal library optimization site within the starting EGFR peptide.....	160
Figure 25. Results of the MALDI-TOF screen of the EGFR peptide library.....	161
Figure 26. RP-HPLC-MS analysis of selected library compounds.....	162
Figure 27. Library compound screening by MALDI-TOF MS.. .....	163
Figure 28. Structure of the Syk screen hit 4-10F. ....	164
Figure 29. Mechanism of aspartimide formation and resultant racemization.....	165
Figure 30. Phosphorylation of 4-10F by recombinant Syk. ....	166
Figure 31. Proposed workflow for modified library screen with SAMDI.....	167
Figure 32. Solid Phase Peptide Synthesis (SPPS) workflow .....	183

## LIST OF ABBREVIATIONS AND SYMBOLS

(Ac)	Acetyl
-CONH <sub>2</sub>	Amidation
(aq)	Aqueous
(g)	Gaseous
[X]	Concentration of X
[X] <sub>0</sub>	Total Concentration of X
°C	Degrees C
3D	Three-dimensional
6FAM	6-Carboxyfluorescein
a	Hydrodynamic Radius
a.a.	Amino Acid
Ab	Antibody
Abl	Abelson Tyrosine Kinase
ACN	Acetonitrile
ACT	Activation Solution
ATP	Adenosine triphosphate
BGE	Background Electrolyte
BSA	Bovine Serum Albumen



cDNA	Complementary DNA
CE-LIF	Capillary Electrophoresis with Laser-Induced Fluorescence
CE-MS	Capillary Electrophoresis-Mass Spectrometry
CHEF	Chelation Enhanced Fluorescence
cIEF	Capillary Isoelectric Focusing
CMC	Critical Micellar Concentration
COOH	Carboxylic Acid
CRC	Colorectal Cancer
CZE	Capillary Zone Electrophoresis
D	Diffusion Coefficient
Dap	Diaminopropionic Acid
DCM	Dichloromethane
DEP	Deprotection Solution
DIC	Diisopropyl carbodiimide
DIEA	Diisopropylethylamine
DMEM	Dulbecco's Modified Eagles Medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol

DTT	Dithiothreitol
E	Electric Field
E	Enzyme
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
ELISA	Enzyme-Linked Immunosorbent Assay
EOF	Electroosmotic Flow
EpCAM	Epithelial Cell Adhesion Molecule
eq	Equivalents
ESI	Electrospray Ionization
ESI	Electrospray Ionization
<i>f</i>	Drag Forces
FACS	Fluorescence-Activated Cell Sorting
FAM	Carboxyfluorescein
FBS	Fetal Bovine Serum
FLIM	Fluorescence Lifetime Imaging Microscopy
Fmoc	Fluorenylmethyloxycarbonyl
FRET	Förster Resonance Energy Transfer
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo

	[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl)uronium hexafluorophosphate
HCl	Hydrochloric acid
HCTU	2-(6-Chloro-1 <i>H</i> -benzotriazole-1-yl) - 1,1,3,3-tetramethylaminium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER	Human Epidermal Growth Factor Receptor
HOBt	Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRP	Horseradish Peroxidase
Htc	7-( <i>S</i> )-hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid
HTRF	Homogeneous Time-Resolved Fluorescence
i.d.	Internal Diameter
IHC	Immunohistochemistry
IPA	Isopropyl alcohol
KAYAK	Kinase Activity Assay for Kinome Profiling
$k_{cat}$	Turnover Number
KD	Kinase Domain
$K_M$	Michaelis-Menten Constant
L	Capillary Length

LC-MS	Liquid Chromatography-Mass Spectrometry
MALDI	Matrix Assisted Laser Desorption Ionization
MEKC	Micellar Electrokinetic Chromatography
MeOH	Methanol
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass Spectrometry
MW	Molecular Weight
N	Number of Theoretical Plates
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NMM	N-Methylmorpholine
NMP	N-methylpyrrolidone
NMR	Nuclear Magnetic Resonance
NSCLC	Non-Small Cell Lung Cancer
OMpe	O-Methylphenylester
OtBu	O-tert-Butyl
P	Product
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Poly-Ethylene Glycol

PEG	Polyethylene glycol
PTK	Protein Tyrosine Kinase
PTP	Protein Tyrosine Phosphatase
PTP1B	Protein Tyrosine Phosphatase 1B
pVO <sub>4</sub>	Pervanadate
q	Charge
rcf	Relative Centrifugal Units
RP	Reversed Phase
R <sub>s</sub>	Resolution
RTK	Receptor Tyrosine Kinase
S	Substrate
S/N	Signal-to-Noise Ratio
SAM	Self-Assembled Monolayer
SH2	Src Homology 2
SPPS	Solid Phase Peptide Synthesis
sQSSA	Standard Quasi-Steady State Assumption
Src	Src Tyrosine Kinase
SYK	Spleen Tyrosine Kinase
t	Time
TCPTP	T-Cell Phosphatase

TFA	Trifluoroacetic acid
TIS	Triisopropylsilane
TKI	Tyrosine Kinase Inhibitor
TOF	Time-of-Flight
TSTU	<i>O</i> -( <i>N</i> -Succinimidyl)- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
UV	Ultraviolet
V	Applied Voltage
<i>v</i>	Rate
$V_{\max}$	Maximal Rate (Michaelis-Menten Model)
W(x)	Lambert W Function
$\alpha$ -C	$\alpha$ -Cynaohydroxycinnamic Acid
$\epsilon$	Permittivity
$\zeta$	Zeta Potential
$\eta$	Viscosity
$\mu_{\text{eff}}$	Effective Mobility
$\mu_{\text{eo}}$	Electroosmotic Mobility
$\mu_{\text{ep}}$	Electrophoretic Mobility
$\Sigma$	Standard Deviation

## CHAPTER 1: IMPORTANCE OF EGFR KINASE ACTIVITY MEASUREMENTS AT THE SINGLE CELL LEVEL

### 1. Introduction

#### 1.1 Enzymes and Their Role in Oncogenic Transformation

Enzymes are protein catalysts that enable the chemical reactions necessary for cellular function. The chemical reactions catalyzed vary widely, including oxidation/reductions, structural rearrangements, and hydrolyses. The basic model of an enzymatic reaction is outline in Equation (1).



In this model, the enzyme E recognizes the substrate S and catalyzes its conversion to the product P. The activity of an enzyme, defined as moles of product generated per unit time, is a complex function of protein conformation, protein-protein interactions, substrate and product concentrations, and the presence of co-factors, activators, and inhibitors.<sup>1</sup> In healthy cells, enzyme activity levels are dynamic and tightly controlled, poised to respond to a stimulus and subsequently to return to basal levels.<sup>2,3</sup> Disease phenotypes can often be traced to one or more abnormal enzyme activities. Perhaps the best understood example of this is the hyper-

-activation of protein kinases in cancer.<sup>4</sup>

Protein kinases effect the phosphorylation of protein or peptide substrates by catalyzing the transfer of the  $\gamma$ -phosphoryl of ATP to the hydroxyl of functional serine, threonine, or tyrosine residues. Protein phosphorylation functions to modulate enzyme activity and is a primary mechanism of information propagation in intracellular signaling networks.<sup>5,6</sup> Abnormally high activity of some protein kinases can lead to dysregulated signaling in key networks governing such cellular functions as growth, proliferation, and survival, ultimately leading to malignant cellular transformation and disease.<sup>4</sup>

## **1.2 Oncogenic EGFR Activity**

The Epidermal Growth Factor Receptor (EGFR) is a receptor tyrosine kinase that promotes growth and proliferation in response to growth factor binding.<sup>7</sup> EGFR is comprised of three domains, represented in Figure 1A: an extracellular ligand-binding domain, a transmembrane domain, and an intracellular tyrosine kinase domain with an unstructured C-terminal tail. EGFR is one of a family of four structurally related transmembrane receptor tyrosine kinases (RTKs) that make up the Human Epidermal Growth Factor Receptor (HER) Family. These kinases are highly homologous in the kinase domain,<sup>8</sup> but differ significantly in their ligand selectivity.<sup>9</sup> Ligand binding to EGFR, HER3, or HER4 induces a conformational change in the extracellular domain and promotes receptor dimerization with one of



the other HER family members.<sup>10</sup> Homo- or hetero-dimerized receptors then perform reciprocal trans-phosphorylation of tyrosine residues in the intracellular tail region of the kinase domain.<sup>11</sup> These phosphotyrosine residues serve as docking sites for downstream signaling proteins, which bind at these sites through their SH2 domains and are recruited to the kinase domain for phosphorylation and activity modulation.<sup>7</sup>

The identity of the activating ligand and the composition of the receptor dimer determines which tyrosine residues within the tail are phosphorylated and, therefore, which intracellular signaling pathways are activated.<sup>5</sup> EGFR-containing dimers initiate four key signaling pathways controlling growth and proliferation: PI3K/Akt/mTor, Ras/MEK/ERK, JAK/STAT, and PLC (Figure 1B).<sup>12</sup> In healthy cells, EGFR signaling is kept in check by protein phosphatases, which catalyze dephosphorylation, and by the cell's protein metabolism machinery, which recycle proteins into their constituent amino acids.<sup>5</sup> Uncontrolled EGFR signaling can lead to inappropriate hyperactivation of EGFR-dependent signaling networks, resulting in uncontrolled proliferation and oncogenesis. Aberrant EGFR signaling has been documented in a range of epithelium-derived cancers including non-small cell lung, breast, colorectal, pancreatic, head-and-neck, and some gliomas.<sup>9</sup> Oncogenic EGFR signaling typically results either from overexpression of the EGFR protein or from one or more activating mutations in the EGFR gene, leading to constitutive, growth

factor-independent activation of the EGFR kinase domain and unregulated signaling in downstream networks.

### **1.3 Targeting EGFR in the Clinic**

In a healthy cell, loss of activity in one signaling network is compensated by redundant parallel pathways to maintain homeostasis.<sup>13</sup> Tumor cells, in a phenomenon termed oncogene addiction, rely so heavily on one aberrant activity that inhibition of that activity results in cell death.<sup>14,15</sup> This constitutes an important therapeutic opportunity in tumors where the oncogene can be identified. For EGFR-reliant tumors, a set of targeted EGFR inhibitors is routinely employed in the clinic. These drugs fall into two classes: monoclonal antibodies and small molecule tyrosine kinase inhibitors.<sup>16</sup> Anti-EGFR monoclonal antibodies bind the EGFR extracellular domain to prevent receptor dimerization and activation, whereas the small molecule TKIs dock in the kinase domain (KD) ATP binding site to directly inhibit EGFR kinase activity. Theoretically, these targeted inhibitors should selectively induce apoptosis in EGFR-dependent tumor cells, leaving normal cells (with normal signaling redundancies) unharmed and resulting in fewer deleterious side effects compared with cytotoxic chemotherapy. In reality, while they are accompanied by fewer and less severe side effects compared with untargeted cytotoxic chemotherapy, these drugs fail to live up to their full potential.<sup>17</sup> A review of studies of a small-molecule EGFR inhibitor found that 32% of patients with

advanced EGFR mutation-positive non-small cell lung cancer failed to respond to treatment.<sup>18</sup>

The variable clinical success of EGFR inhibitors can in part be attributed to inadequate selection criteria when determining patient eligibility for anti-EGFR therapy. The most recent guidelines for clinical EGFR testing recommend analysis of primary tissue with PCR amplification followed by mutation specific sequencing (e.g. Sanger, capillary electrophoresis, real-time PCR) to test for the presence of activating mutations in the EGFR gene.<sup>18</sup> Despite a positive correlation between detection of EGFR mutations and response to EGFR TKIs, a subpopulation of patients who are EGFR-positive fail to benefit from targeted therapy. Additionally, a small number of patients receive a survival benefit from anti-EGFR therapy despite testing negative for EGFR mutations.<sup>18</sup>

A major flaw of the current testing methodology is the use of genetic information as a proxy for actual enzyme activity. Abnormal EGFR signaling driven by hyperactivation of the kinase is not directly related to the presence of an activating mutation; rather, many layers of regulation exist between genetic mutation and observed activity including transcriptional and translational mechanisms and functional post-translational processing of the protein. Therefore, the simple presence of an activating EGFR mutation in a tumor genome is a potentially inaccurate gauge of abnormal kinase activity and, consequently, an

inadequate predictor of response to anti-EGFR therapy. Genetic analyses need to be supplemented with complementary assays that directly quantify EGFR kinase activity in patient samples. Additionally, the determinants of kinase activity can differ in each cell, resulting in highly heterogeneous signaling even in apparently homogeneous samples. Therefore, activity measurements need to be adaptable to the single cell level so that rare drug- sensitive or drug-resistant populations are not obscured.

#### **1.4 Overview of Kinase Activity Assays**

In this section, the basic theory of kinase activity measurements is explored. Practical considerations for experimental design and implementation are discussed, followed by a brief treatment of the theoretical foundations underpinning current modeling and interpretation of kinase activity data.

##### **1.4.1 Basic Requirement for Kinase Activity Measurements**

Direct measurements of kinase activity, or kinase assays, can be accomplished by an impressive range of technologies developed over the course of the last century. Fundamentally, these technologies measure either the consumption of substrate (dS) or the evolution of phosphorylated substrate (dP), typically as a function of time ( $-dS/dt = dP/dt$ ). Kinase assay design is a complex bioanalytical goal, the pursuit of which has resulted in the development of an enormous variety of assay platforms. The components of a kinase assay depend on the sample matrix and assay platform;

a brief summary of common assay platform components is presented in Table 1. The minimum requirements for any kinase assay are as follows:

- Kinase
- Phosphorylatable substrate
- ATP
- Divalent metal cation ( $Mg^{2+}$  or  $Mn^{2+}$ )
- Buffer/stabilizers
- Detection method

The source of the kinase to be assayed will largely determine the assay sample matrix and the available detection platforms. Many assay platforms are not compatible with physiological samples (cell lysates, intact cells) because endogenous species in the sample matrix interfere either with the kinase reaction itself or with the detection method. Therefore assay platforms are generally divided into three classes: *in vitro* assays with purified kinase, *in cellulo* assays in cellular lysates, and *in cellulo* assays in living cells.

#### **1.4.2 Source of Kinase and Sample Matrix**

In choosing a source of kinase and assay platform, the biological relevance of the experiment must be balanced against the need for experimental transparency, as well as practical considerations of difficulty, cost, and available instrumentation.

Clearly, if true biological kinase activity needs to be measured, an *in cellulo* or *in vivo* model is the best option. However, *in cellulo* assays are highly complex, not only in their implementation, but also in interpretation of results. Additionally, control of

experimental variables is limited, thereby limiting the precision of the hypotheses that may be explored. Often, *in vitro* assays with purified kinase prove to be the most practicable choice for routine biochemical analyses such as kinetic experiments and inhibitor screens because they afford the researcher full control over sample matrix, assay conditions, and assay variables.

### 1.4.3 Kinase Substrate Choice

The kinase substrate may be chosen from a range of options, depending upon the particular assay format and other experimental requirements and constraints. These include: 1) full-length or partial protein, 2) random amino acid polymers,<sup>19</sup> and 3) short tryptic or synthetic peptides.<sup>20-22</sup> Early kinase assays employed full-length proteins such as histone as the kinase substrate<sup>20</sup>; where the protein used is an endogenous substrate for the kinase of interest, native specificity can be approximated. Specificity is particularly an important consideration in complex sample matrices such as cellular lysates where the substrate is subject to interference from off-target enzymatic activities such as phosphorylation by off-target kinases, dephosphorylation by protein phosphatases, and metabolism by proteolytic enzymes. Where specificity is not a consideration or is undesirable, random amino acid polymers can be an inexpensive and robust alternative. These are large macromolecular polymers containing a fixed percentage of phosphorylatable residues in a simple background of nonreactive amino acids. The most popular

example of this is poly-Glu(Tyr) (4:1), a random polymer containing one mole of tyrosine for every 4 moles of glutamic acid. Such polymers show very broad specificity profiles due to a lack of refined structural characteristics. However, prior to the availability of inexpensive peptide substrates, poly-Glu(Tyr) and similar polymers were commonly employed when protein substrates were cost or time-prohibitive.

Macromolecular substrates have generally fallen out of favor as technology has progressed, particularly for *in vitro* work. First, advances in peptide synthesis, specifically the invention of solid phase peptide synthesis by Merrifield in the 1960s,<sup>23,24</sup> have made custom peptide substrates widely commercially available and affordable. Second, whereas the macromolecular kinase substrates often contain multiple phosphorylation sites which make kinetic analyses difficult,<sup>25</sup> peptide substrates have user-defined sequences with any desired number of phosphoacceptors. Peptide substrates are usually quite small, containing 15 or fewer amino acids. This size regime is advantageous because it results in excellent solubility and very high stability over the course of the assay, improving experimental robustness. Perhaps the most important advantage of peptide substrates over other classes of substrates is the ease with which peptides can be customized to suit both the target kinase and the experimental technique. For example, labeling of synthetic peptides with fluorophores or other moieties to

enable detection has become a trivial task for which most labs are sufficiently equipped. Importantly, the wide availability of fluorophore-labeled peptide substrates has precipitated the development of a host of fluorescence and luminescence-based assay platforms and greatly facilitated massively high throughput screening assays that have become foundational in academic and industrial drug discovery labs.

#### 1.4.4 Basic Kinetic Theory

A theoretical understanding of kinase activity is grounded in measurements of reaction rate. Consequently, most kinase assays incorporate some temporal component, whether by continuously monitoring the kinase reaction over time, or by taking measurements at multiple fixed time points. A robust kinetic model of a kinase's activity allows perturbations to that activity to be accurately assessed. This is particularly important when evaluating potential therapeutic interventions such as kinase agonists and inhibitors. The Michaelis-Menten model of enzyme kinetics<sup>26</sup> is often invoked to model kinase reactions. Dating to the early twentieth century, the Michaelis-Menten model describes the kinetics of an enzymatic reaction of the form:



Here the enzyme E and substrate S form a reversible complex ES, followed by the conversion of substrate to product P and dissolution of the ES complex. The



Michaelis-Menten model assumes that the enzymatic reaction conforms to the standard quasi-steady state assumption (sQSSA), which requires that the concentration of the enzyme substrate complex is essentially constant in time.<sup>27</sup> In the traditional treatment, this is interpreted as a requirement that the consumption of substrate is negligible ( $\leq 5\%$ ) over the course of the experiment. An additional requirement for this treatment is that the initial reaction rate must increase linearly with enzyme concentration. Invoking the sQSSA with regard to (2), the following equation describing the reaction rate  $v$  can be derived:

$$v = \frac{dP}{dt} = \frac{k_{cat}[E]_0[S]}{K_M + [S]} = \frac{V_{max}[S]}{K_M + [S]} \quad (3)$$

where  $[E]_0$  is the total enzyme concentration. Two derived parameters are used to describe reactions conforming to the Michaelis-Menten model. The  $V_{max}$  is the maximal rate of reaction, equivalent to the  $k_{cat}$  (2) multiplied by the total enzyme concentration and corresponding to the maximal rate limited by the number of available enzyme active sites. The  $K_M$  is a ratio of kinetic constants that is often treated analogously to a binding equilibrium constant, although it is not in fact an equilibrium constant unless the system being modeled is at equilibrium under the experimental conditions used. It is more correct to interpret the  $K_M$  as the substrate concentration at which the enzyme's catalytic power is most efficiently utilized. It is therefore a good approximation of the endogenous intracellular substrate concentration. Generally speaking, a substrate of a given enzyme is considered

“good” if the  $K_M$  describing their interaction is in the low micromolar or smaller. The parameters  $K_M$  and  $k_{cat}$  are useful for comparing enzyme and substrate behavior under varied conditions. The ratio  $k_{cat}/K_M$  is commonly invoked to directly compare multiple substrates of one enzyme or the efficiency of phosphorylation of one substrate by multiple enzymes. For this reason, the ratio is often referred to as the “specificity constant”. A large specificity constant implies that the substrate is phosphorylated rapidly and with a good degree of selectivity. When comparing one substrate and two enzymes, the substrate is said to be selective for one enzyme over another when the specificity constant for that enzyme is significantly larger.

Perturbations to an enzymatic reaction, e.g. addition of agonists or inhibitors, will affect the measured kinetic parameters. This allows the experimenter to derive mechanistic information about the perturbing species. This is particularly useful when evaluating kinase inhibitors because the mechanism of inhibition will affect the inhibitor’s behavior and efficacy *in vivo*. A competitive inhibitor, one that competes with endogenous substrates for binding to the enzyme active site, will diminish the measured  $V_{max}$  and increase the measured  $K_M$ . A non-competitive inhibitor, on the other hand, will reduce the measured  $V_{max}$ , but will not affect the measured  $K_M$ . An uncompetitive inhibitor, one that binds and inactivates the ES complex, will decrease both the measured  $K_M$  and  $V_{max}$ .<sup>28</sup> Clearly, robust kinetic

assays interpreted through the appropriate model can yield a wealth of information about enzymatic behavior and response to perturbation.

The classical approach to estimating Michaelian kinetic parameters requires measuring initial reaction rate at a series of substrate concentrations to generate a Michaelis-Menten kinetic curve (Figure 2). The experimental points are then fit using nonlinear regression to (3), using  $K_M$  and  $V_{\max}$  as fit parameters. The  $k_{\text{cat}}$  is derived from  $V_{\max}$  and the total concentration of enzyme. In order to satisfy the sQSSA with this approach, the experiment must be engineered such that the total substrate present is approximately equal to the amount of free substrate available:

$$[S]_0 = [S] + [ES] + [P] \approx [S] \quad (4)$$

Consequently, the reaction should only be followed to  $\leq 5\%$  completion in order to ensure that the rate measured is truly an initial rate. The initial rates method is the gold standard because it is mathematically straightforward and experimentally simple to implement. Additionally, because product formation is only followed over a very small portion of the reaction course, or “progress curve”, confounding experimental problems such as reagent instability and product inhibition are avoided.<sup>29</sup> However, the initial rate method is costly in terms of time, labor, and reagents, and it cannot exploit the information contained in the discarded 95% of the progress curve. Furthermore, since this approach essentially involves taking the

derivative of progress curve data, it is fundamentally prone to magnifying experimental error.

Efforts to derive  $K_M$  and  $k_{cat}$  by fitting complete progress curves date back to the seminal report by Michaelis and Menten in 1913, in which the authors integrated equation (3) to describe time-dependent formation of product. The integrated or time-dependent Michaelis-Menten (5) comprises a linear term describing product formation at high  $[S]$  (e.g.  $\gg K_M$ ), and a logarithmic term which models product formation at low  $[S]$  ( $\ll K_M$ ).

$$K_M \cdot \ln\left(\frac{[S]_0}{[S]_0 - [P](t)}\right) + [P](t) = V_{max} \cdot t \quad (5)$$

This equation is difficult to fit accurately because the independent variable cannot be isolated from the dependent variable. Attempts to derive  $K_M$  and  $V_{max}$  from linearized plots of equation (5) have been made but are error-prone and unpopular.

An alternative solution to the integration of (5) is given in (6).<sup>30,31</sup>

$$[P](t) = [S]_0 - K_M \cdot W\left\{\frac{[S]_0}{K_M} \exp\left(\frac{[S]_0 - V_{max} \cdot t}{K_M}\right)\right\} \quad (6)$$

where  $W$  is the Lambert function, such that:

$$\text{For } y + \ln(y) = \ln(x) \quad (7)$$

$$y = W(x)$$

Use of the Lambert function allows the dependent variable  $P(t)$  to be isolated from the independent variable, thus reducing error in modeling. Few non-linear

regression programs are equipped to solve the Lambert function, but several useful analytical approximations have been published which are simple to implement in common software such as OriginLab or GraphPad.<sup>30-35</sup> Because progress curve analysis takes into account the entire reaction course, an individual progress curve experiment contains >90% more information than a corresponding initial rates experiment. Theoretically, this means that fewer progress curve experiments should be needed, compared with initial rates experiments, to obtain the same information. The shape of a kinase reaction curve describes not only the phosphorylation event, but also contributions from activators, inhibitors, interferents, competing enzymatic reactions, substrate inhibition, reagent instability, and any other phenomena observable by the assay technique employed. This leads to fewer overall experiments performed, less reagent consumption, and less time and labor expenditure. The barrier to widespread implementation of progress curve analysis has undoubtedly been the need for complex mathematical models to describe the shape of the observed reaction curves, accompanied by the requirement for considerable computational resources. However, technological advances in the last decade have made essentially negligible the computational burden for progress curve analysis. Concomitantly, many specialized software packages have been made available, with varying degrees of user-friendliness, to facilitate complex model development for progress curve analysis.<sup>29,36-40</sup> Nevertheless, progress curve analysis

takes time to implement successfully—time which many laboratories are not willing to expend in order to “fix what isn’t broken”. Progress curve analysis has remained largely a niche technique, applied in situations where the benefits of the technique sufficiently outweigh the inertial burden of supplanting initial rates analyses.

Most *in vitro* assay platforms will support either initial rates or progress curve analyses, although continuous assays are particularly well-suited to progress curve analysis because a large number of data points can be collected with minimal effort on the part of the researcher. The majority of *in vitro* assays are fixed time-point assays, which are compatible with kinetic and mechanistic studies as described provided that it is possible to make measurements at multiple time points. The following sections describe the common assay platforms employed to measure EGFR kinase activity *in vitro* using purified kinase, *in cellulo* using cellular extracts or fixed cells, and *in cellulo* in live cells.

### **1.5 Assay Platforms for Direct Measurement of EGFR Kinase Activity**

In the following sections, a brief review of EGFR kinase activity assays is provided. The review focuses on how these assays are practically utilized and aims to illustrate the types of experimental questions each technique is useful to answer. Assay limitations are also discussed. Table 2 summarizes the assays discussed.

### 1.5.1 *In Vitro* Radiometric Assays

*In vitro* EGFR assays fall into two classes based on the detection method used: radiometric assays using radiolabeled ATP and fluorescence/luminescence-based assays. Radiometric assays have historically been the preferred platform for kinase activity assays due to their simplicity, robustness, and sensitivity, and are considered the gold standard for measuring EGFR kinase activity.<sup>25</sup> However, radiometric assays are expensive and hazardous to human health and to the environment. Therefore, non-hazardous assays using fluorescence or luminescence readouts have supplanted radiometric assays in recent years, particularly for high throughput applications.

Radiometric kinase assays measure the incorporation of <sup>32</sup>P or <sup>33</sup>P from radiolabeled ATP into peptide or protein substrates. In a typical assay, kinase, substrate, ATP, and radiolabeled ATP are co-incubated in solution, followed by spotting of the assay mixture onto a negatively-charged phosphocellulose membrane. The anionic membrane captures the protein or peptide substrate, which either natively or by design contains a cationic binding sequence. The membrane can then be extensively washed to remove excess radiolabeled ATP before treating the membrane with scintillant and detecting the emitted photons. This approach was applied by Rusnak and colleagues to characterization of an EGFR/Her2 inhibitor GW2016 (lapatinib, Tykerb<sup>®</sup>, GlaxoSmithKline).<sup>41</sup> Radiometric assays using peptide

substrates were used to compare inhibition of a panel of kinases by GW2016 and explore inhibitor selectivity. GW2016 has since become a successful targeted drug used to treat various forms of metastatic breast cancer.

More recently, Anastassiadis and colleagues<sup>42</sup> reported the use of a radiometric “HotSpot” assay to generate a comprehensive map of kinase inhibitor selectivity. The authors assayed a panel of 300 kinases including EGFR for inhibition by nominally selective targeted inhibitors using an automated radiometric assay platform. The authors compared the inhibitor selectivity scores obtained with their radiometric assays to similar studies using kinase-inhibitor binding assays and found that the binding assays were prone to both false positives and false negatives. The authors were also able to comprehensively profile the selectivity of 178 kinase inhibitors with good coverage across the kinome and to generate selectivity maps to guide future use of these inhibitors. Interestingly, the authors found that a 4,6-dianilinopyrimidine EGFR inhibitor displayed the greatest degree of “uni-specificity,” inhibiting substrate phosphorylation by EGFR by more than 94% while inhibiting the next most susceptible target by only 22%. Because the authors used radiometric assays, they were able to directly quantify substrate turnover by the kinase panel. Therefore they were able to unambiguously and accurately assess the influence of the inhibitor panel on kinase-mediated substrate turnover with minimal interference to confound results and generate false positives or negatives.



### 1.5.2 *In Vitro* Fluorescence/Luminescence-Based Assays

Use of radiometric assays for high-throughput studies such as that described above are uncommon because disposal of such a large quantity of hazardous waste can become cost prohibitive. More commonly, fluorescence/luminescence based platforms are employed. An impressive variety of such platforms have been reported<sup>25,43,44</sup>—far too many to be described here. The reader is referred to a series of reviews which comprehensively cover *in vitro* measurement of kinase activity using fluorescence or luminescence-based platforms. Here, selected examples of the most common assay platforms that have been applied to measurement of EGFR activity are recommended to the reader's attention.

Generally, fluorescence/luminescence kinase assays are divided into those that use phospho-specific antibodies to detect phosphorylation and antibody-free methods. Of these, the antibody-dependent methods are the most common. Li and colleagues applied a commercially available homogeneous time-resolved fluorescence (HTRF) assay (KinEASE, CisBio) to the *in vitro* characterization of novel EGFR inhibitors.<sup>45</sup> Here, a short peptide substrate, biotinylated at the N-terminus, is phosphorylated by EGFR. Streptavidin labeled with a specialized FRET acceptor binds the biotinylated peptide, while an anti-phosphotyrosine antibody binds only phosphorylated peptide. The antibody is labeled with a Europium cryptate FRET donor. Excitation of the Eu complex around 320 nm leads to short-range energy

transfer (< 9 nm) to the FRET acceptor, which emits long-lived fluorescence at 665 nm. Because the emission lifetime is comparatively long, time-resolved measurements which allow for decay of short-lived background fluorescence are possible. This significantly improves the signal-to-background ratio of the assay. Li and colleagues applied the HTRF assay to determine IC<sub>50</sub> values for a large panel of novel EGFR inhibitors coupled to bioprobes and were able to show improved potency of bioprobe-coupled inhibitors compared with their uncoupled parent compounds. Because the HTRF assay is homogeneous and highly miniaturizable, it is well suited to such large-scale studies where reagent consumption and laborious wash steps can be prohibitively resource expensive. Nevertheless, antibody-dependent kinase assays are limited by the availability of selective detection antibodies. Furthermore, these assays are generally endpoint-only, so kinetic experiments are laborious.

The Imperiali lab has published a variety of continuous kinase activity assays based on chelation-enhanced fluorescence (CHEF).<sup>46-48</sup> The technology, commercialized as the Omnia® assay by Life Technologies, has recently been applied to mechanistic studies of irreversible EGFR inhibitors.<sup>49</sup> The Omnia EGFR assay uses a specially-designed peptide sensor consisting of a kinase recognition sequence, including phosphorylation site, a beta-turn motif, and a specialized Sox fluorophore. Phosphorylation of the peptide leads to chelation of a Mg<sup>2+</sup> cation,

bridging the phosphoryl with the Sox label to enhance fluorescence 2-10 fold. The Sox sensor provides a direct, kinetic readout of EGFR kinase activity. Schwartz and colleagues determined the mechanism of covalent EGFR inhibition by modeling changes in time-dependent *in vitro* Sox sensor phosphorylation as a function of EGFR inhibition.<sup>49</sup>

Chang and colleagues applied another popular continuous EGFR assay to screen for irreversible EGFR inhibitors.<sup>50</sup> The Z'-Lyte assay, available from Life Technologies, is an *in vitro* FRET-based platform. A peptide substrate is engineered with a phosphorylation site located proximal to a protease cleavage site. The peptide is labeled at N- and C-termini with a FRET dye pair. Phosphorylation of the peptide by EGFR inhibits proteolytic cleavage in a coupled reaction, while the non-phosphorylated peptide is cleaved. Therefore, FRET is allowed to occur only when the peptide is phosphorylated, generating a signal that is directly proportional to substrate turnover by EGFR. The authors applied the Z'-Lyte assay to determine IC<sub>50</sub> values for a series of EGFR inhibitors and to compare inhibitor potency against several clinically important EGFR mutants.

A common feature of the Omnia and Z'-Lyte assays is the need to design peptide substrates that are both efficiently phosphorylated by the target kinase and highly compatible with the assay platform. Because substrate design is non-trivial, fairly promiscuous substrates which are suitable for multiple kinases are desirable.

Therefore, selectivity is problematic, limiting the use of these assays in more complex matrices such as cellular extracts.

### 1.5.3 *In Vitro* Mobility Shift Assay

Caliper's mobility shift assay platform stands out among *in vitro* kinase assays because it quantifies both phosphorylated and non-phosphorylated substrates by incorporating an analytical separation step.<sup>25</sup> The kinase reaction is typically carried out off-line, followed by analysis on a specialized microfluidic chip. The key to this technology is the ability separate non-phosphorylated and phosphorylated peptide substrates based on differential mobility in an electric field. The substrates are pre-labeled with a fluorophore to permit on-chip fluorescence detection. Automated sampling of the off-line kinase reaction at regular intervals enables essentially continuous assay monitoring. Because this approach measures both substrate and product, no calibration curve is required to quantify either species. Rather, the absolute quantity of phosphorylated substrate can be calculated from the fraction of substrate turnover and the known quantity of total substrate in the reaction. Xie and colleagues took advantage of the capability for continuous assay monitoring afforded by the Caliper chip to acquire progress curves of peptide phosphorylation by EGFR with and without a novel EGFR inhibitor or a control inhibitor.<sup>51</sup> They demonstrated based on the shape of the progress curves that the novel EGFR inhibitor binds and irreversibly inhibits the kinase, while the control

inhibitor binds but eventually dissociates from the kinase, allowing it to recover activity.<sup>51</sup>

#### **1.5.4 *In Cellulo* Assay Platforms for Cell Lysates**

Undoubtedly the most common and popular assay of EGFR activity in cell lysates is the classical Western blot; here, proteins from a cellular extract are pre-separated by denaturing gel electrophoresis, transferred to an ionic membrane, and probed with phospho-specific antibodies. Typically, phosphorylation of several endogenous EGFR substrates in the sample of interest is compared to controls to qualitatively assess the activity of the EGFR signaling pathway. The phosphorylated substrates measured typically include autophosphorylation sites within the EGFR kinase domain and EGFR-dependent phosphorylation sites in downstream signaling proteins such as members of the MAPK and PI3K/Akt pathways.<sup>52,53</sup> Western blotting is the gold standard biochemical EGFR assay for cellular lysates.

Yu, *et al.* reported a mass spectrometry-based method for monitoring EGFR-mediated signaling in a massively parallel manner.<sup>54</sup> The KAYAK (Kinase Activity Assay for Kinome profiling) was used to monitor phosphorylation of 90 unique peptide substrates covering a broad range of kinase activities downstream of EGFR. Peptide substrates were incubated individually with cell lysate and ATP, then the quenched reactions mixed with heavy isotope standards for each phosphopeptide. The 90 assays were then combined into two pools and separated by nano-LC prior to

MS analysis; inclusion of heavy isotope standards for each substrate enabled quantitative analysis of substrate phosphorylation. The authors were able to monitor pathway-specific changes in phosphorylation in response to targeted EGFR inhibition in immortalized tumor cell lines. KAYAK is a promising approach to studying network behavior downstream of EGFR, particularly in response to stimulation and inhibition. It provides information similar to large-scale western blotting, but does not require antibodies and requires only two LC-MS experiments to analyze 90 pathway components. A drawback of this method is the use of heavy isotope internal standards required for quantitation by mass spectrometry, which significantly increases the cost of the assay.

Ghosh and colleagues at the University of Wisconsin, Madison, published a novel and sensitive peptide-based EGFR assay designed to quantify EGFR activity in lysates of primary cells.<sup>55</sup> The authors modified magnetic beads with a set of EGFR peptide substrates and incubated the beads in 96-well plates with inhibitor-treated lysates of immortalized triple negative breast cancer cells lines. Following the kinase reaction, the beads were washed and phosphorylated peptide labeled with an HRP-conjugated anti-phosphotyrosine antibody. After further washing, the phosphorylation was detected by monitoring HRP-mediated conversion of Amplex Red to Resorufin in the presence of H<sub>2</sub>O<sub>2</sub>. The authors demonstrated reasonable selectivity of their assay for EGFR by showing differential assay response in lysates

of cell lines with known high or low expression of EGFR. The authors also showed good sensitivity of the assay, successfully measuring peptide phosphorylation with only 10 µg of lysate protein, or about 500 cancer cells.<sup>55</sup> This assay has yet to be validated in primary cells, but the possibility of easily measuring kinase activity in small numbers of cells is intriguing. It should be noted, however, that assay performance in primary cells is unlikely to match performance in tissue cultured cell lines where the cell populations are comparatively homogenous and EGFR activity is artificially high. It is probable that much larger primary cell populations will be required to reliably generate quantifiable signal. Furthermore, a major concern in the analysis of primary cancer cells is intratumoral heterogeneity which, undetected, can lead to disease relapse. Any technique measuring EGFR activity in a bulk lysate risks averaging out and obscuring rare or small subpopulations, missing critical therapeutic information. Therefore, techniques which measure EGFR activity in single cells are highly desirable.

Shi, *et al.* reported a microfluidic “single cell barcode chip” which can be used essentially to perform multiplexed ELISA assays in single cell lysates generated *in situ*.<sup>56</sup> One or a few EGFR-positive cultured glioblastoma cells were captured in wells of a two-layer microfluidic chip. Each well contained a spatially-defined array of eleven capture antibodies directed against EGFR and downstream signaling proteins. Cells in each well were lysed *in situ* using lysis buffer containing protease

and phosphatase inhibitors to minimize interference from off-target enzymes. After an incubation period to allow analyte proteins to bind to their respective capture antibodies and subsequent wash steps, biotinylated primary antibodies (e.g. directed against specific phosphorylation sites) were introduced to complete the ELISA “sandwich”. Detection was accomplished by labeling the biotinylated antibodies with fluorophore-labeled streptavidin, followed chip reading in a high resolution fluorescence scanning instrument.

### **1.5.5 *In Cellulo* Assay Platforms for Fixed Cells**

Leuchowius and colleagues developed a flow cytometry-based method for quantifying EGFR dimer formation and phosphorylation in single cells using a novel *in situ* proximity ligation method.<sup>57</sup> Fixed and permeabilized cells were labeled with EGFR-specific antibodies conjugated to oligonucleotides. Treatment with DNA ligase to create a cDNA bridge between co-localized antibodies was followed by DNA amplification and labeling with fluorophores. The fluorescence intensity, proportional either to EGFR dimer formation or EGFR phosphorylation depending on the detection antibodies chosen, was then quantified by flow cytometry. This method is very sensitive compared to traditional immunostaining because the DNA amplification step allows many fluorophores to bind to each antibody pair. The authors successfully applied this method to measure EGF-dependent EGFR phosphorylation in tissue-cultured cells.



Intercellular heterogeneity in EGFR expression and activation is of particular interest for clinical samples. Generally, cells of interest are obtained as fixed tissue slices that have been prepared by a clinical pathologist for immunohistochemistry. Tissue fixatives covalently crosslink intracellular proteins, essentially freezing the tissue in time. The tissue may then be stained with antibodies directed against EGFR or phosphorylated EGFR to qualitatively assess kinase overexpression and/or hyperactivation. Sonnweber and colleagues performed such immunostaining on tumors obtained from patients with early stage non-small cell lung cancer (NSCLC) and showed that EGFR phosphorylation, but not overexpression, correlated with poor prognosis.<sup>58</sup>

A serious drawback of immunohistochemical analyses is the subjective nature of the actual data analysis. Skilled pathologists assign a numerical score based on their assessment of the stained tissue; because each pathologist has different experience and qualifications, this leads to variable results. Kong and colleagues sought to provide a reliable alternative to IHC by developing a quantitative, automated fluorescence lifetime imaging (FLIM) method for determining EGFR phosphorylation in fixed tissue microarrays.<sup>59</sup> The authors converted fixed, paraffin embedded malignant tissue into rehydrated tissue microarrays, then stained the arrays with an anti-EGFR antibody and an anti-phosphotyrosine antibody. The antibodies were labeled with a cyanine (“Cy”) dye pair that, when co-localized,

could undergo FRET upon excitation of one of the dyes. Automated FLIM imaging of the tissue arrays allowed the authors to quantify EGFR phosphorylation in the primary tissues. The authors showed that the mean FRET efficiency from the brightest 10% of cells from biopsy cores correlated significantly with reduced disease-free survival, but did not correlate with overall survival. The authors were also able to demonstrate that total EGFR expression did not significantly correlate with either disease-free survival or overall survival. This finding was in agreement with other non-quantitative studies using immunostaining and western blot analyses.

Use of fixed tissue for single cell analysis of clinical samples is standard because fixation and subsequent processing allows tissue to be preserved for storage and pathological analysis at a convenient time. However, fixation terminates intracellular reactions by cross-linking intracellular proteins. Since the timescale of fixation may not match with the timescale of the biochemical reaction to be measured, measurement artifacts may be introduced that would be avoided if measurements were taken in intact living cells. Therefore, assay platforms have been developed for quantification of EGFR activity in single living cells.

### **1.5.3 *In Cellulo* Assay Platforms for Live Cells**

Approaches to monitoring EGFR activity in living cells vary widely. A particularly elegant approach was presented by Roger Tsien's lab, where a FRET

sensor comprised of an EGFR substrate peptide and a phosphotyrosine-binding SH2 domain separating a FRET pair of fluorescent proteins was genetically encoded in a cell line of interest.<sup>60</sup> Upon phosphorylation of the substrate region, the SH2 domain binds the phosphorylation site, bringing the fluorescent proteins into proximity and causing an increase in measured FRET that is proportional to the degree of phosphorylation. This approach enabled the authors to image the dynamics of sensor phosphorylation in response to EGFR stimulation. Other examples of genetically encoded FRET sensors of EGFR activity in living cells have been published<sup>61</sup>, inspired by Tsien's design. These sensors are invaluable research tools, but are poorly compatible with real clinical samples, which are heterogeneous and small, usually consisting of tens of thousands of cells or fewer and heavily contaminated with stromal cells, fibroblasts, and other cells from the tumor microenvironment. Furthermore, sensor encoding and validation can take days or weeks to complete; ideally, activity measurements in clinical specimens should be performed as soon as possible after collection to minimize changes in activity due to *ex vivo* sample handling and culture.

This work details the development of a single cell EGFR kinase assay which is compatible with small, heterogeneous populations of live cells. The assay is based on a microelectrophoretic chemical cytometry platform, described in subsequent sections, coupled to a capillary electrophoresis-based EGFR biochemical assay in

which a rationally designed peptide substrate is employed to directly report kinase activity at the single cell level.

## **1.6 Chemical Cytometry with Single Cell Capillary Electrophoresis**

### **1.6.1 Definition and Applications of Chemical Cytometry**

Dovichi defines “chemical cytometry” as the application of analytical chemistry to quantify the contents of single cells.<sup>62</sup> This requires highly specialized instrumentation capable of handling minute sample volumes (typical mammalian cell  $\approx 10^{-12}$  L) coupled to ultrasensitive detection methods. Chemical cytometry methods involve a separation component for multiplexed analysis of cellular contents. The separation technique best suited to handling the contents of single cells is capillary electrophoresis (CE). CE was first applied to chemical cytometry by Jorgenson in 1989.<sup>63</sup> In that early work, single giant neurons were manually lysed and their contents labeled with a fluorescent tag prior to separation by CE coupled to an arc lamp fluorescence detector. Shortly thereafter, Ewing demonstrated the use of CE to directly sample portions of the cytoplasm from large neurons by inserting the etched inlet of the capillary directly into the cell.<sup>64</sup> Ewing used electrochemical detection to quantify neurotransmitters from the neuronal cytoplasm. In recent years, significant advancements in detection sensitivity have enabled CE-based chemical cytometry of primary mammalian cells, including undirected (e.g. proteomics) and directed (target-specific) analyses. Dovichi’s group has demonstrated ultrasensitive

two-color CE-LIF analysis of lipid metabolism in single cells using fluorescent lipid reporters, with limits of detection as low as  $10^{-22}$  mol.<sup>65</sup> Sweedler's group has applied CE-MS to quantitation of > 100 small molecule metabolites in single neurons, with detection limits in the high attomole range ( $10^{-16}$  mol).<sup>66</sup> Mass spectrometry detection has yet to catch up with LIF detection in terms of sensitivity, but has the advantage of being label-free. The Allbritton group has developed a toolbox of fluorescent enzyme activity reporters for single cell biochemical assays in mammalian cells, using an ultrasensitive CE-LIF chemical cytometry platform with detection limits in the range of  $10^{-19}$ - $10^{-20}$  mol, depending upon the reporter used. Single cell biochemical assays reported by the Allbritton group include those for protein serine/threonine kinases,<sup>67,68</sup> intracellular proteases,<sup>69-71</sup> lipid kinases,<sup>72,73</sup> and protein tyrosine phosphatases.<sup>74,75</sup> In all cases, these technologies rely heavily on the exceptional separation power of CE to resolve analytes of interest in a highly complex sample matrix. The following section provides an overview of the theoretical aspects of CE and its figures of merit.

### **1.6.2. Theoretical Overview of Capillary Electrophoresis<sup>76,77</sup>**

CE is an analytical separation technique in which analytes are resolved based on differential mobility in an electric field which is generated by applying a voltage across an electrolyte-filled capillary (Figure 3). The rate of migration,  $v$ , of a given

analyte is determined by the product of the analyte's effective mobility,  $\mu_{eff}$ , and the applied field  $\mathbf{E}$ :

$$v = \mu_{eff} \mathbf{E} \quad (8)$$

The effective mobility of an analyte is the sum of its electrophoretic ( $\mu_{ep}$ ) and electroosmotic ( $\mu_{eo}$ ) mobilities (9).

$$\mu_{eff} = \mu_{ep} + \mu_{eo} \quad (9)$$

The electrophoretic mobility depends upon properties of the analyte itself within the context of the separation background electrolyte (BGE) being employed. It is directly proportional to the charge on the analyte and indirectly proportional to drag forces experience by the analyte. While the analyte charge is a function of the chemistry of the analyte and BGE, the drag forces are related to the size and shape of the analyte in solution. The electrophoretic mobility can be written as

$$\mu_{ep} = \frac{q}{f} \quad (10)$$

where  $q$  is the net charge on the analyte and  $f$  describes the drag forces experienced by the analyte. For a spherical analyte, the mobility is written

$$\mu_{ep} = \frac{q}{6\pi\eta a} \quad (11)$$

where  $\eta$  is the viscosity of the BGE and  $a$  is the hydrodynamic radius of the analyte.

For many analytes, particularly biopolymers such as DNA and proteins, the relationship of molecular weight to electrophoretic mobility is of interest. Several

models, based on geometric parameters associated with various solvated polymer conformations, have been proposed to predict the effect of molecular weight on mobility. No “one size fits all” model has emerged which can reliably predict mobility, particularly for peptides. For small spherical analytes, the electrophoretic mobility is proportional to  $(MW)^{-1/3}$ . The relationship for proteins, described by the Offord model, predicts the mobility to vary with  $(MW)^{-2/3}$ . The MW-mobility relationship for peptides of moderate length (< 40 a.a.) lies somewhere between these two extremes.<sup>78</sup>

The electroosmotic mobility is the mobility of an analyte with a net charge of zero. It is due to the phenomenon of electroosmosis: bulk fluid flow based on mobility of BGE ions in the electrical double layer (Figure 3B). In typical capillary zone electrophoresis (CZE), the separation capillary is an open tubular column made of unmodified fused silica, which presents ionizable silanol groups at the surface of the capillary wall. Deprotonation of surface silanol groups gives the capillary wall a net anionic charge (Figure 3C). Cationic components of the BGE specifically adsorb at the capillary surface, forming an immobile, cationic layer called the Stern layer. The local concentration of cations extends into the mobile (“diffuse”) layer. The Stern and diffuse layers together comprise the electrical double layer at the capillary surface, forming a potential gradient extending from the surface into the bulk solution. The potential at the outer edge of the diffuse layer is called the zeta

potential ( $\zeta$ ). In an applied field, cations in the diffuse layer migrate in the direction of the cathode. Cohesive forces in the BGE cause the bulk solution to migrate with the mobile layer, inducing bulk flow with a characteristic flat flow profile which facilitates the exceptional separation efficiencies for which CE is known. The electroosmotic flow (EOF) depends upon the surface chemistry of the capillary wall and the composition and pH of the BGE. The electroosmotic mobility can be written as

$$\mu_{eo} = \frac{\epsilon}{4\pi\eta} \zeta \mathbf{E} \quad (12)$$

where  $\epsilon$  is the dielectric constant of the BGE. The electroosmotic mobility ( $\mu_{eo}$ ) is inversely proportional to the natural logarithm of the ionic strength of the BGE.<sup>79</sup>

For CZE in open tubular columns in the absence of hydrodynamic or convective flow, band broadening depends only on molecular diffusion of the analyte as described by the Einstein equation:

$$\sigma = \sqrt{2Dt} \quad (13)$$

Therefore, the number of theoretical plates  $N$  achievable in CZE is given by

$$N = \frac{L^2}{\sigma^2} = \frac{\mu_{eff}V}{2D} \quad (14)$$

where  $L$  is the length of the capillary column and  $V$  is the applied voltage. In terms of the Van Deemter model of theoretical plate height ( $H$ ),  $H$  in CE depends only on the longitudinal diffusion ( $B$ ) term and is inversely proportional to migration rate.



Since rate is proportional to the applied voltage, it is apparent that the separation efficiency in CZE benefits from the highest possible applied voltage. Maximizing the applied voltage also has the benefit of increasing the resolution  $R_s$  between analyte bands:

$$R_s = 0.177(\mu_{ep,1} - \mu_{ep,2}) \left[ \frac{V}{\bar{D}(\bar{\mu}_{ep} + \mu_{eo})} \right]^{1/2} \quad (15)$$

From this equation, one may infer that maximal resolution is achieved when the electroosmotic mobility perfectly counterbalances the average electrophoretic mobility of the analytes ( $R_s \rightarrow \infty$ ). Obviously, both the applied voltage and factors governing  $\mu_{eo}$  and  $\mu_{ep}$  (BGE chemistry, capillary treatment) should be optimized to provide acceptable resolution in the shortest separation time possible to maximize separation efficiency. The upper limit of  $V$  in most cases is efficient heat dissipation across the capillary to prevent radial temperature gradients which cause band broadening and convective effects leading to nonlinearity of the voltage-current relationship (Joule heating). The Joule heating limit for a given capillary length and internal diameter depends strongly on the BGE used and varies widely for different buffer systems at the same concentration. Best practice for achieving a highly efficient CZE separation is to develop an optimized BGE system (buffer composition, pH, additives) permitting acceptable resolution of all analytes in the shortest

possible time, and to operate under the highest applied voltage allowed by the optimized BGE.

CE is amenable to a variety of operation modes determined by the BGE and capillary modifications. CE in open tubular columns with typical homogeneous BGE, termed CZE as previously discussed, is the most versatile CE mode. Other specialized modes have been developed which introduce different mechanisms of selectivity to the separation. Micellar electrokinetic chromatography (MEKC), introduced by Terabe in the early 1980s<sup>80</sup>, is a useful CE variant in which the normal BGE is supplemented with a surfactant at concentrations well above the surfactant's critical micelle concentration (CMC). The micelles act as a pseudostationary phase which is usually engineered to migrate in the opposite direction of the EOF. Analytes partition in and out of the micellar phase, such that their effective mobility is influenced by their affinity for the pseudostationary phase. MEKC is viewed as a hybrid electrophoretic/chromatographic technique and has been proposed as a flexible, economical and environmentally-friendly alternative to traditional analytical HPLC.<sup>81</sup> This mode adds an orthogonal dimension of selectivity to CZE and can significantly augment separation power, particularly in the case of analytes with similar charge-to-mass ratios.

Capillary isoelectric focusing (cIEF) is an important CE mode widely applied to biological macromolecule characterization. In cIEF, the capillary is filled with a

pH gradient-forming BGE. Charged analytes migrate in the applied field until they reach a pH zone corresponding to their isoelectric point; at this point, the analyte's net charge is zero, causing its mobility to go to zero. Bands can be detected either *in situ* by whole-column imaging, or conventionally at a detection window or flow cell following analyte mobilization. CIEF is a powerful technique for high resolution determination of protein charge variants, for example due to post-translational modification.<sup>82,83</sup> It has therefore increased in popularity with the recent growth in the therapeutic protein market.

Exceptional separation efficiencies can routinely be achieved in CE analyses, with theoretical plates in excess of  $10^6/m$  in many cases.<sup>84</sup> Since  $N$  is independent of capillary length ( $L$ ), increasing this parameter leads to very large peak capacities, theoretically without sacrificing separation quality. Thus, CE is ideal for analysis of complex samples such as biologic cells. CE supports very small sample volumes ( $< 1 \mu\text{L}$ ); the minute sample consumption facilitates analysis of precious/rare samples, such as small populations of primary cells from tumor biopsies. Achieving limits of detection suitable to such small sample sizes has been a major challenge in developing CE instrumentation. Many detectors have been successfully coupled to CE, including UV, fluorescence, and mass spectrometry.<sup>85,86</sup> Laser-induced fluorescence (LIF) has been employed to achieve the lowest limits of detection ( $5 \times 10^{-23}$  mol) of any CE detector.<sup>87</sup> The caveat to LIF is the need for fluorescent analytes.

Since most analytes of interest, particularly in biological matrices, are not natively fluorescent, derivatization may be required. Mass spectrometry is label-free, but is difficult to couple to CE due to electric field issues and BGE constraints and is only sufficiently sensitive to detect the most abundant species in the cell (e.g. ATP).

Electrochemical detection has been coupled to single cell CE and is suitable for sensitive detection of electroactive analytes such as dopamine<sup>88</sup> and nitric oxide<sup>89</sup>. In general, however, CE-LIF is employed for single cell analysis of low-abundance species.<sup>62</sup>

### **1.6.3 Single Cell Capillary Electrophoresis Assay of EGFR Kinase Activity**

The following chapters detail the development and implementation of a single-cell biochemical assay of EGFR activity optimized for use with chemical cytometry by CE-LIF. Figure 4A and B provides a diagrammatic overview of the assay. Briefly, a carboxyfluorescein-labeled peptide substrate (“reporter”) of EGFR is loaded into a population of cells of interest, followed by an incubation period to allow any EGFR in the cell to phosphorylate the reporter. The cells are seeded on a glass coverslip coated with extracellular matrix to promote rapid adhesion. To analyze an individual cell, the cell chamber is mounted in a flow chamber over the objective of an inverted microscope, flushed for a few minutes with 37 °C extracellular buffer, and a cell chosen for analysis by microscopy. The cell is then lysed very rapidly (< 5 ms) by a pulse from a Nd:YAG laser focused adjacent to the

cell. The laser forms a plasma in the buffer near the cell, which generates a cavitation bubble. Collapse of the cavitation bubble creates shear forces which rapidly disrupt the cellular membrane and terminate intracellular biochemical reactions through dilution. The cellular contents are loaded electrokinetically into the lumen of the separation capillary, positioned a few microns above the cell. The capillary inlet is moved from the cell chamber to the inlet BGE reservoir and separation initiated by application of a high voltage. The unmodified reporter is separated from phosphorylated reporter and any other metabolic products by CE and detected by ultrasensitive LIF. The EGFR kinase activity is calculated from the ratio of phosphorylated to total peptide in the electropherograms; this ratiometric readout is very robust to experimental variability and allows absolute quantification by comparison to a standard.

The described assay is ideal for single cell analysis of primary cells. The single cell CE platform supports very small and/or very heterogeneous cell populations and is easily coupled to immunofluorescence microscopy to allow identification of cells of interest based on surface markers. For example, staining with fluorophore-labeled anti-EpCAM facilitates differentiation of epithelium-derived tumor cells from fibroblasts and stromal cells that can comprise the bulk of biopsy samples. The exceptional separation power of CE allows reporter multiplexing to measure multiple signaling pathways simultaneously. The combination of very small sample

volumes for CE and excellent mass sensitivity for LIF detection results in limits of detection on the order of  $10^{-20}$  mol. For a typical mammalian cell with a volume of 1 pL, this corresponds to an intracellular reporter concentration of 10 nM. This exquisite sensitivity allows reporter loading into cells to be kept to a minimum to prevent measurement artifacts due to reporter interference in native signaling interactions. The single cell CE system is amenable to automation and high-throughput analysis. While the best published throughput for such a system does not rival massively high throughput techniques such as FACS, it is nevertheless possible to analyze hundreds of cells per hour.<sup>90</sup> Further increasing throughput is an area of active investigation in the Allbritton lab. With higher throughput it is possible to analyze populations of cells large enough to allow statistical analyses to tease apart relevant biochemical subpopulations. This is a particularly pressing need in the clinical treatment of EGFR-dependent solid tumors, where single cell analysis is essential to provide information about clinically relevant heterogeneity affecting response to targeted inhibitors. Flow cytometry is not suitable in general for biopsies from solid tumors because the technique requires millions of cells for robust analysis, while biopsies from solid tumors may contain as few as ten thousand cells. Of this small population, tumor cells may comprise only a small fraction. Single cell CE can address each cell in the small population with minimal waste or information loss to provide a holistic picture of biochemical heterogeneity in clinical samples.

The foundation of a successful single cell biochemical assay is the activity reporter. For EGFR and other tyrosine kinases, this is a fluorophore-labeled, acidic peptide with a tyrosine phosphoacceptor surrounded by recognition elements tailored to the kinase of interest. The reporter peptide is synthesized and characterized *in vitro* under controlled conditions, providing a robust reference for behavior *in cellulo*. This stands in stark contrast to genetically encoded reporters that must be characterized *in cellulo* with very limited control over experimental variables. The exogenous reporter can be introduced into the cytosol of intact cells by any of a variety of methods, including pinocytosis, electroporation, or microinjection. The reporter may alternatively be conjugated to a lipid or cell penetrating peptide to promote passive transport into the cytosol. The reporter sequence may be easily optimized to suit experimental needs. For example, native amino acids within the reporter may be replaced with unnatural analogs to increase selectivity for the target kinase and improve reporter stability in the cytosol. Finally, the separation component of the assay permits direct detection of the reporter in all its forms, including the unmodified reporter, its phosphoform, and any proteolytic products. Importantly, since both the unmodified and phosphorylated reporter are detected in the same experiment, the assay directly quantifies EGFR biochemical activity in each cell. It is therefore complementary to existing clinical EGFR assays which measure aberrancies in the EGFR pathway in single primary cells at the

genetic or protein expression level. The single cell EGFR kinase assay may be implemented in conjunction with these existing clinical assays to provide a holistic picture of EGFR activation in primary cells and to improve identification of patients likely to respond targeted EGFR inhibitors.



## REFERENCES

- (1) Price, N. C.; Stevens, L. *Enzyme technology, fundamentals of enzymology: The cell and molecular biology of catalytic proteins*; Oxford University Press, 1999.
- (2) Purvis, J. E.; Lahav, G. *Cell* **2013**, *152*, 945–956.
- (3) Kholodenko, B. N.; Hancock, J. F.; Kolch, W. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 414–426.
- (4) Blume-Jensen, P.; Hunter, T. *Nature* **2001**, *411*, 355–365.
- (5) Lemmon, M. A.; Schlessinger, J. *Cell* **2010**, *141*, 1117–1134.
- (6) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912–1934.
- (7) Lemmon, M. A.; Schlessinger, J.; Ferguson, K. M. *Cold Spring Harbor perspectives in biology* **2014**, *6*, a020768.
- (8) Bose, R.; Zhang, X. *Experimental Cell Research* **2009**, *315*, 649–658.
- (9) Yarden, Y.; Sliwkowski, M. X. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 127–137.
- (10) Linggi, B.; Carpenter, G. *Trends in Cell Biology* **2006**, *16*, 649–656.
- (11) Leahy, D. J. *Cell* **2010**, *142*, 513–515.
- (12) Han, W.; Lo, H. W. *Cancer Letters* **2012**.
- (13) Avraham, R.; Yarden, Y. *Nat. Rev. Mol. Cell Biol.* **2011**.
- (14) Levitzki, A. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 161–185.
- (15) Sharma, S. V.; Fischbach, M. A.; Haber, D. A.; Settleman, J. *Clin. Cancer Res.* **2006**, *12*, 4392s–4395s.
- (16) Lucchini, E.; Pilotto, S.; Spada, E.; Melisi, D.; Bria, E.; Tortora, G. *Expert Opin Drug Saf* **2014**, *13*, 535–549.

- (17) Baselga, J. *Science* **2006**, *312*, 1175–1178.
- (18) Lindeman, N. I.; Cagle, P. T.; Beasley, M. B.; Chitale, D. A.; Dacic, S.; Giaccone, G.; Jenkins, R. B.; Kwiatkowski, D. J.; Saldivar, J.-S.; Squire, J.; Thunnissen, E.; Ladanyi, M. *The Journal of Molecular Diagnostics* **2013**, *15*, 415–453.
- (19) Braun, S.; Raymond, W. E.; Racker, E. *J. Biol. Chem.* **1984**, *259*, 2051–2054.
- (20) Casnellie, J. E. *Meth. Enzymol.* **1991**, *200*, 115–120.
- (21) Hunter, T. *J. Biol. Chem.* **1982**, *257*, 4843–4848.
- (22) Pearson, R. B.; Kemp, B. E. *Meth. Enzymol.* **1991**, *200*, 62–81.
- (23) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- (24) Merrifield, R. B. *Biochemistry* **1964**, *3*, 1385–1390.
- (25) Jia, Y.; Gu, X.-J.; Brinker, A.; Warmuth, M. *Expert Opin. Drug Discov.* **2008**, *3*, 959–978.
- (26) Michaelis, L.; Menten, M. *FEBS Lett.* **2013**.
- (27) Schnell, S.; Maini, P. K. *Comm. Theoret. Biol* **2003**, *8*, 169–187.
- (28) Segel, I. H. *Biochemical Calculations (2nd edit.)* Wiley; New York, 1976.
- (29) Duggleby, R. *Methods* **2001**, *24*, 168–174.
- (30) Goličnik, M. *Analytical Biochemistry* **2010**, *406*, 94–96.
- (31) Goličnik, M. *Biochemical Engineering Journal* **2012**, *63*, 116–123.
- (32) Goličnik, M. *Eng. Life Sci.* **2011**, *12*, 104–108.
- (33) Goličnik, M. *Analytical Biochemistry* **2011**, *411*, 303–305.
- (34) Goličnik, M. *Biochemical Engineering Journal* **2011**, *53*, 234–238.

- (35) Goličnik, M. *Biochem. Mol. Biol. Educ.* **2011**, *39*, 117–125.
- (36) Johnson, K. A. *FEBS Lett.* **2013**, *587*, 2753–2766.
- (37) Zavrel, M.; Kochanowski, K.; Spiess, A. C. *Eng. Life Sci.* **2010**, *10*, 191–200.
- (38) Goudar, C. T.; Harris, S. K.; McInerney, M. J.; Suflita, J. M. *Journal of Microbiological Methods* **2004**, *59*, 317–326.
- (39) Goudar, C. T. *Journal of Biotechnology* **2012**, *159*, 56–60.
- (40) Goudar, C. T. *Biotechnol. Bioeng.* **2011**, *108*, 2499–2503.
- (41) Rusnak, D. W.; Lackey, K.; Affleck, K.; Wood, E. R.; Alligood, K.; Rhodes, N.; Keith, B. R.; Murray, D. M.; Knight, W. B.; Mullin, R. J.; Gilmer, T. M. *Molecular Cancer Therapeutics* **2001**, *1*, 85–94.
- (42) Anastassiadis, T.; Deacon, S. W.; Devarajan, K.; Ma, H.; Peterson, J. R. *Nature Biotechnology* **2011**, *29*, 1039–1045.
- (43) Jia, Y.; Quinn, C. M.; Kwak, S.; Talanian, R. V. *Curr Drug Discov Technol* **2008**, *5*, 59–69.
- (44) Morris, M. C. *Biochim. Biophys. Acta* **2013**, *1834*, 1387–1395.
- (45) Li, S.; Guo, C.; Sun, X.; Li, Y.; Zhao, H.; Zhan, D.; Lan, M.; Tang, Y. *Eur J Med Chem* **2012**, *49*, 271–278.
- (46) Shults, M. D.; Imperiali, B. *J. Am. Chem. Soc.* **2003**, *125*, 14248–14249.
- (47) Shults, M. D.; Janes, K. A.; Lauffenburger, D. A.; Imperiali, B. *Nat Meth* **2005**, *2*, 277–283.
- (48) Rothman, D. M.; Shults, M. D.; Imperiali, B. *Trends in Cell Biology* **2005**, *15*, 502–510.
- (49) Schwartz, P. A.; Kuzmic, P.; Solowiej, J.; Bergqvist, S.; Bolanos, B.; Almaden, C.; Nagata, A.; Ryan, K.; Feng, J.; Dalvie, D.; Kath, J. C.; Xu, M.; Wani, R.; Murray, B. W. *Proceedings of the National Academy of Sciences* **2014**, *111*, 173–

178.

- (50) Chang, S.; Zhang, L.; Xu, S.; Luo, J.; Lu, X.; Zhang, Z.; Xu, T.; Liu, Y.; Tu, Z.; Xu, Y.; Ren, X.; Geng, M.; Ding, J.; Pei, D.; Ding, K. *J. Med. Chem.* **2012**, *55*, 2711–2723.
- (51) Xie, H.; Lin, L.; Tong, L.; Jiang, Y.; Zheng, M.; Chen, Z.; Jiang, X.; Zhang, X.; Ren, X.; Qu, W.; Yang, Y.; Wan, H.; Chen, Y.; Zuo, J.; Jiang, H.; Geng, M.; Ding, J. *PLoS ONE* **2011**, *6*, e21487.
- (52) Guo, A.; Villen, J.; Kornhauser, J.; Lee, K. A.; Stokes, M. P.; Rikova, K.; Possemato, A.; Nardone, J.; Innocenti, G.; Wetzell, R.; Wang, Y.; MacNeill, J.; Mitchell, J.; Gygi, S. P.; Rush, J.; Polakiewicz, R. D.; Comb, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 692–697.
- (53) Amann, J.; Kalyankrishna, S.; Massion, P. P.; Ohm, J. E.; Girard, L.; Shigematsu, H.; Peyton, M.; Juroske, D.; Huang, Y.; Stuart Salmon, J.; Kim, Y. H.; Pollack, J. R.; Yanagisawa, K.; Gazdar, A.; Minna, J. D.; Kurie, J. M.; Carbone, D. P. *Cancer Research* **2005**, *65*, 226–235.
- (54) Yu, Y.; Anjum, R.; Kubota, K.; Rush, J.; Villen, J.; Gygi, S. P. *Proceedings of the National Academy of Sciences* **2009**, *106*, 11606–11611.
- (55) Ghosh, G.; Yan, X.; Kron, S. J.; Palecek, S. P. *Assay Drug Dev Technol* **2013**, *11*, 44–51.
- (56) Shi, Q.; Qin, L.; Wei, W.; Geng, F.; Fan, R.; Shin, Y. S.; Guo, D.; Hood, L.; Mischel, P. S.; Heath, J. R. *Proceedings of the National Academy of Sciences* **2012**, *109*, 419–424.
- (57) Leuchowius, K. J.; Weibrecht, I.; Landegren, U.; Gedda, L.; Söderberg, O. *Cytometry Part A* **2009**, *75A*, 833–839.
- (58) Sonnweber, B.; Dlaska, M.; Skvortsov, S.; Dirnhofer, S.; Schmid, T.; Hilbe, W. *Journal of Clinical Pathology* **2006**, *59*, 255–259.
- (59) Kong, A.; Leboucher, P.; Leek, R.; Calleja, V.; Winter, S.; Harris, A.; Parker, P. J.; Larijani, B. *Cancer Research* **2006**, *66*, 2834–2843.
- (60) Ting, A. Y.; Kain, K. H.; Klemke, R. L.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.*

- 2001, 98, 15003–15008.
- (61) Offferdinger, M.; Georget, V.; Girod, A.; Bastiaens, P. I. H. *J. Biol. Chem.* **2004**, 279, 36972–36981.
- (62) Dovichi, N. J. In *Chemical Cytometry*; Lu, C., Ed.; 2010; pp. 1–20.
- (63) Kennedy, R. T.; Oates, M. D.; Cooper, B. R.; Nickerson, B.; Jorgenson, J. W. *Science* **1989**, 246, 57–63.
- (64) Olefirowicz, T. M.; Ewing, A. G. *Anal. Chem.* **1990**, 62, 1872–1876.
- (65) Essaka, D. C.; Prendergast, J.; Keithley, R. B.; Palcic, M. M.; Hindsgaul, O.; Schnaar, R. L.; Dovichi, N. J. *Anal. Chem.* **2012**, 84, 2799–2804.
- (66) Nemes, P.; Knolhoff, A. M.; Rubakhin, S. S.; Sweedler, J. V. *Anal. Chem.* **2011**, 83, 6810–6817.
- (67) Proctor, A.; Herrera-Loeza, S. G.; Wang, Q.; Lawrence, D. S.; Yeh, J. J.; Allbritton, N. L. *Anal. Chem.* **2014**, 86, 4573–4580.
- (68) Li, H.; Sims, C. E.; Kaluzova, M.; Stanbridge, E. J.; Allbritton, N. L. *Biochemistry* **2004**, 43, 1599–1608.
- (69) Kovarik, M. L.; Dickinson, A. J.; Roy, P.; Poonnen, R. A.; Fine, J. P.; Allbritton, N. L. *Integr Biol (Camb)* **2014**, 6, 164–174.
- (70) Kovarik, M. L.; Shah, P. K.; Armistead, P. M.; Allbritton, N. L. *Anal. Chem.* **2013**, 85, 4991–4997.
- (71) Proctor, A.; Wang, Q.; Lawrence, D. S.; Allbritton, N. L. *Analyst* **2012**, 137, 3028–3038.
- (72) Dickinson, A. J.; Hunsucker, S. A.; Armistead, P. M.; Allbritton, N. L. *Anal Bioanal Chem* **2014**, 406, 7027–7036.
- (73) Jiang, D.; Sims, C. E.; Allbritton, N. L. *Faraday Discussions* **2011**, 149, 187–200.
- (74) Phillips, R. M.; Dailey, L. A.; Bair, E.; Samet, J. M.; Allbritton, N. L. *Anal. Chem.* **2014**, 86, 1291–1297.

- (75) Phillips, R. M.; Bair, E.; Lawrence, D. S.; Sims, C. E.; al, E. *Anal. Chem.* **2013**, *85*, 6136–6142.
- (76) Jorgenson, J. W. CHEM 444. Separations, 2009.
- (77) Weinberger, R. *Practical Capillary Electrophoresis*; 2nd ed.; Harcourt, Inc.: San Diego, CA, 2000.
- (78) Adamson, N. J.; Reynolds, E. C. *J. Chromatogr. B Biomed. Sci. Appl.* **1997**, *699*, 133–147.
- (79) VanOrman, B. B.; Liversidge, G. G.; McIntire, G. L.; Olefirowicz, T. M.; Ewing, A. G. *Journal of Microcolumn Separations* **1990**, *2*, 176–180.
- (80) Terabe, S.; Otsuka, K.; Ando, T. *Anal. Chem.* **1985**, *57*, 834–841.
- (81) Deeb, S. E.; Iriban, M. A.; Gust, R. *ELECTROPHORESIS* **2011**, *32*, 166–183.
- (82) Deeb, S. E.; Wätzig, H.; El-Hady, D. A.; Albishri, H. M.; de Griend, C. S.-V.; Scriba, G. K. E. *ELECTROPHORESIS* **2014**, *35*, 170–189.
- (83) Staub, A.; Guillarme, D.; Schappler, J.; Veuthey, J.-L.; Rudaz, S. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *55*, 810–822.
- (84) Li, S. F. Y. *Capillary Electrophoresis*; Elsevier, 1992.
- (85) Swinney, K.; Bornhop, D. J. *ELECTROPHORESIS* **2000**.
- (86) Frost, N. W.; Jing, M.; Bowser, M. T. *Anal. Chem.* **2010**, *82*, 4682–4698.
- (87) Yong Chen, D.; Dovichi, N. J. *J. Chromatogr. B Biomed. Sci. Appl.* **1994**, *657*, 265–269.
- (88) Olefirowicz, T. M.; Ewing, A. G. *Journal of Neuroscience Methods* **1990**, *34*, 11–15.
- (89) Metto, E. C.; Evans, K.; Barney, P.; Culbertson, A. H.; Gunasekara, D. B.; Caruso, G.; Hulvey, M. K.; Fracassi da Silva, J. A.; Lunte, S. M.; Culbertson, C. T. *Anal. Chem.* **2013**, *85*, 10188–10195.

- (90) Dickinson, A. J.; Armistead, P. M.; Allbritton, N. L. *Anal. Chem.* **2013**, *85*, 4797–4804.

**Table 1. Common Components of Kinase Assays**

Source of Kinase	Substrate	Sample Matrix	Assay Format	Detection Mode
<ul style="list-style-type: none"> <li>• Recombinant purified enzyme</li> <li>• Cell lysate</li> <li>• Immuno-precipitated enzyme from cell lysate</li> <li>• Intact cell</li> </ul>	<ul style="list-style-type: none"> <li>• Full-length protein</li> <li>• Protein domain</li> <li>• Tryptic fragment peptide</li> <li>• Random amino acid polymer</li> <li>• Synthetic peptide</li> <li>• <math>\gamma</math>-<sup>32</sup>P-ATP</li> <li>• Unlabeled ATP</li> </ul>	<ul style="list-style-type: none"> <li>• Buffer (HEPES, MOPS, Tris-HCl)</li> <li>• Salt (high concentration NaCl)</li> <li>• Surfactants (SDS, NP-40, Tween-20)</li> <li>• Stabilizers (BSA, glycerol)</li> <li>• Divalent cations (Mg<sup>2+</sup>, Mn<sup>2+</sup>)</li> <li>• Reducing agents (DTT, 2-mercaptoethanol)</li> <li>• Cell components (lipids, nucleotides, proteins, ions, small molecules)</li> <li>• Inhibitors (protease, phosphatase, kinase)</li> </ul>	<ul style="list-style-type: none"> <li>• Homogeneous with <i>in situ</i> detection (mix-and-read)</li> <li>• Homogeneous with <i>ex situ</i> detection (separations-based methods)</li> <li>• Heterogeneous (liquid and solid phase with wash steps, e.g. ELISA)</li> <li>• Live cell and/or live tissue (genetically engineered) coupled to microscopy or flow cytometry</li> <li>• Live cell coupled to separation technique</li> <li>• Fixed, permeabilized cell coupled to cytometry</li> </ul>	<ul style="list-style-type: none"> <li>• Radiometric</li> <li>• Luminescence (chemiluminescence, fluorescence)</li> <li>• TR-FRET</li> <li>• Fluorescence Polarization</li> <li>• Absorbance</li> <li>• Mass spectrometry</li> </ul>
	<b>Label</b>			<b>Instrumentation</b>
	<ul style="list-style-type: none"> <li>• <sup>32</sup>P</li> <li>• Fluorophore</li> <li>• Biotin</li> <li>• Quantum dot</li> <li>• Lanthanide-, metal cation-, or fluorophore- doped nanoparticle</li> <li>• Anti-phospho antibody</li> </ul>			<ul style="list-style-type: none"> <li>• Plate reader (fluorescence, absorbance)</li> <li>• Spectrophotometer</li> <li>• Gel electrophoresis + Western blotting</li> <li>• HPLC/UV/Fluorescence/MS</li> <li>• CE-LIF</li> <li>• Mass Spectrometry</li> <li>• Flow cytometry</li> <li>• Mass cytometry</li> <li>• Fluorescence microscopy</li> </ul>



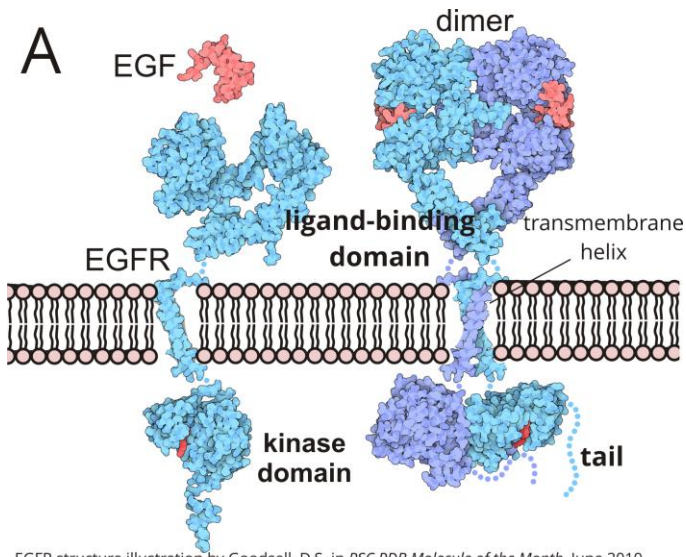
**Table 2. Assay Platforms for Measuring EGFR Tyrosine Kinase Activity**

<b>IN VITRO ASSAY PLATFORMS</b>				
<b>ASSAY</b>	<b>REFERENCES</b>	<b>CLASS</b>	<b>MECHANISM</b>	<b>APPLICATIONS/COMMENTS</b>
radiometric filter-binding	41,42	radiometric	kinase-mediated transfer of <sup>33</sup> P-phosphate to protein or peptide substrate; heterogeneous	<i>in vitro</i> ; kinetics and mechanistic studies; gold standard
HTRF	45	fluorescence/ luminescence, Ab-dependent	FRET between labeled peptide and anti-phosphotyrosine antibody; FRET signal proportional to phosphopeptide	<i>in vitro</i> ; homogeneous, mix-and-read; high throughput; popular in drug discovery setting; usually peptide substrates only due to acceptable distance for FRET
ELISA		fluorescence/ luminescence, Ab-dependent	capture antibody isolated phosphorylated substrate, secondary detection antibodies allow assessment of phosphorylation by fluorescence, chemiluminescence or UV	<i>in vitro</i> or bulk <i>in cellulo</i> ; high throughput
Z'-LYTE	50	fluorescence/ luminescence, Ab-independent	engineered peptide substrate contains FRET pair, phosphorylation site, and proteolytic cleavage site; phosphorylation protects from cleavage, leading to high FRET	<i>in vitro</i> ; high throughput; antibody-independent; substrate design is non-trivial

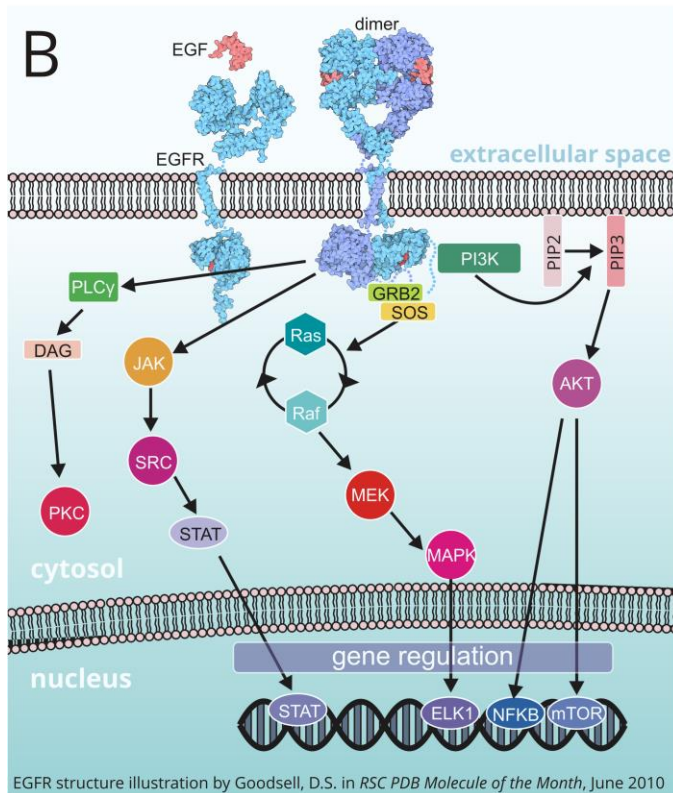
Omnia	49	fluorescence/ luminescence, Ab-independent	Chelation-Enhanced Fluorescence (CHEF); engineered peptide substrate labeled with Sox fluorophore chelates divalent cation when phosphorylated, enhancing Sox fluorescence	<i>in vitro</i> ; high throughput; progress curves; substrate design is non-trivial
Caliper Mobility Shift	51	fluorescence/ luminescence, Ab-independent	phosphorylated peptide substrate separated from non-phosphorylated substrate by differential mobility in a microfluidic chip	<i>in vitro</i> ; high throughput; progress curves; substrate design is non-trivial
<b>BULK IN CELLULO ASSAY PLATFORMS</b>				
western blot	52,53	fluorescence/ luminescence, Ab-dependent	proteins transferred to membrane, washed, and probed with primary and detection Abs; low throughput	phosphorylation of endogenous substrates
magnetic bead	55	fluorescence/ luminescence, Ab-dependent	peptide substrate appended to magnetic bead; detection with phospho-specific Ab; beads permit miniaturization higher throughput	miniaturizeable for HT applications or small numbers of cells (e.g. clinical tissue samples)

KAYAK	54	mass spectrometry	label-free; massively parallel multiplexing; high throughput; large sample consumption	proteomics/research applications; network biology
<b><i>IN CELLULOASSAYS IN SINGLE CELL LYSATES</i></b>				
SC barcode chip	56	fluorescence/ luminescence, Ab-dependent	single cell; essentially on chip ELISA	single cell version of western blot/ELISA type data; potentially applicable to clinical samples
<b><i>IN CELLULO ASSAYS IN SINGLE FIXED CELLS</i></b>				
immunohistochemistry (IHC)	58	fluorescence/ luminescence, Ab-dependent	single cell; non-quantitative; poor reproducibility; usually not for phosphoproteins	accessible to clinical pathologists, widely used and accepted for primary clinical specimens
Flow proximity ligation	57	fluorescence/ luminescence, Ab-dependent	nucleotide-labeled detection antibodies bind EGFR and/or phosphorylated EGFR; co-localized antibodies joined by cDNA after treatment with DNA ligase; amplified cDNA labeled with fluorophore; many mol of fluorophore for each mol of Ab pair results in high sensitivity	single cell; quantitative; high throughput; requires large cell numbers; suitable for clinical samples

FRET/FLIM for pEGFR	59	fluorescence/ luminescence, Ab-dependent	co-immunostaining for EGFR and pTyr with Ab labeled with FRET pair; high throughput; requires capability for FLIM imaging; complex sample prep	requires specialized imaging; tissue preparation is laborious
<b>IN CELLULO ASSAYS IN SINGLE LIVE CELLS</b>				
genetically encoded FRET sensor	60	fluorescence/ luminescence, Ab-independent	phosphorylation of peptide substrate forces intermolecular binding of SH2 domain, bringing CFP and YFP into proximity for FRET	requires genetic engineering; suitable for research purposes; provides dynamic spatial and temporal information
chemical cytometry	17-20	fluorescence/ luminescence, Ab-dependent	fluorescent peptide reporter is loaded into single cells; phosphorylated and non-phosphorylated reporter are separated by single cell CE-LIF following individual cell lysis	exquisitely sensitive; requires very few cells; suitable to multiplexing and automation

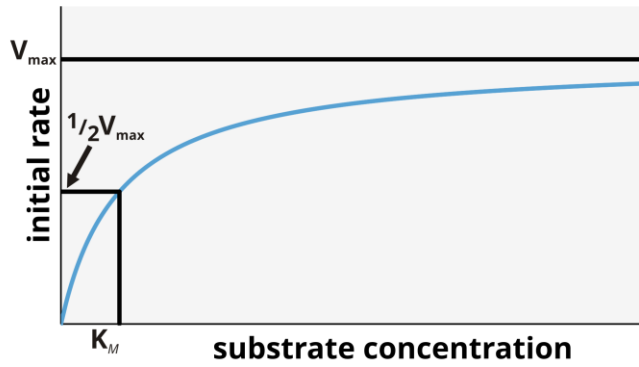


EGFR structure illustration by Goodsell, D.S. in *RSC PDB Molecule of the Month*, June 2010

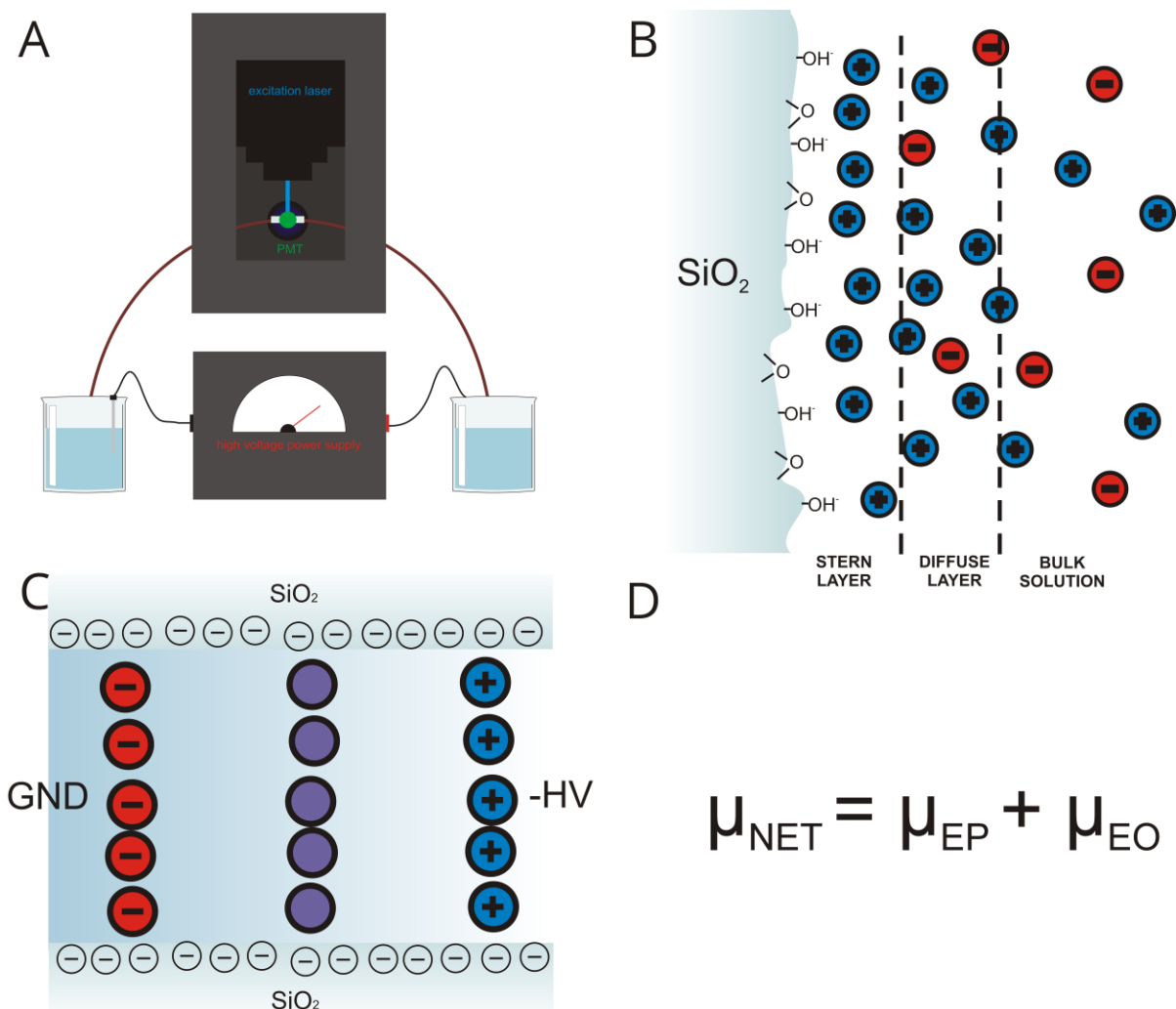


EGFR structure illustration by Goodsell, D.S. in *RSC PDB Molecule of the Month*, June 2010

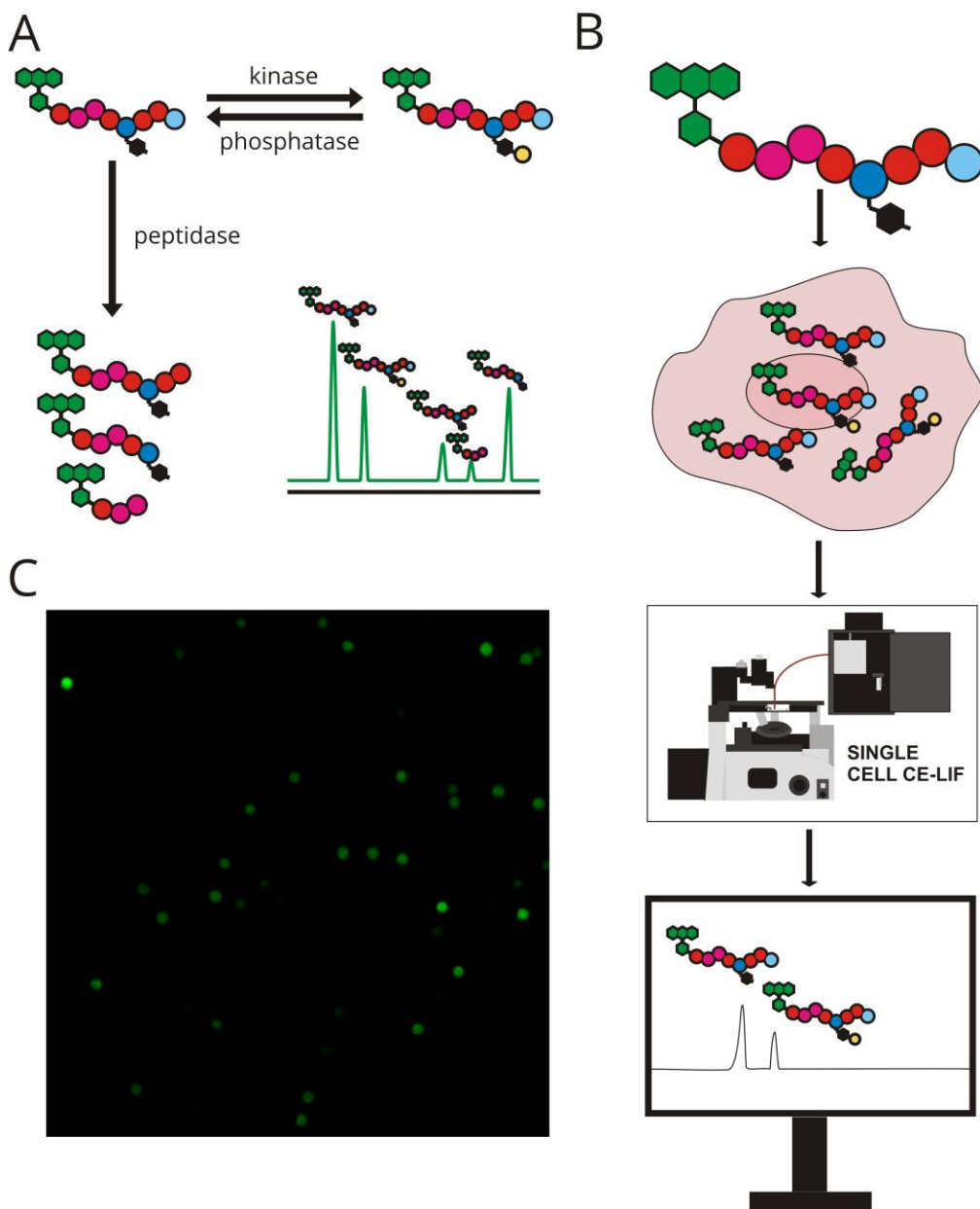
**Figure 1. Overview of EGFR biology.** (A) Structure of EGFR in its unbound, inactive state and its ligand-bound, dimerized active state. Ligand binding and dimerization induce structural changes in the intracellular kinase domain which activate EGFR kinase activity. (B) The major EGFR-dependent intracellular signaling pathways.



**Figure 2. Kinetic curve based on the Michaelis-Menten model of enzyme kinetics.** Initial reaction rate varies characteristically with substrate concentration. The  $V_{\max}$  is the maximum rate of reaction, reached when the substrate concentration equals or exceeds the concentration of available enzyme active sites. The substrate concentration at  $1/2 V_{\max}$  is defined as the Michaelis-Menten constant,  $K_M$ .



**Figure 3. Overview of capillary electrophoresis** (A) Diagram of a typical CE-LIF instrument. The ends of a buffer-filled glass capillary are immersed in buffer reservoirs connected to a high-voltage power supply. A window burned in the capillary coating is aligned with an excitation laser and a ball lens, oriented at right angles to each other, to allow excitation of fluorescent analytes and orthogonal collection of emission. (B) Schematic representation of the electrical double layer at the surface of a fused silica capillary. The anionic surface charge on the capillary wall attracts an immobile layer of cations from the BGE. Further out in solution, a mobile layer of net cationic character is responsible for bulk BGE flow in an electric field. Charge neutrality is re-established in the bulk solution. (C) Ion separation in normal mode capillary electrophoresis. Cations migrate ahead of the EOF toward the anode at the outlet, while neutral species migrate with the EOF. Anion migration opposes the EOF, so these species will only reach the detector if the electroosmotic mobility is greater than the electrophoretic mobility. (D) Net mobility in capillary zone electrophoresis is the vector sum of the electrophoretic and electroosmotic mobilities.



**Figure 4. Overview of single cell EGFR assay.** (A) Detection of EGFR reporter and its enzymatic products by CE-LIF. Optimal separation conditions allow all metabolites to be resolved from the parent reporter. (B) Schematic of the single-cell EGFR biochemical assay. Reporter is loaded into intact cells. After incubation, the cells are analyzed by single-cell CE-LIF, and reporter modification is quantified to yield a direct readout of EGFR kinase activity. (C) Fluorescence microscopy of intact cells loaded with a fluorescein-labeled EGFR reporter.



## CHAPTER 2: CONFORMATIONAL CONSTRAINT AS A SOLUTION TO PROTEIN TYROSINE PHOSPHATASE INTERFERENCE IN INTRACELLULAR PEPTIDE-BASED KINASE ASSAYS

### 2.1 Background

Protein tyrosine kinases (PTKs) propagate intracellular signaling through selective phosphorylation of regulatory tyrosine residues within protein substrates.<sup>1-3</sup> Abnormal activation of one or more PTKs represents an important phenotype in a variety of cancers, and, increasingly, clinically efficacious anti-cancer treatment regimens include some form of targeted PTK inhibition.<sup>4-6</sup> Successful treatment with targeted PTK inhibitors clearly presupposes the presence of cells whose malignancy is driven by the PTK to be inhibited. In practice, direct assessment of PTK activity in clinical specimens is a daunting analytical challenge, in particular because the inherent heterogeneity of clinical samples necessitates that measurements be taken at the level of the individual cell<sup>7</sup>. Bulk population measurements are much simpler to implement analytically, but may miss rare druggable subpopulations, ultimately resulting in disease relapse.<sup>7,8</sup> Direct measurement of PTK activity involves monitoring phosphorylation of a substrate of the PTK of interest, typically via some kind of fluorescence readout. For activity measurements in intact cells, the “reporter” substrate must either be genetically encoded and expressed in the cell, or must be

synthesized exogenously and introduced into the cytosol prior to measurement. The first approach is advantageous because endogenous protein substrates may be used as reporters, preserving native structural and spatiotemporal regulatory mechanisms.<sup>9</sup> However, genetic encoding is experimentally difficult to implement,<sup>10</sup> particularly in small, heterogeneous clinical samples, and may not permit reliable quantitative analyses in all samples.<sup>11</sup> In contrast, short phosphorylatable peptide substrates, which have traditionally been used to monitor PTK activity *in vitro*, are efficiently and cost-effectively synthesized exogenously, can be introduced into intact cells through a wide range of techniques, and are adaptable to ultrasensitive, quantitative analytical platforms.<sup>12,13</sup> Peptide reporters have been successfully coupled to ultrasensitive single-cell microelectrophoresis to quantify a variety of enzymatic activities in intact single cells, including peptide metabolism,<sup>14-17</sup> serine/threonine kinase activity,<sup>18</sup> and protein tyrosine phosphatase activity.<sup>19,20</sup> However, this platform has been difficult to adapt to PTK activity measurements due to the susceptibility of PTK peptide reporters to interference from protein tyrosine phosphatase (PTP) activity. PTPs catalyze the hydrolysis of the phosphotyrosine phosphate monoester and are some of the most efficient and fastest-acting enzymes currently known.<sup>21-23</sup> The ability of PTPs to act rapidly and comparatively indiscriminately upon phosphorylated peptide substrates often results in peptide reporters being overwhelmed by PTP activity. The practical

implication of this phenomenon is that the lifetime of phosphoreporter in biologically relevant systems (cell lysates, intact cells) is typically too short to capture experimentally. Where the phosphoreporter is the species of interest (e.g. PTK activity assays), this short lifetime is a serious analytical obstacle. While it is common to overcome this challenge in cell lysates by employing one or more chemical PTP inhibitors, this approach is not ideal for assays in intact cells because PTP inhibition is likely to disrupt the delicate signaling architecture being studied. An inhibitor-free solution to PTP interference would be ideal.

We report that substitution of a constrained tyrosine analog, 7-hydroxy-(S)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Htc), in the place of natural tyrosine within a validated EGFR kinase activity reporter reduces PTP assay interference by three orders of magnitude while retaining comparable capacity to monitor EGFR activity and response to inhibition *in vitro*. We further apply this enhanced reporter to quantify EGFR activity in single, intact A431 epidermoid carcinoma cells using ultrasensitive single cell microelectrophoresis. We demonstrate that the Htc-containing reporter, but not the native tyrosine reporter, is suitable for quantifying EGFR biochemical activity in single intact cells.

## 2.2 Materials and Methods

### 2.2.1 Peptide Synthesis and Purification

Peptides were prepared using Fmoc Solid Phase Peptide Synthesis (SPPS) protocols, either by hand or using a Protein Technologies, Inc. PS3 automated peptide synthesizer (Tucson, AZ). Standard Fmoc amino acids were purchased from NovaBiochem as the t-Butyl-protected forms where appropriate, with the exception of aspartic acid. The peptide sequences in this report were highly susceptible to aspartic acid racemization via aspartimide hydrolysis. Therefore, the OMpe-protected form of aspartic acid was employed where economically feasible.<sup>24</sup> The solid phase was chosen for increased hydrophilicity to promote peptide chain solvation.<sup>25</sup> Either Nova-PEG Amide (AnaSpec, Fremont, CA) or CLEAR Amide (Peptides International, Louisville, KY) resin was used. Before use, resin was allowed to reach room temperature and was rinsed sequentially with methanol, isopropyl alcohol, dichloromethane, and DMF to remove as much water as possible and reduce aspartic acid racemization. Fmoc removal was achieved with 20% (v/v) piperidine in DMF supplemented with 0.1 M HOBt to further combat aspartic acid racemization.<sup>26</sup> Native Fmoc-amino acids were coupled using HCTU in 0.4 M N-methylmorpholine in DMF. Complete coupling was verified using a Ninhydrin test (AnaSpec, Fremont, CA) for primary amines or a chloranil/acetaldehyde test for

secondary amines. Special attention was required for the coupling of the constrained tyrosine analog Htc (Fmoc-7-hydroxy-(S)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, Sigma, St. Louis, MO). Best results were obtained using the DIC/Oxyma coupling system in DMF, using a slight excess of DIC and Oxyma with respect to the amino acid and reacting at least 18 h. After successful Htc coupling, the hydroxyl at the 7 position of the tetrahydroisoquinoline moiety was protected by 2 h or overnight treatment with excess 2-chlorotriyl chloride in DIEA/DMF. Addition of the subsequent glutamic acid required overnight coupling with a large excess of amino acid and use of HATU/Oxyma in NMM/DMF. All peptides were labeled at the N-terminus with 6-carboxyfluorescein (Anaspec) using an overnight treatment in DIC/Oxyma in N-methylpyrrolidone. Labeled peptides were treated briefly with 20% (v/v) piperidine + 0.1 M HOBt in DMF prior to cleavage to remove fluorescein esters. Peptides were cleaved from the resin and side chain protecting groups were removed in 95% trifluoroacetic acid with 2.5% triisopropylsilane and 2.5% water as scavengers. TFA was removed via evaporation under a gentle stream of N<sub>2</sub>(g), and peptides were either isolated by precipitation in cold diethyl ether or directly reconstituted in water following TFA removal. All products were purified by reversed phase HPLC with a 2-3.33 %/min water-acetonitrile gradient. Mobile phases were supplemented with 0.1% TFA. Purified products were verified by mass spectrometry, using either ESI ionization coupled to a quadrupole mass analyzer or

by MALDI-TOF MS. Purified peptides were isolated by lyophilization and reconstituted in pH 7.5 Tris buffer prior to concentration determination by UV-Vis.

### **2.2.2 Cell Culture**

All cell lines were maintained at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub> in the recommended medium supplemented with 10% FBS, 100 mg/mL penicillin, and 100 U/mL streptomycin. A431: DMEM (Cellgro, Manassas, VA); SKBR3: McCoy's 5A (Cellgro, Manassas, VA); NCI-N87: RPMI 1640 (Cellgro, Manassas, VA). For stimulation experiments in single A431 cells, cells were starved overnight in low serum (0.2% FBS) DMEM. Immediately prior to analysis, cells were stimulated 20-30 min with 200 ng/mL EGF (Life Technologies, Carlsbad, CA) in low serum media.

### **2.2.3 *In Vitro* EGFR Assays**

Purified EGFR kinase domain was purchased from EMD Millipore (catalog # 14-431, Billerica, MA) and was used at 2-5 ng/μL in EGFR dilution buffer (20 mM MOPS, pH 7.5; 1 mM EDTA; 0.01% Brij 35; 0.1% β-mercaptoethanol; 1 mg/mL BSA; 5% glycerol). Reactions were carried out at 30 °C in kinase reaction buffer (5 mM MOPS, pH 7.2; 0.4 mM EDTA; 1 mM EGTA; 10 mM MnCl<sub>2</sub>; 4 mM MgCl<sub>2</sub>; 8 μg/mL BSA) with 1 mM ATP and the indicated concentrations of peptide. Time points were terminated by 1:5 dilution in 0.2 M HCl (aq) and stored at -20 °C until analysis by CE-LIF.

### **2.2.4 Phosphopeptide Preparation**

The phosphorylated forms of Tyr-tide and Htc-tide were prepared enzymatically using EGFR under the conditions indicated above, but with a high concentration of EGFR. Peptide phosphorylation was monitored by CE-LIF; reactions were quenched with EDTA after the peptides reached approximately 90% phosphorylation. Phosphopeptide mixtures were inactivated by incubation for 4 min at 90 °C to denature the EGFR, aliquoted at working volumes, and stored at -20 °C.

### **2.2.5 *In Vitro* PTP Assays**

Purified recombinant PTP1B and TCPTP were purchased from EMD Millipore (Billerica, MA). PTP reactions were carried out at room temperature in PTP reaction buffer (60 mM HEPES, pH 7.2; 150 mM NaCl; 1 mM EDTA; 0.17 mM DTT; 0.83% glycerol; 0.017% BSA; 0.002% Brij-35); time points were quenched, stored, and analyzed as above.

### **2.2.6 Cell Lysate Preparation**

A431 cell lysates were prepared by rinsing confluent monolayers 3X with phosphate-buffered saline (PBS, pH 7.4) and scraping the cells into fresh lysis buffer (20 mM HEPES, pH 7.4; 5 mM  $\beta$ -glycerophosphate; 200 mM NaCl; 1% NP-40; 10% glycerol with 1X Roche cOmplete EDTA-Free protease inhibitor) over ice. Cell suspensions were transferred to low-retention microcentrifuge tubes and incubated

on ice 20 min with occasional gentle mixing, followed by a 15 min clarification at 14000 rcf (4 °C). Supernatants were reserved, aliquoted at working volumes, and stored up to 6 months at -20 °C. The total protein content of the lysates was determined using a fluorescamine assay.<sup>17</sup>

### **2.2.7 Lysate Phosphorylation Assays**

To measure reporter phosphorylation in cell lysates, A431 lysates were thawed on ice and diluted to 4 mg/mL (4X) in fresh lysis buffer with 0.1 mM sodium pervanadate prior to mixing with kinase reaction buffer and peptide reporter (final concentration 10 µM). Lysates were treated with 10 µM EGF for 5 min prior to assay initiation. Assays were initiated by adding ATP to a final concentration of 1 mM. For inhibitor studies, lysates were pre-treated 10 min with indicated concentrations of gefitinib or lapatinib, followed by stimulation and assay initiation. All assays were performed at 37 °C; time points were collected, quenched, and analyzed as above.

### **2.2.8 Lysate Dephosphorylation Assays**

Stock cell lysates were thawed on ice and diluted to 10 mg/mL with fresh lysis buffer. 10 mg/mL lysates were diluted to 1.25 mg/mL (1.25X) in PTP reaction buffer (above) supplemented with 3.3X Sigma protease inhibitor cocktail in DMSO and with or without 150 µM fresh sodium pervanadate. Lysates were incubated 20 min on ice with inhibitors prior to addition of phosphoreporter to a final



concentration of 10  $\mu$ M and incubation at 37 °C. Time points were collected, quenched, and analyzed as indicated above.

### **2.2.9 Lysate Peptidolysis Assays**

To quantify reporter degradation in cell lysates, confluent monolayers of A431 cells were harvested with trypsin, pelleted, washed 2X with PBS, resuspended in a small volume of PBS and lysed by multiple rounds of freezing in liquid nitrogen and rapidly thawing at 37 °C. Lysate concentrations were determined as indicated above and diluted to a final concentration of 18 mg/mL in PBS prior to addition of 10  $\mu$ M reporter to initiate the reaction. Reactions were carried out at 37 °C and quenched as above.

### **2.2.10 Capillary Electrophoresis**

Capillary electrophoretic analyses were performed using commercial Beckman Coulter (Pasadena, CA) automated PA800 CE-LIF instruments outfitted with 488 nm solid state lasers (Coherent, Santa Clara, CA). Separations were carried out in 30  $\mu$ m internal diameter bare silica capillaries filled with electrophoretic buffer (0.2 M boric acid/NaOH, pH 9.5) under a 600 V/cm applied electric field. Prior to analysis, samples were diluted to 100-200 nM in 50% electrophoretic buffer/50% water to promote sample stacking.

### 2.2.11 *In Vitro* Data Analysis

Electropherograms were integrated using commercial software (32Karat, Beckman Coulter, Pasadena, CA). For kinase assays, percent phosphorylation was calculated from the corrected area under the phosphoreporter peak divided by the total reporter corrected area multiplied by 100%. For phosphatase assays, the percent phosphorylation was normalized to the initial percent phosphorylation at time zero. Percent dephosphorylation was calculated as the normalized percent phosphorylation at a given time subtracted from the normalized percent phosphorylation at time zero. For phosphorylation and dephosphorylation in cell lysates, values are given with respect to intact reporter.

To estimate Michaelis-Menten kinetic constants  $K_M$  and  $k_{cat}$ , progress curves of reporter phosphorylation by EGFR or of phosphoreporter dephosphorylation by PTP1B or TCPTP were generated by treating as above with recombinant enzyme followed by CE-LIF. Progress curves were fit to the time-dependent Michaelis-Menten equation (Eq 1) using an analytical approximation of the Lambert function.<sup>27</sup> Curves at 2-3 substrate concentrations were fit globally in OriginLab (Northampton, MA). In the case of EGFR, multiple replicates were fit globally, as this resulted in better fits than weighted fitting of averaged data with standard deviations. In all cases, the standard errors reported are those generated by OriginLab's non-linear regression tool. Fits and residual plots are available in the supporting information.

The  $K_M$  and  $k_{cat}$  for pHtc-tide dephosphorylation by recombinant phosphatases was measured by traditional initial rates analysis due to economic constraints on enzyme consumption. A plot of initial rate versus substrate concentration was fit to the Michaelis-Menten equation using OriginLab's nonlinear regression tool.

### **2.2.12 Pinocytic Loading for Single Cell EGFR Activity Assays**

Monolayers of A431 cells were serum starved overnight in DMEM medium containing 0.2% serum. Immediately before pinocytic loading with peptide reporter, the monolayers were stimulated 10 min with 200 ng/mL EGF to initiate high levels of EGFR activity. Monolayers were then disadhered using 0.05% trypsin with EDTA and suspended in complete medium. Approximately  $1 \times 10^6$  cells were isolated in a 1.5 mL tube and loaded with Htc-tide or Tyr-tide by pinocytosis. The cells were incubated for 10 min with hypertonic loading solution (Influx™, Life Technologies) containing 100  $\mu$ M peptide reporter (20  $\mu$ L per  $10^6$  cells), followed by pinosome lysis in hypotonic DMEM to release reporter into the cytosol and initiate the EGFR biochemical assay. Cells were washed once with complete medium and immediately plated on custom glass chambers coated with CellTak™ (Corning, Corning, NY). Viability, assessed by trypan blue staining immediately after loading, was 81%. Cells were plated at a low density resulting in approximately 30 adherent cells/mm<sup>2</sup> after rinsing. Cells were allowed to adhere at 37 °C for approximately 35 min, followed by single cell analysis.

### **2.2.12 Single Cell Analysis with Microelectrophoresis**

Single cell microelectrophoresis was performed as previously described.<sup>17-20</sup> Cell chambers containing loaded A431 cells adhered to CellTak were placed on the stage of an inverted microscope. Extracellular buffer (33-37 °C) was flowed over the cells for several minutes to remove weakly adhered or non-adherent cells. Cells were chosen for individual analysis based on acceptable morphology (round shape, intact membrane, lack of blebbing) and absence of other cells in the field of view. Selected cells were lysed individually by a rapid pulse from an Nd:YAG laser as previously described, then the contents electrokinetically loaded into the lumen of a 30  $\mu\text{m}$  i.d. fused silica capillary positioned directly above the cell. The cellular contents including unmodified and modified peptide reporter were separated by capillary electrophoresis under a field strength of 200 V/cm and detected by laser-induced fluorescence. The optimal separation buffer for single cell analysis was 0.2 M boric acid/NaOH with 2 mM SDS, pH 9.5. The number of moles of reporter loaded into each cell was calculated based on the response from a known standard run at the beginning of each day.

## **2.3 Results and Discussion**

### **2.3.1 Mitigation of Phosphoreporter Dephosphorylation *in vitro* via Conformational Constraint**

In previous efforts to develop native tyrosine-containing peptide reporters for single-cell PTK assays, our lab has consistently observed the rapid

dephosphorylation of these peptides upon phosphorylation, both *in vitro* and *in vivo*. In all cases, we observed that phosphoreporter dephosphorylation occurs on a significantly shorter time scale than phosphorylation of the corresponding tyrosine reporter (Table 3). *In vitro* dephosphorylation of the phosphorylated EGFR kinase reporter pTyr-tide (6FAM-EDDEYEEV-NH<sub>2</sub>) by recombinant TCPTP occurs with a rate constant  $k_{\text{cat}}$  168X greater than that of Tyr-tide phosphorylation by EGFR. Based on the observation that introduction of conformational constraint to a peptide substrate can impede its interaction with certain enzymes,<sup>28-31</sup> we hypothesized that substitution of the native tyrosine (Figure 5A) within the reporter sequence with a conformationally constrained, phosphorylatable analog would slow dephosphorylation by PTPs, increasing phosphoreporter lifetime in the cytosol to measureable levels. The constrained tyrosine analog L-Htc (Figure 5B) has previously been included in an efficient substrate of spleen tyrosine kinase (SYK) and is variably tolerated by other PTKs.<sup>28,30</sup> We observed that substitution of L-Htc into our EGFR reporter Tyr-tide resulted in comparable phosphorylation by EGFR (Table 3). Similar to tyrosine, the Htc molecule contains a phosphorylatable hydroxyl positioned on a rigid conjugated 6-member ring. In contrast to tyrosine, however, this ring is cyclized via a methylene bridge to the  $\alpha$ -nitrogen, forming a fused bicyclic system with limited flexibility. NMR studies of L-Htc in the context of similar peptides have shown that this residue can adopt either a gauche (+) or

gauche (-) conformation about the  $\chi_1$  torsional angle formed by the  $N\alpha-C\alpha-C\beta-C\gamma$ , with the gauche (+) being preferred.<sup>28</sup> This cyclization severely restricts the flexibility of the entire residue, limiting the conformations that the peptide is able to adopt within the enzyme active site. To test the hypothesis that decreased flexibility of the phosphorylated reporter would discourage dephosphorylation by PTPs, we performed a series of *in vitro* dephosphorylation assays with the model PTPs PTP1B and TCPTP, which have been shown to preferentially recognize highly acidic peptide substrates such as pTyr-tide.<sup>32,33</sup> The data in Figure 5C illustrate the different timescales of dephosphorylation of pTyr-tide and pHtc-tide (6FAM-EDDE-**Htc**-EEV-NH<sub>2</sub>) by PTP1B at a range of enzyme concentrations. We observed very little dephosphorylation of pHtc-tide, even at the highest enzyme concentration. In contrast, pTyr-tide is fully dephosphorylated within 5 min at this enzyme concentration. This differential susceptibility to PTP interference is further illustrated in Figure 5D and E. Phosphoreporter dephosphorylation in 10 min was monitored as a function of increasing PTP1B or TCPTP concentration. While it is possible by using large quantities of either enzyme to force dephosphorylation of pHtc-tide, it is significantly less susceptible to dephosphorylation compared with the native phosphotyrosine peptide.

### 2.3.2 Phosphoreporter Lifetime in Epithelial Tumor Lysates

We next verified that this phenomenon was not limited to the two recombinant PTPs tested by comparing dephosphorylation of pHtc-tide and pTyr-tide in lysates of a panel of model tumor lines overexpressing one or more of the EGFR family kinases. A431 cells, derived from an epidermoid carcinoma of the vulva and overexpressing full length wild type EGFR, have previously been used to study PTP1B activity in cell lysates.<sup>34,35</sup> The data in Figure 6A clearly demonstrate that, while pTyr-tide is completely dephosphorylated within 5 min, approximately 90% of pHtc-tide remains even after 30 min. We verified that the decrease in phosphoreporter in both cases was due to dephosphorylation and not to peptidolysis or another competing process by pre-treating the lysates with 0.1 mM sodium pervanadate, a pan-PTP inhibitor that irreversibly oxidizes the catalytic cysteine within the PTP active site.<sup>36</sup> Loss of both reporters was nearly eliminated upon pervanadate treatment, confirming that this phenomenon is due to PTP activity. The resistance of pHtc-tide to dephosphorylation is not limited to A431 lysates. We observed similar rates of dephosphorylation of both pTyr-tide and pHtc-tide in A431 epidermoid carcinoma, SKBR3 breast carcinoma, and NCI-N87 gastric carcinoma cell lysates (Figure 6B), suggesting that Htc-tide can be generally applied to measure EGFR kinase activity in cell lysates.

### 2.3.3 Kinetics of Phosphoreporter Dephosphorylation

We further investigated the mechanism of Htc reporter resistance to PTP interference by comparing the kinetics of dephosphorylation of the two phosphoreporters by recombinant PTP1B and TCPTP. The interaction of both phosphoreporters with PTP1B and TCPTP can be modeled according to the standard Michaelis-Menten model of enzyme kinetics.<sup>32,37</sup> We used a combination of progress curve analysis and initial rate measurements to estimate the kinetic parameters  $K_M$  and  $k_{cat}$  from dephosphorylation time courses. For both PTPs studied, the  $k_{cat}$  for pTyr-tide dephosphorylation was 3 orders of magnitude larger than the corresponding value for pHtc-tide. Interestingly, the  $K_M$  values for pTyr-tide and pHtc-tide were very similar for both PTPs, suggesting that the differential dephosphorylation observed is primarily driven by differences in the rate of phosphate monoester hydrolysis, rather than in the ability of the two reporters to bind to the phosphatase active site. Examining the vacuum electrostatics of the substrate binding surfaces of published X-Ray crystal structure of PTP1B<sup>38</sup> in PyMol<sup>39</sup> indicates large patches of enriched positive charge near the phosphate monoester binding pocket (Figure 7). We posit that the overwhelming acidity of the peptide reporters drives electrostatic interactions with the substrate binding regions of the phosphatases; this electrostatic affinity is reflected in the low measured  $K_M$  values. This hypothesis is supported by recent work demonstrating the



overwhelming preference of common PTPs for acidic substrates.<sup>33</sup> We speculate that the introduction of conformational constraint into the backbone of pHtc-tide positions the phosphate monoester unfavorably with respect to the conserved PTP catalytic residues, thus significantly decreasing the rate of phosphate monoester hydrolysis. The ability of EGFR to recognize Htc-tide may be a function of a more balanced surface charge distribution near the substrate binding cleft of the EGFR KD<sup>38</sup> (Figure 7A), resulting in a greater number of accessible binding modes for Htc-tide and a higher probability of catalytically favorable reporter binding. X-Ray crystallographic or NMR studies would be required to verify this hypothesis and may be an interesting avenue for future research.

### **2.3.4 Increased Resistance to Peptidolysis in Cell Lysates**

It has been previously reported that modification of short peptides with unnatural amino acids discourages peptidolysis and increases intact peptide lifetime in cell lysates and intact cells.<sup>31,40-42</sup> We hypothesized that substitution of tyrosine with L-Htc would confer the additional advantage of improved stability in the face of peptidolysis in cell lysates. To test this possibility, we first synthesized standards of the 6FAM-labeled fragments of Tyr-tide and verified that no fragments co-migrated with intact unmodified or phosphorylated Tyr-tide or Htc-tide. Tyr-tide and Htc-tide share four fragments N-terminal to the phosphorylation site. The three Htc-containing fragments of Htc-tide were not synthesized because similarities in

electrophoretic mobility between Htc-tide and Tyr-tide allow us to infer approximate migration times for the remaining Htc-containing fragments. Given the excellent separation between the intact reporters and all fragments (Figure 8A), it is improbable that the Htc-containing fragments would co-migrate with either intact reporter. Having demonstrated the ability to separate intact reporter from its metabolic products by CE-LIF, we evaluated breakdown of these reporters in high-concentration, inhibitor-free A431 lysates (Figure 8B). As expected, replacement of natural tyrosine with the constrained analog increased the stability of the reporter almost 2-fold, increasing the half-life of intact reporter from 193 ( $\pm$  53) min to 328 ( $\pm$  87) min in 18 mg/mL cell lysates.

### 2.3.5 Quantifying Response to EGFR Inhibition

Having demonstrated comparable *in vitro* phosphorylation of Tyr-tide and Htc-tide by recombinant EGFR kinase domain, we proceeded to compare the utility of the two reporters to monitor EGFR activity in a more biologically relevant system. We therefore explored the capacity of Htc-tide and Tyr-tide to act as reporters of EGFR kinase activity in A431 cell lysates (Figure 9). We monitored reporter phosphorylation in lysates of A431 cells that had been serum starved overnight, treated with EGFR TKI, and stimulated with the EGFR ligand EGF immediately prior to assay initiation. The reporters respond equivalently to selective EGFR

perturbation, indicating that Htc-tide may be used in place of Tyr-tide to monitor EGFR activity *in cellulo*.

### 2.3.6 Single Cell Analysis

The ultimate goal of this work is to develop a peptide reporter suitable for single cell analysis of EGFR biochemical activity in intact cells. To that end, the compatibility of Htc-tide versus Tyr-tide for single cell EGFR kinase assays was assessed (Figure 10). Either Htc-tide or Tyr-tide was loaded by pinocytosis into serum-starved, stimulated A431 cells and incubated for 35-45 min at 37 °C. Individual loaded cells were lysed and analyzed by single cell microelectrophoresis as previously described.<sup>17-20</sup> Both reporters were successfully loaded into the cytosol, as confirmed by fluorescence microscopy (data not shown). In keeping with previous work, both reporters underwent variable peptidolysis. Figure 10B shows a linear increase in intracellular Htc-tide metabolism with the amount of Htc-tide loaded into the cells. Consistent with lysate data, Htc-tide showed marginally better resistance to peptidolysis compared with Tyr-tide. This improved stability is likely due to the increased rigidity of the peptide backbone of Htc-tide compared with Tyr-tide.

While Tyr-tide was successfully loaded into cells and underwent expected levels of peptidolysis, it was not detectably phosphorylated in any of the cells analyzed (Figure 10A). This result is in keeping with previous observations of rapid

dephosphorylation of phosphotyrosine peptides in intact single cells.<sup>19,20</sup> In contrast, Htc-tide was measurably phosphorylated in all of the cells analyzed, indicating excellent resistance to assay interference from PTPs. Quantification of Htc-tide phosphorylation in single A431 cells shows a linear increase in moles of pHtc-tide detected in response to increased total Htc-tide loaded (Figure 10B). This is consistent with estimated intracellular Htc-tide concentrations (0.001-0.1  $\mu\text{M}$ ) well below the *in vitro* measured  $K_M$  of 18  $\mu\text{M}$ .

## 2.4 Conclusions

This work presents a novel solution to protein tyrosine phosphatase interference in peptide reporter-based intracellular EGFR kinase assays whose efficacy does not depend on use of inhibitors or any other intervention beyond chemical modification of the peptide reporter. Simply replacing natural tyrosine with the conformationally constrained analog Htc sufficiently altered reporter recognition by PTPs to enable intracellular reporter phosphorylation. This approach to mitigating assay interference is minimally invasive to the cell. To the best of our knowledge, the data presented here comprise the first examples of direct measurement of tyrosine kinase activity in intact single cells using a short synthetic peptide reporter. Future work will explore the broader applicability of conformationally constrained tyrosine analogs in reducing PTP interference in a range of intracellular PTK assays. Conformational constraint of the reporter may

prove to be a general solution to assay interference for chemical cytometry of PTK activity.

## REFERENCES

- (1) Lemmon, M. A.; Schlessinger, J. *Cell* **2010**, *141*, 1117–1134.
- (2) Bose, R.; Holbert, M. A.; Pickin, K. A.; Cole, P. A. *Current Opinion in Structural Biology* **2006**, *16*, 668–675.
- (3) Blume-Jensen, P.; Hunter, T. *Nature* **2001**, *411*, 355–365.
- (4) Baselga, J. *Science* **2006**, *312*, 1175–1178.
- (5) Krause, D. S.; Van Etten, R. A. *N. Engl. J. Med.* **2005**, *353*, 172–187.
- (6) Levitzki, A. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 161–185.
- (7) Almendro, V.; Marusyk, A.; Polyak, K. *Annu. Rev. Pathol. Mech. Dis.* **2013**, *8*, 277–302.
- (8) Bhatia, S.; Frangioni, J. V.; Hoffman, R. M.; Iafrate, A. J.; Polyak, K. *Nature Biotechnology* **2012**, *30*, 604–610.
- (9) Sample, V.; Mehta, S.; Zhang, J. *Journal of Cell Science* **2014**, *127*, 1151–1160.
- (10) Lu, S.; Wang, Y. *Clinical Cancer Research* **2010**, *16*, 3822–3824.
- (11) Hertel, F.; Zhang, J. *Biopolymers* **2013**, *101*, 180–187.
- (12) Rubakhin, S. S.; Romanova, E. V.; Nemes, P.; Sweedler, J. V. *Nat Meth* **2011**, *8*, S20–S29.
- (13) Borland, L. M.; Kottegoda, S.; Phillips, K. S.; Allbritton, N. L. *Annual Review of Analytical Chemistry* **2008**, *1*, 191–227.
- (14) Brown, R. B.; Hewel, J. A.; Emili, A.; Audet, J. *Cytometry* **2010**, *77A*, 347–355.
- (15) Kovarik, M. L.; Shah, P. K.; Armistead, P. M.; Allbritton, N. L. *Anal. Chem.* **2013**, *85*, 4991–4997.
- (16) Kovarik, M. L.; Dickinson, A. J.; Roy, P.; Poonnen, R. A.; Fine, J. P.; Allbritton,

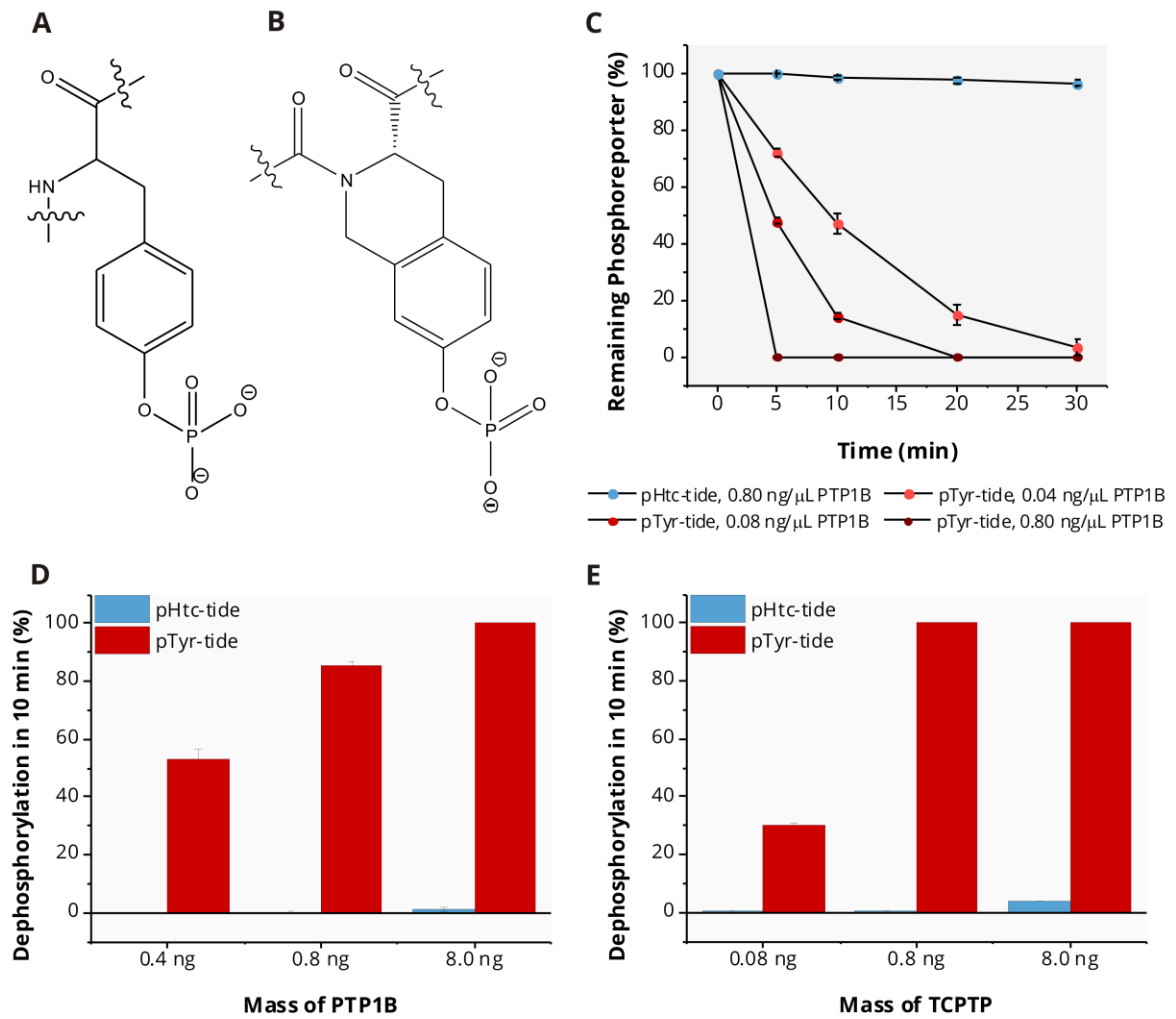
- N. L. *Integr Biol (Camb)* **2014**, *6*, 164–174.
- (17) Proctor, A.; Wang, Q.; Lawrence, D. S.; Allbritton, N. L. *Analyst* **2012**, *137*, 3028–3038.
- (18) Proctor, A.; Herrera-Loeza, S. G.; Wang, Q.; Lawrence, D. S.; Yeh, J. J.; Allbritton, N. L. *Anal. Chem.* **2014**, *86*, 4573–4580.
- (19) Phillips, R. M.; Bair, E.; Lawrence, D. S.; Sims, C. E.; al, E. *Anal. Chem.* **2013**, *85*, 6136–6142.
- (20) Phillips, R. M.; Dailey, L. A.; Bair, E.; Samet, J. M.; Allbritton, N. L. *Anal. Chem.* **2014**, *86*, 1291–1297.
- (21) Lad, C.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5607–5610.
- (22) Zhang, Z. Y. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 1–52.
- (23) Zhang, Z.-Y.; Wang, Y.; Dixon, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1624–1627.
- (24) Karlström, A.; Undén, A. *Tetrahedron Letters* **1996**.
- (25) Kempe, M.; Barany, G. *J. Am. Chem. Soc.* **1996**, *118*, 7083–7093.
- (26) Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W. *Org. Lett.* **2012**, *14*, 5218–5221.
- (27) Goličnik, M. *Analytical Biochemistry* **2010**, *406*, 94–96.
- (28) Ruzza, P.; Calderan, A.; Donella-Deana, A.; Biondi, B.; Cesaro, L.; Osler, A.; Elardo, S.; Guiotto, A.; Pinna, L. A.; Borin, G. *Biopolymers* **2003**, *71*, 478–488.
- (29) Ruzza, P.; Cesaro, L.; Tourwé, D.; Calderan, A.; Biondi, B.; Maes, V.; Menegazzo, I.; Osler, A.; Rubini, C.; Guiotto, A.; Pinna, L. A.; Borin, G.; Donella-Deana, A. *J. Med. Chem.* **2006**, *49*, 1916–1924.
- (30) Donella-Deana, A.; Ruzza, P.; Cesaro, L.; Brunati, A. M.; Calderan, A.; Borin, G.; Pinna, L. A. *FEBS Lett.* **2002**, *523*, 48–52.

- (31) Proctor, A.; Wang, Q.; Lawrence, D. S.; Allbritton, N. L. *Anal. Chem.* **2012**, *84*, 7195–7202.
- (32) Zhang, Z.-Y.; Thieme-Seffler, A. M.; Maclean, D.; McNamara, D. J.; Dobrusin, E. M.; Sawyer, T. K.; Dixon, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4446–4450.
- (33) Selner, N. G.; Luechapanichkul, R.; Chen, X.; Neel, B. G.; Zhang, Z.-Y.; Knapp, S.; Bell, C. E.; Pei, D. *Biochemistry* **2014**, *53*, 397–412.
- (34) Lee, S. R. *Journal of Biological Chemistry* **1998**, *273*, 15366–15372.
- (35) Lou, Y.-W.; Chen, Y.-Y.; Hsu, S.-F.; Chen, R.-K.; Lee, C.-L.; Khoo, K.-H.; Tonks, N. K.; Meng, T.-C. *FEBS Journal* **2007**, *275*, 69–88.
- (36) Huyer, G.; Liu, S.; Kelly, J.; Moffat, J.; Payette, P.; Kennedy, B.; Tsaprailis, G.; Gresser, M. J.; Ramachandran, C. *Journal of Biological Chemistry* **1997**, *272*, 843–851.
- (37) Johnson, K. A. *FEBS Lett.* **2013**, *587*, 2753–2766.
- (38) Sarmiento, M.; Puius, Y. A.; Vetter, S. W.; Keng, Y.-F.; Wu, L.; Zhao, Y.; Lawrence, D. S.; Almo, S. C.; Zhang, Z.-Y. *Biochemistry* **2000**, *39*, 8171–8179.
- (39) The PyMOL Molecular Graphics System.
- (40) Pappa, E. V.; Zompra, A. A.; Spyrali, Z.; Diamantopoulou, Z.; Pairas, G.; Lamari, F. N.; Katsoris, P.; Spyroulias, G. A.; Cordopatis, P. *Biopolymers* **2011**, *96*, 260–272.
- (41) Banerjee, J.; Gyanda, R.; Chang, Y.-P.; Armishaw, C. J. *Methods Mol. Biol.* **2013**, *1081*, 13–34.
- (42) Gentilucci, L.; De Marco, R.; Cerisoli, L. *Curr. Pharm. Des.* **2010**, *16*, 3185–3203.

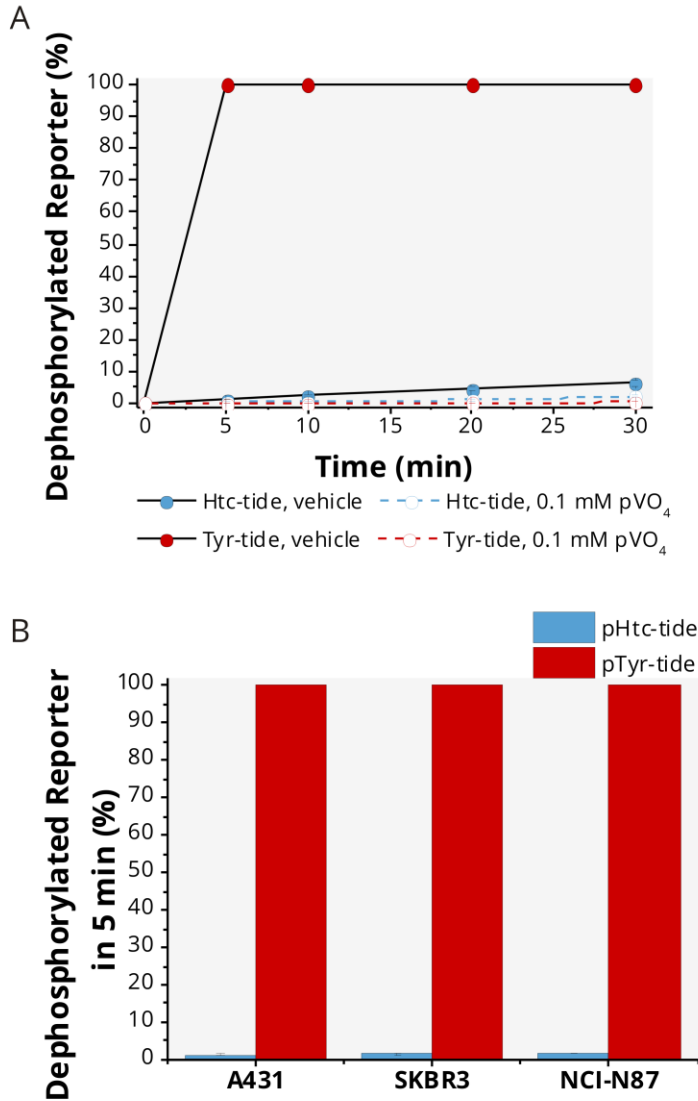


**Table 3. *In vitro* Kinetic Parameters**

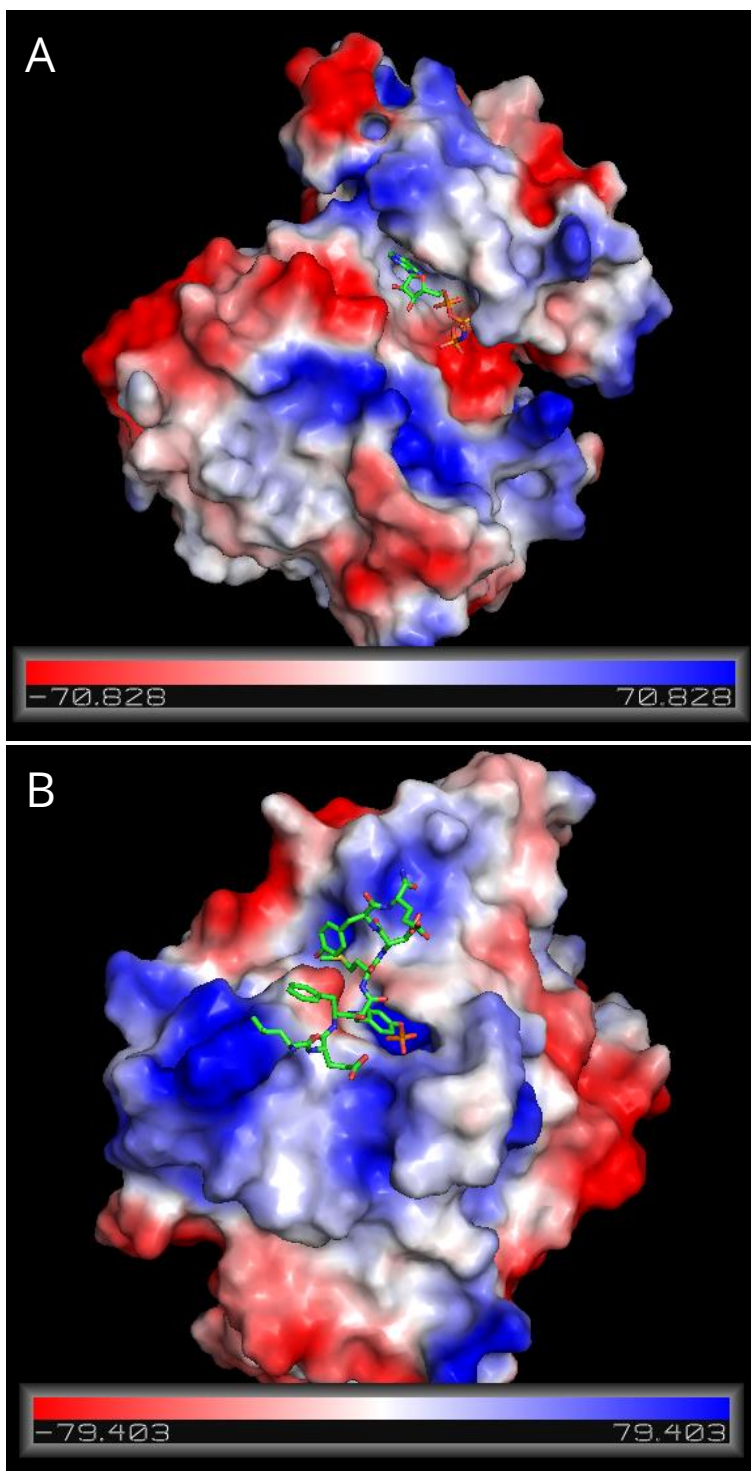
<b>EGFR</b>			
Reporter	$K_M$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $\times 10^{-3}$ )
Htc-tide	18.3 ( $\pm 2.2$ )	0.205 ( $\pm 0.006$ )	11.2
Tyr-tide	40.2 ( $\pm 4.9$ )	0.322 ( $\pm 0.016$ )	8.0
<b>PTP1B</b>			
Htc-tide	11.1 ( $\pm 0.5$ )	0.026 ( $\pm 0.001$ )	2.3
Tyr-tide	3.1 ( $\pm 0.2$ )	10.8 ( $\pm 0.2$ )	3500
<b>TCPTP</b>			
Htc-tide	12.2 ( $\pm 0.9$ )	0.054 ( $\pm 0.003$ )	4.4
Tyr-tide	2.8 ( $\pm 0.3$ )	54.1 ( $\pm 1.7$ )	19000
<b><i>In Cellulo</i> Proteolysis</b>		Htc-tide	Tyr-tide
$t_{1/2}$ (min)	328 ( $\pm 87$ )	193 ( $\pm 53$ )	



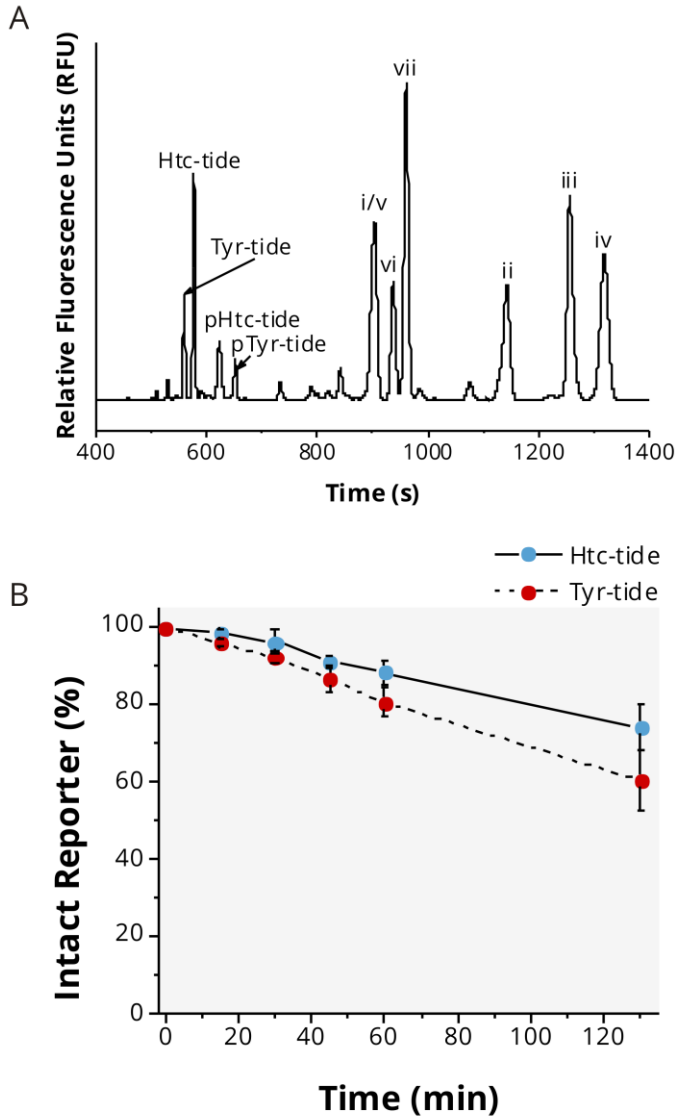
**Figure 5. Differential *in vitro* dephosphorylation of EGFR reporters.** A) Line structure of phosphorylated tyrosine. B) Line structure of phosphorylated L-Htc. C) Time course of phosphorylated EGFR reporter dephosphorylation by PTP1B at a range of concentrations. D)-E) Bar graphs comparing phosphoreporter dephosphorylation as a function of increasing PTP1B (D) or TCPTP (E). N = 3.



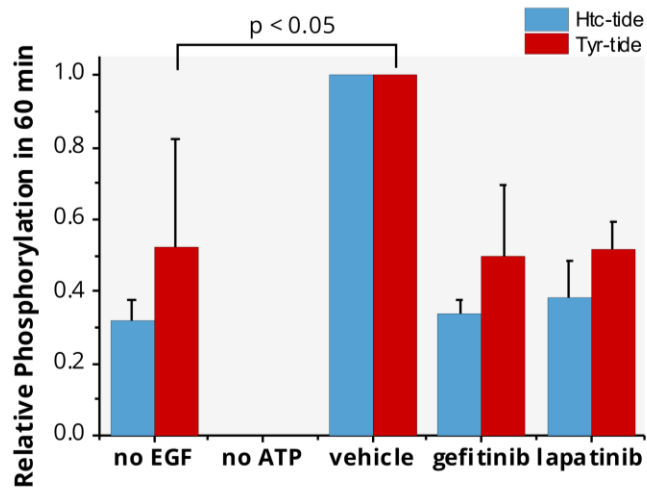
**Figure 6. In cellulo dephosphorylation of phosphorylated EGFR reporters.** A) Dephosphorylation of phosphorylated Htc-tide (blue) or Tyr-tide (red) in A431 lysates with and without pervanadate treatment. B) Dephosphorylation of the phosphoreporters is independent of cell line. A431: epidermoid carcinoma; SKBR3: breast carcinoma; NCI-N87: gastric carcinoma. N = 3.



**Figure 7. Vacuum electrostatic potential maps of EGFR kinase domain (A) and PTP1B (B) bound to substrates.** The structure in (A) was obtained from the NCBI structure database and first published in Yun, *et al.* (PDB 2ITX). The structure in (B) was obtained from the same database and was published in Sarmiento, *et al.* (PDB 1EEO).



**Figure 8. Assessment of reporter proteolysis in cell lysates.** A) Standard separation of Htc-tide, Tyr-tide, and their phosphoforms from the fragments of Tyr-tide. Htc-tide shares four fragments with Tyr-tide (i-iv). i: 6FAM-E-OH; ii: 6FAM-ED-OH; iii: 6FAM-EDD-OH; iv: 6FAM-EDDE-OH; v: 6FAM-EDDEY-OH; vi: 6FAM-EDDEYE-OH; vii: 6FAM-EDDEYEE-OH. B) Reporter breakdown in A431 lysates as a function of time. N = 3.



**Figure 9. EGFR reporter phosphorylation in A431 cell lysates under perturbation.** Both Htc-tide and Tyr-tide phosphorylation faithfully report EGF-induced kinase activity stimulation and EGFR TKI-induced inhibition. N = 3.



## CHAPTER 3: TOWARD HIGH THROUGHPUT EGFR ACTIVITY MEASUREMENTS IN SINGLE TUMOR CELLS

### 3.1 Introduction

In previous chapters the necessity of measuring oncogenic EGFR kinase activity in single intact cells has been extensively discussed. This chapter focuses on validation of the Htc-tide-based single cell EGFR assay developed in Chapter 2, with the aim of adapting the assay to an automated CE-LIF system designed for high throughput serial analysis of single cells. An overarching goal of this dissertation is to create a tool which can be used to explore intercellular heterogeneity in EGFR signaling. Intercellular heterogeneity is a well-established feature of cancer<sup>1,2</sup> and has been proposed as a mechanism underlying disease relapse in cancer patients treated with targeted kinase inhibitors.<sup>3</sup> Tumor cell populations are rapidly evolving in response to Darwinian selection pressures comprising their local microenvironment.<sup>1,4-6</sup> Because the microenvironment of, and therefore the selection pressures experienced by, each cell are unique, intratumoral heterogeneity can be extremely high.

The practical consequence of intratumoral, or intercellular, heterogeneity from a clinical standpoint, is that a tumor may contain subpopulations with a wide



range of responses to targeted drugs.<sup>7,8</sup> The ability to quantify drug target activity in each cell from a tumor biopsy would provide clinicians with a holistic view of aberrant signaling across individual tumors, thus enabling design of more efficacious treatment regimens targeting each subpopulation within the tumor. Statistically robust detection of intratumoral subpopulations is to some extent limited by the number of cells that can be accurately analyzed. Importantly, this number depends not only upon the number of usable cells present in the biopsy sample, but also on the throughput of the single cell analysis platform used.

The Htc-tide-based single cell EGFR assay described in the previous chapter can be adapted to high throughput serial analysis on automated instrumentation developed in the Allbritton lab.<sup>9,10</sup> The highest throughput achieved on this system to date is 3.5 cells/min (210 cells/h), separating fluorescent dyes loaded into non-adherent cells.<sup>9</sup> This chapter details the *in cellulo* validation of the Htc-tide assay and preliminary work on its adaptation to the high throughput serial analysis system.

A three-dimensional (3D) cell culture model was chosen as a model system for assay validation. 3D culture is a convenient research tool that more closely mimics *in vivo* tumor structure and nutrient gradients compared with traditional 2D culture systems, while retaining the convenience and ready availability of 2D systems.<sup>11,12</sup> 3D cultures, or spheroids, generally are composed of

three regions based on nutrient and oxygen permeation across the spheroid diameter.<sup>13</sup> In the outer region, where cells have most access to nutrients, cells are proliferative. As growth factor penetration decreases, cells become quiescent. The core of the spheroid is inaccessible to nutrients, resulting in cell necrosis.<sup>11,13</sup> Interestingly, differences have been reported in EGFR-family kinase activity and response to targeted inhibitors in 3D compared with 2D cultures. Luca and colleagues surveyed EGFR expression and pathway activation in a series of colorectal cancer models, and showed that decreased EGFR expression in 3D culture models using quantitative RT-PCR and immunoblotting.<sup>14</sup> Weigelt, *et al.* found that 3D cultures of breast cancer cell lines overexpressing the EGFR family kinases responded differently to anti-EGFR family kinase inhibitors compared with 2D cultures.<sup>15</sup> Importantly, the authors observed differential sensitivity of 3D cultures for monoclonal antibody inhibitors versus the small molecule TKI lapatinib, which targets EGFR and its relative HER2. We explore the capacity of the Htc-tide-based EGFR assay to report differences in EGFR activity in 2D and 3D cultures of the EGFR-overexpressing A431 cell line, derived from an epidermoid carcinoma of the vulva.

## **3.2 Methods**

### **3.2.1 Materials**

Buffers and additives were purchased from Sigma Aldrich (St. Louis,

MO), Fisher Scientific (Hampton, NH), or AMRESCO (Solon, OH) unless otherwise indicated. Primary antibodies were purchased from Cell Signaling Technology (Danvers, MA). Primary rabbit-anti-human antibodies used were anti-EGFR (D38B1, #4267), anti-pEGFR (Y1068, #2234), and anti-GAPDH (#5174). Secondary goat anti-rabbit antibody was from GE (Fairfield, CT). Gels were purchased from BioRad (Hercules, CA). The A431 cell line was purchased from the ATCC (Manassas, VA). Influx™ pinocytic loading reagent was from Life Technologies (Carlsbad, CA).

### **3.2.2 2D and 3D A431 Cell Culture and Treatment**

2D A431 cultures were maintained as suggested by the ATCC. Growth medium was DMEM supplemented with 10% fetal bovine serum and pen-strep. 3D A431 cultures (“spheroids”) were grown in standard polystyrene 96-well tissue culture plates modified with an agarose meniscus in each well, as previously described.<sup>12,16</sup> Briefly, low-melting point agarose was suspended at 1% (w/v) in water, then dissolved by maintaining in a 100 °C water bath for 30 min. The agarose solution was quickly dispensed in 100 µL portions into each well of a 96-well plate. The agarose was allowed to set at room temperature for at least 2 h before seeding cells. The final agarose gel formed a concave meniscus in each well which facilitates spheroid formation. A431 cells were trypsinized, then counted in a hemocytometer prior to seeding at a density of 50,000 cells/150 µL. For drug treatment studies, spheroids were allowed to coalesce for four days, then treated on the fifth day with

50  $\mu$ L of 4X DMSO (vehicle) or lapatinib in complete DMEM added to each well using a multichannel pipette. Spheroids were treated overnight then analyzed the next day. 2D A431 cultures were treated in parallel by growing monolayers to 80-90% confluence, then treating overnight with 1X vehicle or lapatinib. In all cases, the final concentration of lapatinib was 60 nM, which is in the range of the reported IC<sub>50</sub> values for lapatinib *in vitro* and in A431 monolayers.<sup>17-19</sup> A convenient value on the lower end of this range was chosen in order to maximize the possibility of observing differences in response in the two culture systems.

### 3.2.3 Pinocytic Loading of A431 Cells

Pinocytic loading was performed according to the vendor's instructions and a previously published protocol.<sup>20</sup> A431 monolayers were stripped from the culture flask using trypsin with EDTA, then counted. Approximately 10<sup>6</sup> cells were pelleted 1 min at 2000 rpm, then resuspended in hypertonic loading medium containing 180  $\mu$ M Htc-tide. The cells were incubated 10 min at 37 °C, then pelleted. The loading medium was reserved for re-use up to three times. Loaded cells were immediately resuspended in 1 mL hypotonic lysis medium (60% serum free DMEM, 40% water) to lyse pinosomes and release Htc-tide into the cytosol. The cells in 1 mL hypotonic medium were quickly added to 2 mL hypotonic medium at 37 °C. This point was considered time zero in all cases. The cells were aliquoted into two 1.5 mL tubes, incubated 90 s, then pelleted by centrifugation (1 min at 2000 rpm). The

hypotonic medium was quickly removed and replaced by complete DMEM. The cells were either immediately plated for imaging or single cell analysis, or were allowed to recover at 37 °C until analysis. 3D A431 cultures were loaded as above following disaggregation with warm trypsin-EDTA and gentle vortexing.

Disaggregated spheroids were rinsed with warm media or PBS prior to use to remove necrotic debris from the spheroid core.

#### **3.2.4 Bulk EGFR Assays in Intact Cells**

2D and 3D cultured A431 cells were loaded with Htc-tide as described above, then incubated 50 min at 37 °C. The cells were washed 1X with PBS, pelleted, then resuspended in cold PBS containing 1X cOmplete protease inhibitor (Roche, Penzberg, Germany) and 1X PhosSTOP pan-phosphatase inhibitor (Roche, Penzberg, Germany). Cells were immediately lysed with heat (5 min at 100 °C). The lysates were clarified by centrifugation at 14000 rcf for 15 min in a 4 °C centrifuge. Samples were stored at -20 °C until analysis.

#### **3.2.5 Microscopy**

Brightfield and fluorescence microscopy was performed on a Nikon Eclipse inverted microscope fitted with an arc lamp and filter sets appropriate to the fluorophores used. Images were acquired with a CoolSNAP HQ2 CCD camera (Photometrics, Tucson, AZ) and analyzed in ImageJ (<http://rsb.info.nih.gov/ij>).

### **3.2.6 Western Blotting**

2D and 3D cultured A431 cells treated as indicated were collected by trypsinization, then washed twice with PBS. Cells were lysed by suspending in a minimal volume of NP40 lysis buffer supplemented with 1X cOmplete protease inhibitor and 1X PhosSTOP, then incubating on ice for 20 min with occasional agitation. Lysates were clarified by centrifugation as above then analyzed for total protein content by a Bradford assay (BioRad, Hercules, CA). 20-50  $\mu\text{g}$  total protein/15  $\mu\text{L}$  was denatured in sample buffer containing bromophenol blue and 2-mercaptoethanol by boiling at 100 °C, then separated by SDS-PAGE. Separated denatured proteins were transferred to a PVDF membrane for 1 h at 100 V (4 °C). The membrane was blocked 30 min in 5% (w/v) bovine serum albumen in TBST before incubating overnight with primary antibody at 4 °C. Incubation of secondary antibody was performed for 1 h at 4 °C immediately prior to imaging. Blots were developed using Bio-Rad Clarity Western ECL Substrates (catalog #170-5061) and imaged using an Alpha Innotech Fluor Chem FC2 imager.

### **3.2.7 Capillary Electrophoresis**

Capillary electrophoretic separations were performed in 30  $\mu\text{m}$  i.d. bare silica capillaries. The background electrolyte used was 200 mM boric acid/NaOH, pH 9.5, with 2 mM SDS, unless otherwise noted. Detection was performed with laser-induced fluorescence as described in the preceding chapter. Standards were

loaded hydrodynamically for 10 s. The estimated loading volume, taking into account both hydrodynamic loading and spontaneous fluid displacement, was 530 pL for a 10 s injection.

### **3.2.8 Single Cell Analysis**

Single cell analysis was performed as described in the preceding chapter.

### **3.2.9 Data Analysis**

Electropherograms were processed using custom MATLAB software. Software functions included additional data filtering through a low-pass second order Chebyshev filter, automated baseline determination, and peak integration. Separation efficiencies and peak resolution were calculated based on peak parameters derived from Gaussian peak fitting performed using IgorPro.

## **3.3 Results and Discussion**

### **3.3.1 Proposed EGFR Assay in 3D Cultured A431 Cells**

The proposed workflow for single cell analysis of cells isolated from 3D A431 cell cultures is represented in Figure 11, and is similarly applicable to primary tumor specimens obtained from clinical or research sources. Following treatment of 3D culture with the EGFR TKI lapatinib, the spheroid is disaggregated using trypsin-EDTA to yield a single cell suspension. In the case of clinical specimens, optimized protocols appropriate to the tissue would be implemented. The cells are then loaded with Htc-tide and incubated to allow phosphorylation by EGFR. For this early work,

the cells are loaded using pinocytosis (Figure 12B). Ultimately, Htc-tide will be passively loaded to minimize cellular perturbation and resulting experimental artifacts. Passive loading may be accomplished by a variety of strategies, discussed extensively elsewhere.<sup>21-25</sup> Promising possibilities include covalent conjugation of the reporter to a lipid or cell-penetrating peptide to drive passive diffusion across the outer cellular membrane. Following an appropriate incubation period, the cells are lysed and analyzed as discussed in the preceding chapters. Automated serial analysis will allow hundreds to thousands of cells to be analyzed in a single experiment.

### **3.3.2 Pinocytic Loading of A431 Cells with Htc-tide**

For assay validation, pinocytosis was used to introduce Htc-tide into the cytosol of large populations of cells. Briefly, a hypertonic loading solution was prepared containing 200  $\mu$ M Htc-tide. The hypertonic medium consisted of the Influx<sup>TM</sup> loading reagent (Life Technologies) in DMEM with 5% serum and 50 mM of a HEPES buffer. Incubation of cells in the hypertonic solution caused the cells to equalize the tonicity across the plasma membrane by taking up packets of the surrounding medium into the cytosol. These packets, called pinosomes, were encased in plasma membrane; subsequent incubation of the loaded cells in a hypotonic solution caused the pinosomes to lyse and released the contents into the cytosol, thus initiating the intracellular kinase assay. The cells can then be exchanged



into complete medium to recover prior to imaging and/or analysis. Representative brightfield and fluorescence microscopy images of A431 cells loaded with Htc-tide by pinocytosis are presented in Figure 12A and B, respectively. The cells were co-stained with propidium iodide (PI), a DNA and RNA-intercalating fluorophore excluded by cells with an intact plasma membrane, to indicate dead cells. No evidence of co-localization of Htc-tide and PI is observable in Figure 12B, indicating that Htc-tide-loaded cells remain viable after pinocytosis.

A431 monolayers (2D cultures) were serum-starved overnight and either left untreated or stimulated with EGF immediately prior to loading with Htc-tide. CE-LIF analysis of the cytosolic fraction of bulk homogenates of these cells indicate that Htc-tide phosphorylation in intact A431 cells is EGF-dependent (Figure 12C and D). Taken together with literature evidence for exceptionally high EGFR expression levels in A431 cells, the strong EGF dependence of Htc-tide phosphorylation suggests that EGFR is primarily responsible for observed reporter phosphorylation under these conditions.

### **3.3.3 3D Culture of A431 Cells**

Generation of 3D A431 cultures, or spheroids, was accomplished as outlined in Figure 13. Standard polystyrene 96-well tissue culture plates were modified for culture by adding 100  $\mu$ L of a warm 1% (w/v) agarose solution to each well, then allowing the solution to gel at room temperature for at least two hours. As

the agarose solidified it formed a deep concave meniscus in the bottom of each well. A431 cells are unable to adhere to the agarose surface. Therefore, cells in suspension seeded into the wells coalesced into compact spheroids, aided by the geometry of the well bottom. A431 cells seeded at 50,000 cells/well-formed spheroids within 24 h. The diameter of the spheroids quantified by microscopy gradually decreased over the course of 4 days, presumably due increases in spheroid compactness and intercellular adhesion over time. Spheroids to be analyzed were treated with drug on day 4 and analyzed on day 5 of culture. Prior to analysis either by western blotting or peptide reporter assays, spheroids were disaggregated with trypsin-EDTA and washed to remove the necrotic core.

### **3.3.4 Differences in EGFR Activation in 2D and 3D Cultured A431 Cells**

Relative EGFR expression and autophosphorylation in 2D and 3D A431 cultures were compared by western blotting (Figure 14A). In keeping with published reports of EGFR expression in CRC spheroids, total EGFR expression is slightly decreased in the 3D A431 cultures.<sup>14</sup> EGFR autophosphorylation at Y1068, a proxy for receptor activation, is also lower in the 3D cultures. This observation supports the canonical model of spheroid structure, in which only the outermost ring of cells are actively proliferating. Given that EGFR is a key player in cell growth and proliferation, it is to be expected that its activity will be lower in the quiescent inner ring of the spheroid than in the outer proliferative ring. EGFR

autophosphorylation is inhibited in both models upon treatment with lapatinib at 60 nM.

Surprisingly, EGFR activity assessed by the intracellular Htc-tide assay is higher in the 3D cultured cells (Figure 14B). Treatment with lapatinib reduces Htc-tide phosphorylation to the levels observed in the 2D cultured cells. The discrepancy between the western blotting data and the peptide assay may be due in part to the fact that western blotting cannot directly measure kinase activity. Rather, phosphorylation at one of the receptor's autophosphorylation sites (Y1068) is taken as a proxy for kinase activity. While this is a common approach to estimating EGFR activity, phosphorylation at Y1068 alone provides an incomplete picture of EGFR activation. The receptor contains multiple phosphorylation sites which reflect kinase activity in complex, poorly-understood ways.<sup>26-30</sup> Reports of differential EGFR activation in 2D and 3D cultures rarely monitor phosphorylation of more than one of these sites, opting rather to probe phosphorylation of downstream proteins associated with EGFR signaling.<sup>14</sup> A further complication is differential metabolism of Htc-tide in the two culture systems (Figure 14C). While the extent of reporter metabolism is independent of drug treatment, Htc-tide metabolism is significantly higher in the 2D cultured cells. This may be attributable to generally higher metabolism rates in actively proliferating cells, which probably comprise a much higher percentage of the 2D cultures compared to the 3D cultures given uniform

access of the 2D cultured cells to nutrients and oxygen. This dual functionality of the reporter may be an interesting avenue for future research.

### **3.3.5 Separation Optimization for 4 cm Effective Length CE-LIF System**

CE-LIF analyses of Htc-tide assays in the preceding sections were performed on a conventional single cell microelectrophoresis platform with a capillary effective length of 27 cm, using a BGE of 200 mM boric acid/NaOH, pH 9.5, with 2 mM SDS. Total separation time on this system under optimized conditions is 2.5 h; this significantly limits throughput, particularly for single cell analyses. To enable higher throughput analyses comparing Htc-tide phosphorylation in single 2D and 3D cultured A431 cells, the separation was re-optimized for compatibility with the automated single cell analysis system, which has a capillary effective length of 4 cm. Figure 15A shows the dependence of Htc-tide/pHtc-tide resolution on both the buffer (boric acid) concentration and additive (SDS). Generally, 200 mM boric acid with concentrations of SDS below the CMC (8 mM) gave the best resolution. Addition of SDS to the BGE also greatly improves separation reproducibility (data not shown), presumably due to decreased protein and lipid adsorption to the capillary wall. Optimization of pH (Figure 15B) yielding somewhat puzzling results. Separation efficiencies (z-axis) increased with pH up to 9.5, then dropped abruptly. Resolution was best a pH 8.5 and pH 10.0, and was found to drop between pH 9.5 and pH 10.0. Given that differences in resolution were quite small overall and

consistently greater than 1.0, pH 9.8 was chosen, despite the lower resolution and separation efficiency, for initial testing of single cell analysis. The reason for this is that the conditions are only slightly more basic than the initial BGE but resulted in slightly longer migration times. This was thought to reduce the likelihood of a major change in fragment migration patterns and fragment co-migration with either Htc-tide or pHtc-tide. Furthermore, repeated comparison of pH 9.5 and pH 9.8 showed more reproducible migration times under the pH 9.8 conditions (data not shown). The resulting optimized separation is 35 min total, which is a significant improvement over the conventional separation. This was thought to be acceptable for initial validation of the single cell analysis. However, a more extensive buffer optimization will be necessary to further improve separation reproducibility and efficiency and to reduce separation time below 10 min. A separation on the order of a few minutes or shorter will greatly facilitate high throughput serial analysis. In the future, a range of different buffer systems, pH values, and additives will be tested to achieve the desired separation.

### **3.3.6 EGFR Activity in Single A431 Cells**

Electropherograms from initial tests of single cell analysis on the high throughput system are collected in Figure 16A-H. The standard separation (6A, 6E) is reproducible between days. Htc-tide phosphorylation in serum-starved, EGF stimulated 2D cultured cells (Figure 16B-D) is in keeping with population data

(Figure 2C) and increases with time. Phosphorylation in basal 2D cultured cells is lower as expected and is again time-dependent (Figure 16F-H). Htc-tide resistance to proteolysis is truly impressive, with a significant amount of Htc-tide and pHtc-tide detectable even after 3 h (Figure 16H). These preliminary data demonstrate proof-of-concept for future high throughput single cell analyses in 2D and 3D cultures.

### **3.4 Conclusions and Future Directions**

The data presented in this chapter lay the foundation for future high-throughput analysis of EGFR activity in single cells. Htc-tide is differentially phosphorylated and metabolized in intact cells from 2D and 3D cell cultures. Htc-tide phosphorylation decreases upon EGFR inhibition by the selective inhibitor lapatinib. The separation component of the assay was successfully shortened by two hours, and future work will shorten the total assay time even further. Future work will focus on applying this validated assay to biologically relevant systems. Analyses in 3D cultured cells will continue to serve as a method development and validation platform, particularly as the high throughput analysis workflow is further refined. Ultimately, this assay will be applied to EGFR activity measurements in single primary cells, including those derived from pancreatic xenografts exhibiting elevated expression of EGFR. The ability to assay response to EGFR inhibitors at the single cell level will be a valuable tool for probing heterogeneous EGFR activity and drug sensitivity in real populations of tumor cells.

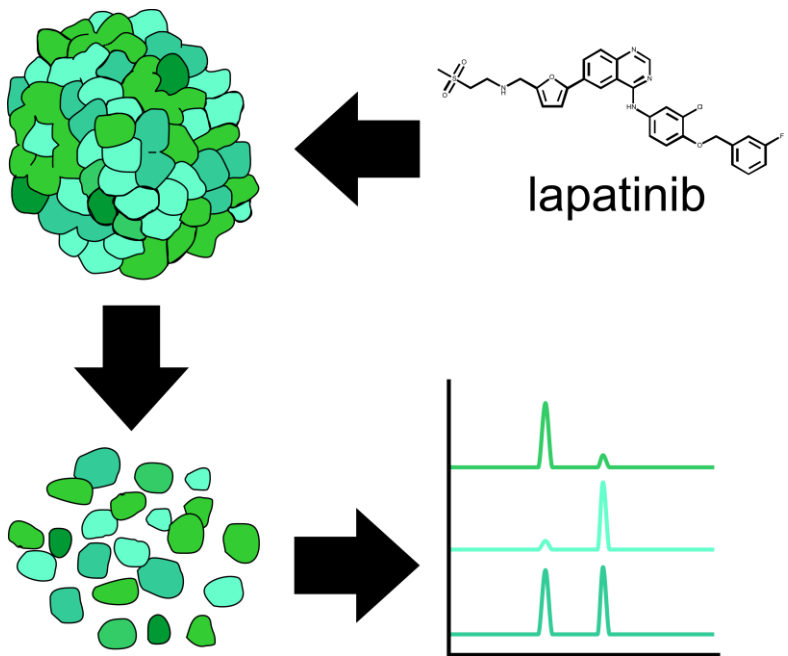
## REFERENCES

- (1) Almendro, V.; Marusyk, A.; Polyak, K. *Annu. Rev. Pathol. Mech. Dis.* **2013**, *8*, 277–302.
- (2) Bhatia, S.; Frangioni, J. V.; Hoffman, R. M.; Iafrate, A. J.; Polyak, K. *Nature Biotechnology* **2012**, *30*, 604–610.
- (3) Yap, T. A.; Gerlinger, M.; Futreal, P. A.; Pusztai, L.; Swanton, C. *Science Translational Medicine* **2012**, *4*, 127ps10.
- (4) Foo, J.; Leder, K.; Mumenthaler, S. M. *Evol Appl* **2012**, *6*, 54–69.
- (5) Gerlinger, M.; Rowan, A. J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; Varela, I.; Phillimore, B.; Begum, S.; McDonald, N. Q.; Butler, A.; Jones, D.; Raine, K.; Latimer, C.; Santos, C. R.; Nohadani, M.; Eklund, A. C.; Spencer-Dene, B.; Clark, G.; Pickering, L.; Stamp, G.; Gore, M.; Szallasi, Z.; Downward, J.; Futreal, P. A.; Swanton, C. *N. Engl. J. Med.* **2012**, *366*, 883–892.
- (6) Martini, M.; Vecchione, L.; Siena, S.; Tejpar, S.; Bardelli, A. *Nature Publishing Group* **2011**, *9*, 87–97.
- (7) Fedele, C.; Tothill, R. W.; McArthur, G. A. *Cancer Discovery* **2014**, *4*, 146–148.
- (8) Kleppe, M.; Levine, R. L. *Nature Publishing Group* **2014**, *20*, 342–344.
- (9) Dickinson, A. J.; Armistead, P. M.; Allbritton, N. L. *Anal. Chem.* **2013**, *85*, 4797–4804.
- (10) Jiang, D.; Sims, C. E.; Allbritton, N. L. *ELECTROPHORESIS* **2010**, *31*, 2558–2565.
- (11) Lin, R.-Z.; Chang, H.-Y. *Biotechnol. J.* **2008**, *3*, 1172–1184.
- (12) Breslin, S.; O’Driscoll, L. *Drug Discovery Today* **2013**, *18*, 240–249.
- (13) Liu, X.; Weaver, E. M.; Hummon, A. B. *Anal. Chem.* **2013**, *85*.

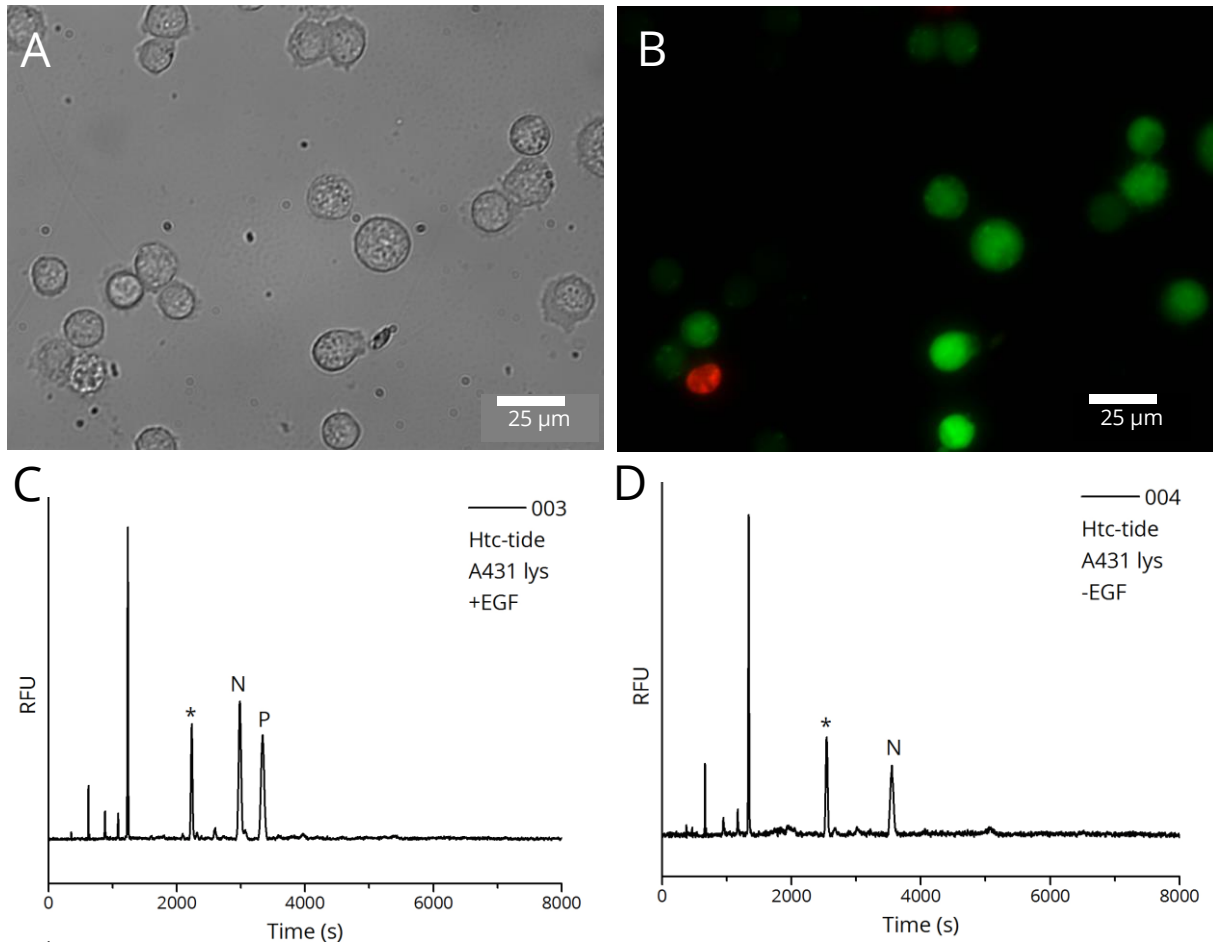
- (14) Luca, A. C.; Mersch, S.; Deenen, R.; Schmidt, S.; Messner, I.; Schäfer, K.-L.; Baldus, S. E.; Huckenbeck, W.; Piekorz, R. P.; Knoefel, W. T.; Krieg, A.; Stoecklein, N. H. *PLoS ONE* **2013**, *8*, e59689.
- (15) Weigelt, B.; Lo, A. T.; Park, C. C.; Gray, J. W.; Bissell, M. J. *Breast Cancer Res Treat* **2009**, *122*, 35–43.
- (16) Ho, W. Y.; Yeap, S. K.; Ho, C. L.; Rahim, R. A.; Alitheen, N. B. *PLoS ONE* **2012**, *7*, e44640.
- (17) Rusnak, D. W.; Alligood, K.; Mullin, R. J.; Spehar, G. M.; Arenas Elliott, C.; Martin, A. M.; Degenhardt, Y.; Rudolph, S. K.; Jr Haws, T. F.; Hudson Curtis, B. L. *Cell Proliferation* **2007**, *40*, 580–594.
- (18) Rusnak, D. W.; Lackey, K.; Affleck, K.; Wood, E. R.; Alligood, K.; Rhodes, N.; Keith, B. R.; Murray, D. M.; Knight, W. B.; Mullin, R. J.; Gilmer, T. M. *Molecular Cancer Therapeutics* **2001**, *1*, 85–94.
- (19) Mahboobi, S.; Sellmer, A.; Winkler, M.; Eichhorn, E.; Pongratz, H.; Ciossek, T.; Baer, T.; Maier, T.; Beckers, T. *J. Med. Chem.* **2010**, *53*, 8546–8555.
- (20) Kovarik, M. L.; Dickinson, A. J.; Roy, P.; Poonnen, R. A.; Fine, J. P.; Allbritton, N. L. *Integr Biol (Camb)* **2014**, *6*, 164–174.
- (21) Koppelhus, U.; Nielsen, P. E. *Advanced Drug Delivery Reviews* **2003**.
- (22) Zorko, M.; Langel, Ü. *Advanced Drug Delivery Reviews* **2005**, *57*, 529–545.
- (23) Nelson, A. R.; Borland, L.; Allbritton, N. L.; Sims, C. E. *Biochemistry* **2007**, *46*, 14771–14781.
- (24) Nasrolahi Shirazi, A.; Tiwari, R. K.; Oh, D.; Banerjee, A.; Yadav, A.; Parang, K. *Mol. Pharmaceutics* **2013**, *10*, 2008–2020.
- (25) Hawiger, J. *Current Opinion in Chemical Biology* **1999**.
- (26) Zhang, X.; Gureasko, J.; Shen, K.; Cole, P. A.; Kuriyan, J. *Cell* **2006**, *125*, 1137–1149.



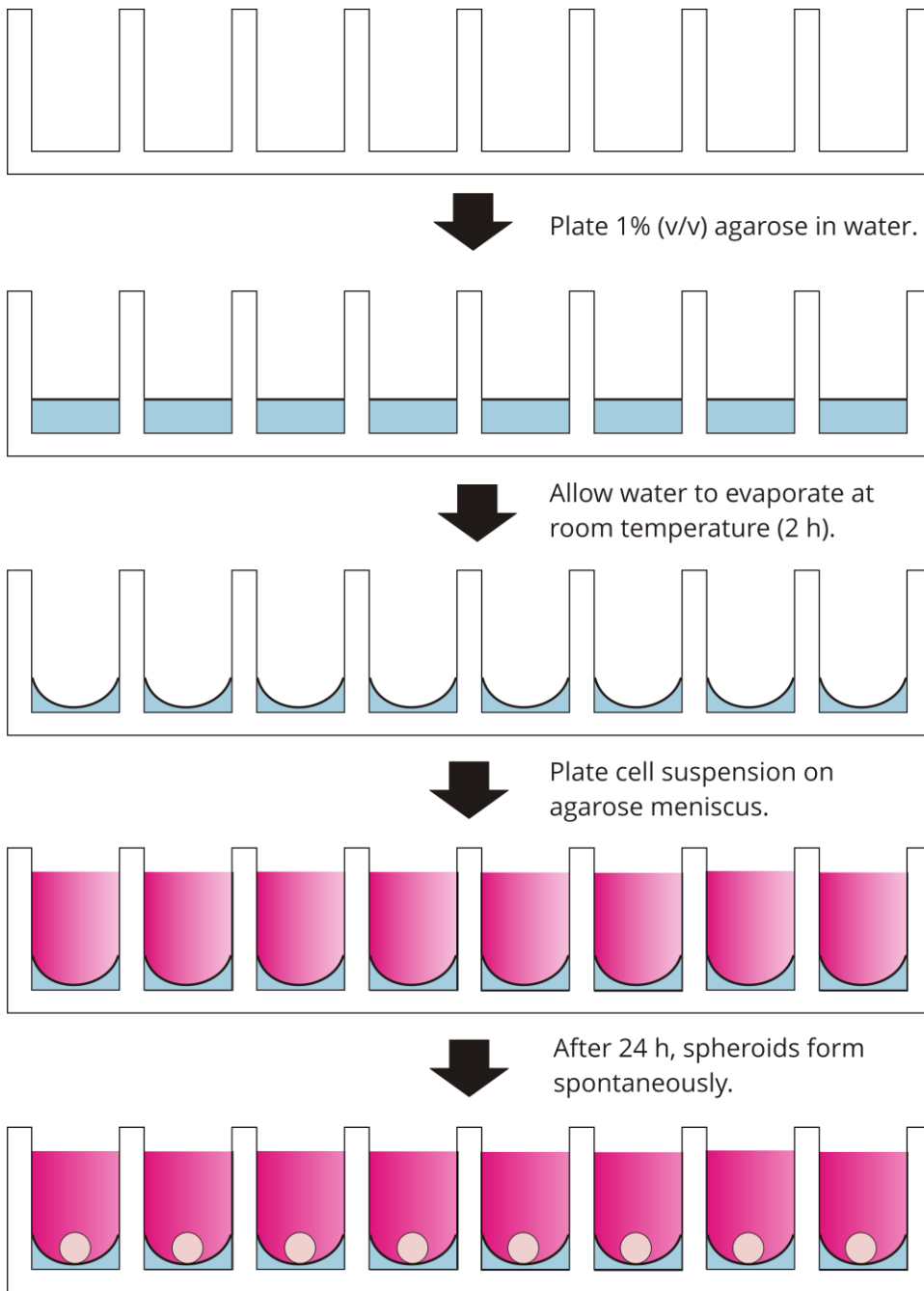
- (27) Jorissen, R. *Experimental Cell Research* **2003**, *284*, 31–53.
- (28) Endres, N. F.; Barros, T.; Cantor, A. J.; Kuriyan, J. *Trends Biochem. Sci.* **2014**, *39*, 437–446.
- (29) Lemmon, M. A.; Schlessinger, J. *Cell* **2010**, *141*, 1117–1134.
- (30) Hubbard, S. R.; Miller, W. T. *Current Opinion in Cell Biology* **2007**, *19*, 117–123.



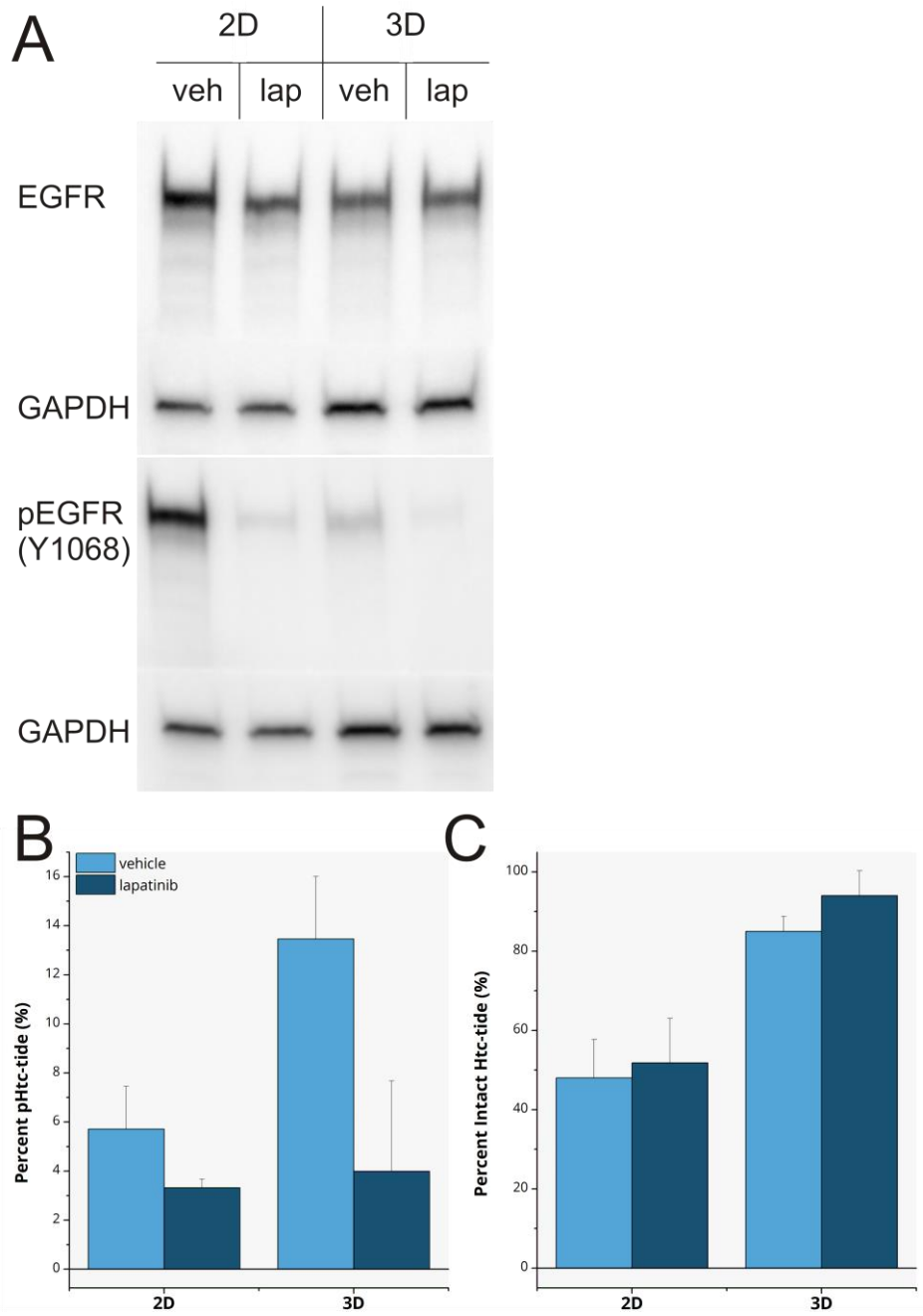
**Figure 11. Proposed workflow for single cell analysis of EGFR activity in response to lapatinib treatment in 3D cell cultures and solid tumors.**



**Figure 12. Pinocytic loading of Htc-tide into A431 cells.** (A) Brightfield microscopy of A431 cells after pinocytic loading and seeding onto CellTak. (B) Composite fluorescence microscopy image of Htc-tide-loaded (green) and propidium iodide-stained (red) A431 cells. (C)-(D) Bulk lysates of A431 cells loaded with Htc-tide after serum starving overnight and stimulating with EGF (C) or leaving untreated (D). Symbols: \* = internal standard; N = Htc-tide; P = pHtc-tide.

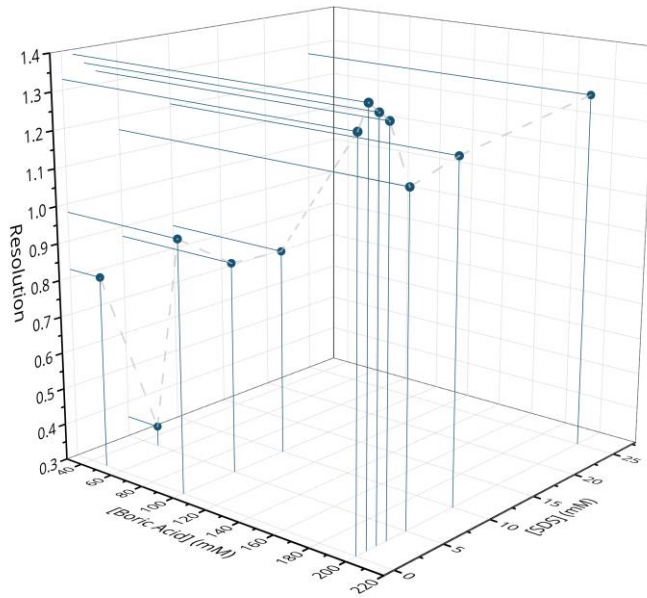


**Figure 13. Schematic of 3D culture technique.**

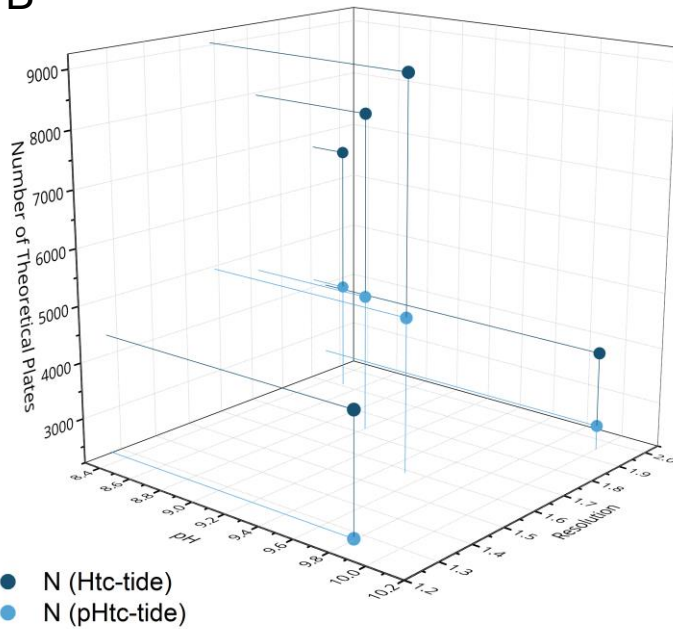


**Figure 14. Differences in EGFR activity in 2D and 3D A431 cultures.** (A) Western blot showing decreased EGFR expression and autophosphorylation in 3D cultures vs. 2D cultures. Blots are representative of triplicate experiments. (B)-(C) Bulk lysate data from 2D and 3D cultured A431 cells loaded with Htc-tide. Htc-tide is phosphorylated to a greater extent in 3D cultures (B), but is metabolized to a lesser extent (C). N = 3. Blots by Mike Lebhar and Josh Mu.

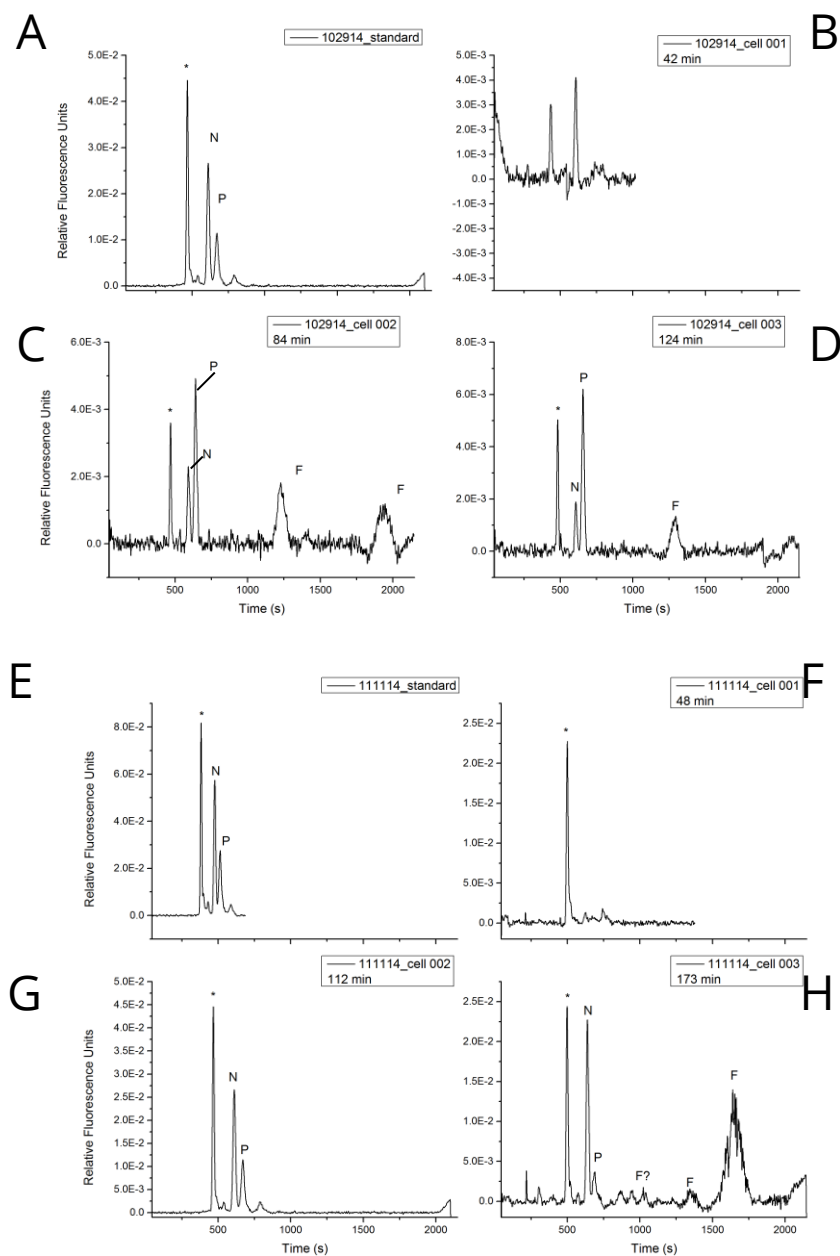
A



B



**Figure 15. CE separation optimization for high throughput serial analysis. (A)** Resolution of Htc-tide and pHtc-tide is dependent on the concentration of boric acid buffer (pH 9.5) and of SDS. **(B)** Htc-tide/pHtc-tide resolution and separation efficiencies depend strongly on pH in a complex manner. N = 1



**Figure 16. Single cell analysis on the high throughput system on two separate days.** (A)-(D) Day one, analysis of single A431 cells that had been serum starved and stimulated with EGF prior to loading. (A) Standard separation on day one. (E)-(H) Day two, analysis of single A431 cells treated with DMSO in complete medium (vehicle). (E) Standard separation on day two. Incubation time for each cell is indicated on the electropherograms. Symbols: \* = Internal Standard; N = Htc-tide; P = pHtc-tide; F = Fragment.

## **CHAPTER 4: A HIGHER THROUGHPUT SITE-DIRECTED COMBINATORIAL APPROACH TO KINASE REPORTER OPTIMIZATION**

### **4.1 Limitations of Kinase Reporter Selectivity**

The specificity of peptide substrates for tyrosine kinases is constrained by the lack of native secondary specificity determinants including spatial and temporal localization, defined secondary structure, and auxiliary domains which promote protein-peptide binding.<sup>1,2</sup> Because the catalytic domain is highly conserved across all known tyrosine kinases, few kinases have unique recognition motifs, and a large degree of overlap in recognition is found even among functionally distinct enzymes.<sup>3,4</sup> Tyrosine kinases in intact cells rely on stimulus-specific spatiotemporal context and highly selective docking domains to recruit appropriate substrates for phosphorylation.<sup>5-7</sup> This exquisitely evolved mechanism of specificity is exceptionally challenging to mimic synthetically. Nevertheless, many researchers have explored rational and semi-rational design of selective peptides.

#### **4.1.1 Synthetic Approaches to Spatiotemporal Control**

An elegant approach to intracellular peptide localization involves synthetically modifying the peptide with a lipid to encourage membrane localization. Lipid modifications that have been reported include myristic acid<sup>8</sup>, cholesterol<sup>9</sup>, and palmitic acid<sup>10</sup>. An alternative to lipid targeting is to



incorporate a peptidic targeting signal into the overall peptide substrate design. These include membrane-crossing sequences<sup>11-13</sup>, nuclear targeting sequences (e.g. TAT)<sup>14-16</sup>, and mitochondria targeting sequences<sup>17-20</sup>. Lipid and peptide targeting groups are usually covalently tethered to the peptide substrate cargo; if desired, this linkage can be engineered to be selectively cleaved following substrate localization. For example, a lipid may be employed to localize a peptide substrate proximal to its membrane-bound target enzyme and to protect from off-target enzymatic activities, followed by release of the substrate cargo in the vicinity of the target enzyme for improved access to the enzyme active site. Cleavable linkages that may be employed include disulfide linkages, which are cleaved in the reducing environment of the cytosol,<sup>21,22</sup> photocleavable linkages that release the cargo upon irradiation with a specific photon energy, and enzyme-cleavable linkages.

Some groups have reported incorporating peptidic docking domains in their overall substrate design to facilitate selective binding to the kinase of interest. Small protein domains such as SH2 and WW bind, or “dock onto”, specific phosphorylated residues within target kinases, thus recruiting their cargo for phosphorylation and activity modulation.<sup>5,23</sup> Full-length docking domains are too large for complete chemical synthesis, but many groups have reported incorporating biosynthetic docking domains in genetically engineered peptide biosensors to increase selectivity<sup>24</sup> or as an integral part of the sensing mechanism.<sup>25-27</sup> Fernandes,

*et al.* reported that appending a short peptide derived from an allosteric binding domain of a MAPK protein substrate to a short synthetic MAPK peptide substrate increased peptide affinity for MAPK 200 fold.<sup>28</sup>

In addition to control of spatial localization, some groups have reported methods to exert temporal control over intracellular peptide substrates. This is achieved by blocking active sites within the peptide substrate with a stimulus-responsive functional group. For example, “photocaging” of the substrate phosphorylation site shields the peptide from being acted upon by the enzyme target. Exposure to the correct wavelength of light releases the protecting group and makes the peptide available for phosphorylation. Thus the experimenter is able to precisely control the experiment “time zero” and potentially derive kinetic information from the experiment.<sup>29</sup>

#### **4.1.2 Modification with Unnatural Amino Acids**

Traditional peptide substrates are limited both in overall size (< 20 a.a.) and in composition. Peptide substrates derived from intracellular protein phosphorylation sites or from combinatorial motif scans are constructed from among the 20 proteinogenic amino acids; often, this limited pool of chemical diversity is further reduced by omitting side reaction-prone amino acids such as tryptophan and cysteine.<sup>30</sup> Attempts to discover highly selective substrates of reasonable (synthesizable) size from such a limited pool of chemical characteristics have largely

failed.<sup>31,32</sup> An alternative approach, and one complementary to those discussed in the previous section, exploits the exponential increase in chemical diversity accessible by synthetically incorporating nonnative and non-proteinogenic amino acids into peptide substrates. Rational substitution of nonnative amino acids into promiscuous peptide substrates has proven to be an effective, but laborious, method to increase substrate specificity of the target kinase over off-target enzymes.<sup>33,34</sup> Lawrence and colleagues developed a more efficient high-throughput optimization technique in which an existing peptide ligand is converted to a scaffold for site-directed combinatorial modification by substituting diaminopropionic acid (Dap) at a selected site within the original substrate. The Dap residue provides a free primary amine which is subsequently acylated with a highly diverse library of synthetic carboxylic acids to generate a peptide library containing novel unnatural side chains at the selected position. The key to this approach is synthesis of the library on a custom cystamine modified resin, allowing DTT-mediated cleavage of the peptide library from the resin in kinase assay buffer. The library can then be immediately screened for activity, eliminating laborious isolation and purification steps. The authors applied this powerful technique to the development of a highly selective and potent peptidic inhibitor of PKC $\alpha$  ( $IC_{50} \approx 2$  nM) capable of discriminating between the target kinase and highly homologous isoforms (PKC $\beta,\gamma$ ).<sup>35</sup>

### 4.1.3 Site Directed Combinatorial Method for Substrate Optimization

The combinatorial optimization scheme developed by Lawrence and colleagues for discovery of high-affinity peptidic inhibitors can be modified to develop selective tyrosine kinase substrates. Peptide substrates of tyrosine kinases are notoriously promiscuous because the high degree of homology among the catalytic domains of tyrosine kinases leads to extensive overlap in recognition motifs. This chapter outlines a combinatorial optimization campaign, based on the work of Lawrence et al, whose purpose is to fine-tune the selectivity of an existing peptide substrate for the Epidermal Growth Factor Receptor (EGFR). This substrate “TyrSub” (EELEDDYEDDNleEE)<sup>36,37</sup> is a previously published sequence pulled from an *in vivo* phosphorylation site and is a substrate for multiple enzymes in addition to EGFR. The 6FAM-labeled version of the substrate (“TyrSub-13”, 6FAM-EELEDDYEDDNleEE-NH<sub>2</sub>) is efficiently phosphorylated by EGFR *in vitro* and exhibits excellent resistance to proteolytic degradation in cell lysates.<sup>38,39</sup> However, its susceptibility to off-target kinase activities limits its usefulness as an analytical tool in biological matrices; increased selectivity for EGFR would enable improved biochemical analyses of EGFR activity in cell lysates and intact cells.

## 4.2 Methods

The workflow for the combinatorial optimization of TyrSub is outlined in Figure 17. First, a truncated version of the reporter “TyrSub-9” (6FAM-

LEDDYEDDNle-NH<sub>2</sub>) was chosen as the scaffold peptide for optimization because it is phosphorylated equivalently to TyrSub-13, is similarly stable in cell lysates, and can be synthesized more efficiently. The scaffold site for acylation with the carboxylic acid library was chosen as described in detail below, followed by synthesis of the total scaffold on DTT-cleavable resin. Lysine was chosen to provide the free amine for acylation, rather than Dap, because its longer side chain reduces the steric barrier to acylation with bulky carboxylic acids at positions proximal to the resin. Selective acylation of the scaffold with a structurally-diverse 90-member carboxylic acid library yielded a 90-member peptide library with novel, unnatural side chain functionalities at the modification site. Cleavage of the peptide library with DTT kinase assay buffer was followed by screening for *in vitro* phosphorylation by EGFR and by the off-target kinase Spleen Tyrosine Kinase (Syk). Hits generated were synthesized on amide resin, extensively purified, and validated with quantitative CE-LIF.

#### **4.2.1 General Synthetic Methods**

C-terminal amide peptides were synthesized by standard Fmoc SPPS on NovaSyn TGR amide resin (NovaBiochem, EMD Millipore, Billerica, MA); synthesis was performed manually in fritted polymer syringes or using a Protein Technologies PS3 semi-automated peptide synthesizer. Amino acid building blocks were orthogonally protected with base-labile Fmoc and acid-labile tBu protecting groups

at the  $\alpha$ -amino and side chain phenol or carboxylic acid groups, respectively, unless otherwise noted. Coupling of sequential amino acids to the resin was effected using the HCTU/DIEA/DMF method (Appendix A). Difficult couplings were repeated with 2-4 fold increases in reagent excess where economically feasible. Completed coupling was verified by a negative Ninhydrin test (AnaSpec, Fremont, CA, Appendix A). Fmoc removal was effected as detailed in Appendix A. The deprotection solution was supplemented with 0.1M HOBt, a weak organic acid, to suppress piperidine-mediated racemization through aspartimide formation. C-terminal thiol peptides were synthesized on custom cystamine-modified resin, described below, using the same protocols as above.

N-terminal fluorescein labeling of Fmoc-deprotected peptides with 6-carboxyfluorescein (6FAM, AnaSpec, Fremont, CA) was accomplished using either an HCTU/DIEA/DMF or DIC/HOBt/NMP method (18-24 h, Appendix A). The DIC/HOBt/NMP method was preferred for improved coupling efficiencies with fewer moles of fluorophore. One to two overnight couplings were generally necessary for complete labeling as monitored by Ninhydrin test. Successfully labeled peptide resin was treated twice (5-15 min) with deprotection solution to cleave fluorescein esters immediately prior to cleavage. N-terminal biotin labeling of deprotected peptides was performed by reacting N-terminal deprotected peptide resin overnight with 10 molar equivalents each of O-Nitrophenyl ester-protected

biotin and HOBt in NMP. Complete labeling was checked by Ninhydrin test and repeated couplings performed as needed. Completed peptides were cleaved from the resin and isolated using the protocol appropriate to the resin used (Appendix A), followed by reconstitution where appropriate in aqueous media.

#### **4.2.2 Cystamine Resin Preparation and Characterization**

DTT-cleavable cystamine resin was prepared as described previously.<sup>35</sup> Briefly, PEG-grafted copolymer polystyrene resin beads modified with free carboxylic acid functional groups (Tentagel-S-COOH, AnaSpec; loading capacity 0.2  $\mu\text{mol}/\text{mg}$ ) were pre-activated 1 h with 5 eq TSTU and 15 eq DIEA dissolved in DMF (50 mL/g resin). Following pre-activation, 10 eq cystamine and 20 eq DIEA dissolved in an equivalent volume of water were added to the reaction vessel, and the reaction allowed to proceed overnight. The resin beads changed color from pale yellow to brown, indicating successful coupling. Coupling was terminated after 16-24 h by washing 6X with each of the following: distilled water, DMF, and DCM. Successful coupling was confirmed by a positive Ninhydrin test. Successfully modified resin was then loaded with Fmoc amino acid (10 eq) using HCTU/DIEA/DMF (Appendix A). Loading was allowed to proceed overnight, followed by sequential rinsing with DMF, IPA, and DCM. Complete loading of resin was verified by a negative Ninhydrin test.

Fmoc-amino acid loading of the cystamine resin was determined as follows. A standard curve of piperidine-dibenzofulvene ( $\lambda_{\text{max}} = 290 \text{ nm}$ ) was prepared by monitoring absorbance as a function of Fmoc-Ala-OH concentration in 30% (v/v) piperidine in DMF. Piperidine-dibenzofulvene forms as a result of piperidine-mediated Fmoc removal from the Fmoc-amino acid. Absorbance measurements at 290 nm were taken in a quartz cuvette using a Molecular Devices SpectraMax M5 spectrometer. Three 10 mg portions of vacuum-dried, loaded cystamine resin were aliquoted into 1.5 mL polypropylene tubes and treated 5 min with 30% (v/v) piperidine in DMF, with gentle shaking, followed by centrifugation to pellet the resin. The absorbance at 290 nm of the supernatant was used to back-calculate the amount of Fmoc loaded onto the resin, and thus the resin loading capacity in  $\mu\text{mol}/\text{mg}$ . Typical loading capacities were in the range of  $0.04 \mu\text{mol}/\text{mg}$ .

#### **4.2.3 Crude Peptide Purification and Characterization**

Crude peptide products were purified by RP HPLC on Phenomenex C18 analytical and/or semi-preparative columns (Torrance, CA) using gradient elution on an Agilent HPLC with UV and fluorescence detection modules (Santa Clara, CA). High yield peptides requiring little purification were eluted using short gradients (2.5-3.3 %/min) from 100% water to 100% acetonitrile (each with 0.1% TFA ion pairing reagent). Low yield/difficult peptides were purified using extended gradients (1.4-2.0 %/min, water to acetonitrile or methanol) on higher efficiency



analytical columns where feasible. Eluted fractions were characterized by soft-ionization mass spectrometry (MALDI-TOF<sup>2</sup> MS or ESI-quadrupole MS) in negative ion mode.

#### 4.2.5 Library Synthesis and Characterization

A selection of 90 structurally diverse carboxylic acids was kindly provided by Dr. David Lawrence (Appendix C). Stock solutions (0.9 M) were prepared in DMF, aliquoted into deep-well solvent-resistant plates, and stored at -20 °C until use. The library scaffold was synthesized as above on cystamine resin and labeled with 5/6-FAM at the N-terminus. The modification site lysine was protected with highly acid-labile Mtt which could be removed with very low concentrations of TFA in DCM, leaving the tBu side chain protecting groups in place.

Acylation of the library site lysine was effected as follows. First, the Mtt protecting group was removed from the peptide scaffold resin by repeated treatment with 2% (v/v) TFA in DCM in 5 min increments over the course of 1 h until the TFA solution no longer turned yellow. For an unknown reason, the Mtt proved very difficult to remove, and the multiple TFA treatments needed almost certainly deprotected some of the tBu group as well. The deprotected scaffold peptide resin was washed sequentially with DMF, 5% (v/v) DIEA in DMF to neutralize TFA remaining, DMF, IPA, and DCM. It was then dried under a stream of N<sub>2</sub>(g) and stored at -20 °C until acylation with the library compounds.

For acylation with the library compounds, the dried, deprotected resin (0.5 g) was swelled 30 min at a density of 50 mg resin/mL DMF then distributed at 5 mg/well into a fritted solvent resistant 96-well plate. The DMF was removed by vacuum filtration, followed by resuspension in 0.05 mL of fresh activation solution (100 eq PyBOP:100 eq HOBt:1000 eq NMM). Acylation was initiated by adding 0.1 mL of a unique 0.9 M carboxylic acid solution (in DMF) to each well. The plate was sealed and gently shaken overnight. Repeated couplings with fresh activation solution were necessary for complete acylation as monitored by spot-checking with a Ninhydrin test. The acylated peptide library-bound resin was washed sequentially 3X with DMF, IPA, and DCM and dried 3 min under vacuum. The dried resin was stored at -20 °C until cleavage.

Cleavage of the peptide library first required removal of remaining acid-labile side chain protecting groups in 95% TFA/2.5% TIS/2.5% water (2.5 h). Fresh TFA needed to be added at intervals over the course of the cleavage because the small volume (0.2 mL) of volatile TFA evaporated rapidly (approximately every 45 min). The library-bound resin was rinsed sequentially with DMF, IPA, and DCM, dried, and stored at -20 °C until further use.

The resin was treated 30 min with 30% (v/v) piperidine in DMF to cleave any esters formed (fluorescein, carboxylic acid, etc.), followed by extensive rinsing with DMF, IPA, and DCM and drying 5 min over vacuum. The peptide library was then

cleaved from the resin using kinase assay buffer (10 mM DTT in 50 mM Tris, pH 7.5, fresh) as follows: 250  $\mu$ L for 1 h, 125  $\mu$ L for 1.5 h, 125  $\mu$ L for 1 h. After each step, the cleaved peptide was filtered into a deep-well 96-well plate. The cleaved library was aliquoted into 96-well plates and frozen at -20 °C until analysis.

The purity of each library compound was estimated by semi-quantitative MALDI-TOF MS analysis on an AB Sciex 4800 Plus instrument (Framingham, MA) using the following matrix:  $\alpha$ -cyanohydroxycinnamic acid saturated in 50% ACN (aq) with 5 mM citric acid and 0.1% TFA. Library compound was spotted 2:1 with matrix on a polished steel substrate (AB Sciex). Mass spectra were acquired in both positive and negative ion modes. Peaks from each spectrum were exported for analysis in MATLAB using the following parameters:

- mass range: 1200-2000 Da
- peak density: 1/5 Da
- $S/N_{\min}$ : 5
- $Area_{\min}$ : 100
- $(Peaks/Spot)_{\max}$ : 500

Peak density was defined as the maximum number of peaks detectable within a given mass window. The  $S/N_{\min}$  was defined as the minimum acceptable signal-to-noise ratio for a detected peak. The  $Area_{\min}$  was defined as the minimum acceptable area for a detected peak. The  $(Peaks/Spot)_{\max}$  was defined as the maximum number of exported peaks in a given sample. The exported lists of peaks and signal intensities were processed with the MATLAB code shown in Appendix B.

Nine compounds were selected randomly for quantitative purity analysis with LC/MS. Analyses were performed on an Agilent 1200 series instrument with an Agilent 6100 single quadrupole spectrometer. The column was a C18 column from Grace (Alltima, 3  $\mu$ m particles, 2.1 X 50 mm; Columbia, MD). Samples were eluted with a 6.1%/min gradient of 3% to 95% acetonitrile/water (0.1% formic acid) following an isocratic desalting step. Mass spectra were acquired in negative ion mode.

#### **4.2.6 High Yield Synthesis of Asp-Containing Scaffold and Library Hit Compounds**

The library scaffold was synthesized on CLEAR amide resin (Peptides International) on a microwave synthesizer (Liberty, CEM Corp., Matthews, NC). The sequence prepared was: Fmoc-LEDDYEDK(Dde)Nle-(CLEAR resin). Fmoc-amino acids were coupled to the resin using HBTU/DIEA in DMF with double couplings at each position. Fmoc removal was accomplished using 20% (v/v) piperidine in DMF supplemented with 0.1 M HOBt. No evidence of aspartimide formation was observed in the intermediate unlabeled product by analytical RP-HPLC-MS. The scaffold was N-terminally deprotected and labeled with 6FAM, followed by three 2 min treatments with deprotection solution to remove FAM esters. Reactive hydroxyl groups on the 6FAM label were protected with an acid labile 2-chlorotriyl group (12 eq 2-chlorotriyl chloride + 12 eq DIEA in DCM, 16 h in a glass vessel). Tritylation of the 6FAM label is necessary to prevent reaction of the label with the hydrazine

deprotection solution in the next step. The scaffold was then selectively deprotected at the library modification site K(Dde) by treating the scaffold-bound resin 3 × 5 min with 3% (v/v) hydrazine in DMF. The library hit peptide 4-10F was prepared by labeling the deprotected K(NH<sub>2</sub>) site with N-(3-indolylacetyl)-L-phenylalanine (5 eq) using a slight excess of Oxyma and DIC in DMF. After a 10 min preactivation, the coupling reaction was allowed to proceed overnight. This was repeated at higher reagent excesses as needed for complete coupling as monitored by Ninhydrin test. Completed peptides were cleaved from the resin and isolated using the TFA cleavage detailed above. Synthesis of the desired product at high purity was verified by RP-HPLC-MS and CE-LIF.

#### 4.2.7 *In Vitro* Kinase Assays

*In vitro* kinase assays were performed under the optimal conditions for each kinase. Kinase buffers were as follows:

- Buffer A: 5 mM MOPS, pH 7.2; 0.05 mM DTT; 1 mM EGTA; 0.4 mM EDTA; 4 mM MgCl<sub>2</sub>; 10 mM MnCl<sub>2</sub>; 0.02% (w/v) BSA; 1 mM ATP
- Buffer B: 50 mM Tris-HCl, pH 7.5; 0.1 mM EGTA; 0.01% (w/v) BSA; 1 mM DTT; 0.1 mM sodium pervanadate; 50 mM MgCl<sub>2</sub>; 1 mM ATP
- Buffer C: 20 mM MOPS, pH 7.2; 1 mM EDTA; 0.1% (w/v) BSA; 10 mM MnCl<sub>2</sub>; 2 mM MgCl<sub>2</sub>; 5% (v/v) glycerol; 0.1 mM sodium pervanadate; 1 mM ATP
- Buffer D: 50 mM Tris-HCl, pH 7.5; 0.5 mM MnCl<sub>2</sub>; 5 mM MgCl<sub>2</sub>; 2 mM DTT; 1 mM ATP
- Buffer E: 50 mM Tris-HCl, pH 7.5; 100 mM NaCl; 0.1 mM EDTA; 5 mM DTT; 50% (v/v) glycerol
- Buffer F: 8 mM MOPS, pH 7.0; 0.2 mM EDTA; 4 mM MgCl<sub>2</sub>; 1 mM ATP

- Buffer G: 20 mM MOPS, pH 7.0; 5% (v/v) glycerol; 0.01% (w/v) Brij-35; 0.1% (v/v)  $\beta$ -mercaptoethanol; 0.1% (w/v) BSA
- Buffer H: 5 mM MOPS, pH 7.2; 1 mM EGTA; 15.3 mM EDTA; 10 mM MnCl<sub>2</sub>; 4 mM MgCl<sub>2</sub>; 0.025 mM DTT; 8 ng/mL BSA; 1 mM ATP
- Buffer I: 5 mM MOPS, pH 7.2; 1 mM EGTA; 15.3 mM EDTA; 10 mM MgCl<sub>2</sub>; 0.025 mM DTT; 8 ng/mL BSA; 1 mM ATP

All *in vitro* kinase assays were performed at 30 °C at the indicated enzyme and peptide concentrations. Assays were terminated either with an equal or greater volume of 0.2 M HCl or by addition of EDTA to a final concentration of 10 mM. Assays of EGFR kinase activity were performed in Buffer A using recombinant purified EGFR kinase domain (residues 696-end). Spleen Tyrosine Kinase (Syk) assays were performed in Buffer B with the full length recombinant protein. HER4/ErbB4 tyrosine kinase assays were performed in Buffer C using the recombinant catalytic domain (residues 706-991). Abelson tyrosine kinase (Abl) assays were performed in Buffer D with recombinant full-length Abl diluted in Buffer E. Src tyrosine kinase assays were performed in Buffer F with full length recombinant Src diluted in Buffer G. All enzymes were purchased from EMD Millipore, aliquoted at working volumes, and stored at -80 °C until use. Assays were performed in low retention polypropylene microcentrifuge tubes at 10-30  $\mu$ L total volumes.

High throughput kinase assays for library compound screening were performed in 40  $\mu$ L-well 96 well polystyrene plates at total assay volumes of 10  $\mu$ L.

Briefly, 2X reaction mixture containing EGFR kinase domain (final concentration: 9 ng/ $\mu$ L) and Buffer A were pipetted into each well of the 96-well plate, followed by assay initiation by addition of 5  $\mu$ L 1X library compound (final concentration: 1/2X). High throughput Syk activity screens were performed in Buffer B to a total assay volume of 14  $\mu$ L with 0.5 ng/ $\mu$ L Syk and 1/2X library compound, as above. Plates were sealed and incubated at 30 °C for 60 min followed by assay termination by addition of an equal volume of 20 mM EDTA. Assays were aliquoted and stored at -20 °C until analysis.

#### **4.2.8 Capillary Electrophoresis**

Quantitative capillary electrophoresis analyses were performed on Beckman Coulter ProteomeLab PA800 (Pasadena, CA) automated instruments in 30  $\mu$ m i.d. bare silica capillaries pretreated with NaOH and HCl. Detection was performed with laser-induced fluorescence using a 488 nm argon ion laser. The optimized electrophoretic buffer was 0.3 M boric acid/NaOH, pH 7.5. Analytes were diluted to 100-250 nM in 1 part electrophoretic buffer to 2 parts water to promote analyte stacking and compensate for the high ionic strength of kinase assay samples. Separations were carried out at 600 V/cm.

#### **4.2.9 MALDI-TOF MS**

Semi-quantitative MALDI-TOF MS analyses were performed on an AB Sciex 4800 Plus instrument (Framingham, MA) using an  $\alpha$ -cyanohydroxycinnamic acid

matrix (saturated in 50% ACN (aq) with 0.1% TFA and 5 mM citric acid). Spots for MALDI ionization were prepared by an air dry method, in which 400-850 nL of sample containing  $10^{-13}$ - $10^{-10}$  mol analyte was spotted onto a patterned steel substrate. Before the sample dried, 400-850 nL matrix was spotted on top of the sample and allowed to co-crystallize slowly at room temperature in a laminar flow hood. This spotting method lead to reproducible, visibly homogeneous spot formation and good analyte ionization, provided adequate dilution of concentrated high-salt samples had been performed. Ionization was performed with a pulsed UV laser, followed by mass analysis by TOF MS in high resolution reflectron mode.

#### **4.2.10 Homogeneous Luminescence Assays**

Two high throughput homogeneous (“mix-and-read”) assays were investigated as possible screening platforms for detecting phosphorylation of library compounds by EGFR and Syk (Figure 18). PerkinElmer’s LANCE and AlphaScreen assays are commonly employed in drug discovery laboratories for early-stage discovery and validation of clinical kinase inhibitors. Kits for the two assays were kindly provided by the vendor. Details of each assay mechanism are provided in the discussion sections below. Assays were performed according to vendor instructions. Briefly, positive and negative synthetic controls were prepared containing either extracellular buffer or Buffer H (for LANCE) or Buffer I (for AlphaScreen) and indicated concentrations of biotinylated TyrSub-9 (biotin-LEDDYEDDNle-NH<sub>2</sub>) or



its synthetically phosphorylated analog (biotin-LEDDpYEDDNle-NH<sub>2</sub>) as negative and positive controls, respectively. For LANCE detection, 60  $\mu$ L of assay sample was mixed with 190  $\mu$ L 1.3X detection mixture (1.3 nM Europium chelate-conjugated anti-phosphotyrosine antibody, 47 nM *ULight*-labeled streptavidin, 1.3X detection buffer) in a white opaque plastic 96-well plate. The plate was sealed and shaken 1 h at room temperature prior to analysis on a Molecular Devices SpectraMax M5 plate reader (Sunnyvale, CA) using optimized parameters. For AlphaScreen detection, assay samples were diluted 1:100 in AlphaScreen control buffer; 20  $\mu$ L of this mixture was dispensed into a PerkinElmer  $\frac{1}{2}$  Area Plate, followed by addition of 20  $\mu$ L 60  $\mu$ g/mL anti-phosphotyrosine-conjugated acceptor beads. Antibody beads were incubated with the assay sample 1 h or overnight at room temperature, followed by addition of biotinylated donor beads (20  $\mu$ L at 60  $\mu$ g/mL) and a 30 min incubation prior to analysis. Plates were read on a specialized EnVision plate reader (PerkinElmer, Waltham, MA) kindly provided by Dr. Bill Janzen in the UNC Drug Discovery Center. Steps including and following addition of the donor beads were carried out in a light-controlled environment to prevent photodegradation of the donor beads.

#### **4.2.11 Library Screen Data Analysis**

MALDI-TOF spectra of library screening assays were processed using MATLAB code modified from a custom program kindly provided by Dr. Ryan

Phillips. The code is provided in Appendix B. The input consists of exported peak intensities for each MALDI spot, and the corresponding masses, within the mass spectra that fall within user defined threshold values. For each spot, corresponding to one library peptide or control, MATLAB opens the appropriate file containing a matrix of masses and peak intensities and cross-references the mass vector with the expected mass of the library peptide, derived from a master file of MALDI spot codes and calculated peptide masses. MATLAB applies a mask to the expected mass, based on whether the spectra were acquired in negative or positive ion mode, to generate the search masses corresponding to the expected ionic species for the library peptide and its phosphorylated form. The program also accounts for neutral water loss and sodium and potassium adducts. MATLAB searches the mass vector for the selected masses within a user-defined tolerance, then simply sums the corresponding values in the intensity vector for the library peptide ions and the phosphorylated peptide ions. The amount of phosphorylated peptide is calculated as a ratio of phosphorylated to total signal and exported. This method was employed to quickly screen the library analyses for “hits” — defined as a peptide that was phosphorylated better than the control by the target kinase and more poorly than the control by the off-target kinase.

## 4.3 Results and Discussion

### 4.3.1 Selectivity of the Starting EGFR Peptide

The selectivity of 6FAM-TyrSub-9 for EGFR was evaluated using *in vitro* enzyme assays analyzed with quantitative CE-LIF. Assays were performed with EGFR, HER4, Syk, Src, and Abl tyrosine kinases as described in the Methods section. The results of those experiments are summarized in Figure 19. 6FAM-TyrSub-9 is phosphorylated appreciably by each of the kinases tested except Abl. This is not surprising given the similarity in consensus sequence for EGFR, Src, and Syk, which all prefer highly acidic substrates (Table 4). Measurements in complex biological matrices, particularly intact cells, with this peptide would suffer from the ambiguity of the protein target, making data analysis difficult. While tyrosine kinase inhibitors are available to suppress off-target kinase activities, none of the available tyrosine kinase inhibitors is completely selective for its target kinase; indeed, the vast majority of tyrosine kinase inhibitors act upon five or more off-target kinases, some with comparable affinity to the target kinase.<sup>40</sup> Therefore, TyrSub-9 is an excellent candidate for combinatorial selectivity optimization.

### 4.3.2 Comparison of Potential Library Screening Platforms

Three high-throughput phosphopeptide detection platforms (Figure 18) were investigated as potential screening assays for the combinatorial peptide optimization campaign: LANCE (PerkinElmer), AlphaScreen (PerkinElmer), and MALDI-TOF MS.

The three detection schemes were chosen for their sensitivity, amenability to parallel, high-throughput analyses, compatibility with available instrumentation, and time/cost efficiency. The bases of the detection platforms are described in detail below.

The LANCE assay is a homogeneous TR-FRET immunoassay that works as follows. A biotinylated peptide substrate is phosphorylated by the kinase of interest, then mixed with a low concentration of streptavidin labeled with a dye optimized for lanthanide-mediated TR-FRET. A Europium chelate-labeled anti-phosphotyrosine antibody binds only phosphorylated peptide, establishing a FRET pair. Stimulation of the Eu chelate at 615 nm within established FRET pairs results in excitation of the dye and emission at 655 nm, which can be easily monitored with low background using standard fluorescence plate readers. The LANCE assay is popular for high-throughput *in vitro* assays due to its mix-and-read format and compatibility with widely-available instrumentation. It is generally restricted to simple sample matrices (e.g. not biological) because endogenous fluorophores in biological samples can interfere with the TR-FRET and adversely affect sensitivity.

The AlphaScreen assay platform is a homogeneous proximity assay similar in design to the LANCE assay. Like LANCE, AlphaScreen employs labeled streptavidin and anti-phosphotyrosine antibodies to selectively label phosphorylated, biotinylated peptide substrate. However, rather than using FRET

labels, AlphaScreen uses chemically doped beads. The Donor bead releases singlet oxygen ( $^1\text{O}_2$ ) upon photoactivation at 680 nm. The singlet oxygen diffuses up to 200 nm to excite the Acceptor beads resulting in broadband emission at 520-620 nm. The emission is detected by specialized high sensitivity plate readers such as PerkinElmer's EnVision instruments. The AlphaScreen assay is very sensitive because the selective emission pathway is not susceptible to interference from endogenous fluorophores, resulting in very low background emission. For this reason, AlphaScreen is one of the few high throughput homogeneous assays to be successfully adapted to biological matrices.

Mass spectrometry is an increasingly popular alternative to luminescence-based high throughput screening assays. In particular, MALDI-TOF MS is well suited to rapid analyses of biomolecules; the soft ionization leads to easy to interpret mass spectra. MS is sensitive and highly versatile. Because it is a label-free technique, the number of analytes that may be monitored simultaneously is largely limited by the range and resolution of the mass analyzer. Sample preparation for MALDI involves co-crystallization of the analyte with a small organic matrix compound with a strong absorption band at the wavelength of the laser to be employed for ionization, usually in the UV. Upon irradiation of the co-crystal, matrix and analyte molecules desorb from the MALDI substrate to form a high-energy plume, where the gas phase molecules are ionized before being directed into the mass analyzer.

Ionization efficiencies vary widely among analytes, so quantitative analyses require an internal standard for each species to be quantified. Therefore, semi-quantitative analyses are better suited to high-throughput screens where synthesis of an internal standard for each compound is not feasible. In this case, semi-quantitative estimates of library peptide phosphorylation are sufficient to identify lead compounds for secondary quantitative verification.

#### **4.3.2 Luminescence Assay Validation with Synthetic Standards**

The luminescence-based commercial assay platforms were tested first as library screening assays due to the capability for quantitation and excellent compatibility with the assay sample matrix. Biotinylated non-phosphorylated and phosphorylated EGFR peptides were synthesized and purified as described in the Methods section above. The peptides were reconstituted in a high-salt buffer and analyzed by the luminescence assays either in buffer alone or in mock assay mixtures comprising all kinase assay components except enzyme. The response of the EGFR control peptides was compared to that of the positive control phosphorylated peptide LCK-P provided by the vendor. The results of those experiments are summarized in Figure 20. Figure 20A shows poor performance of the LANCE assay in detecting the positive EGFR control. Compared with the commercial positive control, the EGFR positive control is indistinguishable from the negative control. The performance of the AlphaScreen platform with the EGFR

peptides was superior, with the positive EGFR control showing similar response to the commercial positive control, though with less sensitivity at the low range of peptide concentration. It is difficult to surmise the reason for poor performance of the LANCE assay with the EGFR peptides. The low signal, indicating poor FRET efficiency, is possibly due to poor binding of the anti-phosphotyrosine antibody to the positive EGFR control, or to the distance between the donor and acceptor of the FRET pair being too large for efficient energy transfer. Given the good performance of the long-range (200 nm) AlphaScreen immunoassay, the second explanation is more likely.

#### **4.3.3 Effect of N-Terminal Label on Scaffold Phosphorylation by EGFR**

Given the excellent sensitivity of the AlphaScreen platform with regard to synthetic controls, its performance with real assay samples was explored. Detection of phosphorylated biotinylated EGFR peptide by AlphaScreen was validated in *in vitro* assays with recombinant EGFR kinase domain (Figure 21). A dose-response of AlphaScreen signal versus peptide concentration, compared to synthetically phosphorylated control, shows poor phosphorylation of the biotinylated peptide by EGFR. This result was unexpected given the excellent phosphorylation kinetics of the 6FAM-labeled version of this peptide. However, a comparison of biotinylated and 6FAM-labeled peptide phosphorylation by EGFR using MALDI-TOF MS shows a significant decrease in affinity of the biotinylated peptide for EGFR (Figure 22).

This result can be understood in the context of the preference of EGFR kinase domain for highly acidic substrates: 6FAM imparts an additional negative charge at physiologic pH, while biotin is uncharged. Furthermore, these data caution against the common assumption that the N-terminal labeling of peptide substrates does not play a role in enzyme-peptide recognition. Because poor phosphorylation of the biotinylated peptide by EGFR precludes use of the AlphaScreen assay as a detection platform for the combinatorial peptide optimization campaign, a hybrid approach was implemented in which MALDI-TOF MS was employed to first semi-quantitatively screen the 6FAM-labeled library for lead compounds, followed by quantitative validation using CE-LIF.

#### **4.3.4 MALDI-TOF for Phosphorylation Detection and Effect of Ionization Bias**

The effect of ionization bias on possible over- or under-estimation of extent of phosphorylation in the peptide library screen was evaluated by comparing MALDI-TOF MS analysis of *in vitro* EGFR assays with 6FAM-TyrSub-9, to quantitative CE-LIF of the same samples. CE-LIF analysis under the optimized separation conditions provides a bias-free ratio of phosphorylated to total reporter in the EGFR assay samples, to which the same parameters derived from MALDI-TOF spectra were compared (Figure 23). The head-to-head comparison of the two methods shows that MALDI-TOF slightly underestimates phosphorylation of 6FAM-TyrSub-9 compared to the CE-LIF values, revealing a small relative decrease in MALDI efficiency of the



peptide phosphoform under the analysis conditions used. The addition of two negative charges imparted by the phosphoryl moiety probably account for this difference. However, since the efficiency differential is small, its effect on assay sensitivity is manageable.

#### **4.3.5 Lysine Scan to Determine Library Site**

As discussed in detail above, the basis of this combinatorial optimization campaign is a semi-rational site-directed coupling of synthetic carboxylic acids to the side chain of a lysine residue to generate a library of peptides containing novel, structurally diverse non-native amino acids at the chosen modification site. Before a full-scale library was prepared, the optimal modification site within the starting 6FAM-labeled EGFR peptide was determined. Potential modification sites were identified at the three positions immediately C-terminal to the tyrosyl phosphorylation site. A comparison of the recognition motifs of EGFR and other tyrosine kinases reveals a nearly universal preference for anionic residues in the Y-N (N-terminal) region proximal to the phosphorylation site. In the Y+N region, greater diversity in terms of side chain polarity and size suggests that this region is preferable for coupling structurally diverse, largely cyclic carboxylic acids comprising the carboxylic acid library.

One modification site was chosen among the three positions immediately C-terminal to the tyrosine by evaluating the effect on EGFR-catalyzed phosphorylation

of substituting acetylated lysine (K(Ac)) at each site. Lysine acetylation is one of the chemically and structurally simplest lysine acylations possible and as such represents the minimum perturbation to the starting EGFR peptide. If substitution of K(Ac) at a given site is not tolerated by EGFR, then that site is not suitable to provide a scaffold for combinatorial optimization. EGFR phosphorylation of the three K(Ac) peptides is summarized in Figure 24. All K(Ac) peptides were phosphorylated more poorly than the starting peptide, an effect probably based on decreased acidity of the sequence. Sites Y+1 and Y+3 affected substrate phosphorylation similarly and to a lesser extent than Y+2. Since many of the carboxylic acids in the compound library are large, bulky polycyclic compounds, the Y+3 site was chosen for combinatorial optimization in order to limit steric hindrance of phosphoryl transfer to the tyrosine. Therefore, the library scaffold was synthesized with the sequence: 6FAM-LEDDYEDK(Mtt)Nle-CONH-(CH<sub>2</sub>)<sub>2</sub>-S-S-cystamine resin.

#### **4.3.6 Library Synthesis and Characterization**

The peptide library was synthesized as described in detail in the methods section above. Following cleavage and isolation of the library in DTT kinase buffer, acylation efficiency at the library site was assessed by MALDI-TOF MS (Figure 25). MALDI-TOF characterization of the library revealed a major impurity in each library peptide, with a mass corresponding to that of the unmodified scaffold minus a neutral loss of water. This mass is consistent with internal cyclization of the library

site lysine with the adjacent aspartic acid residue, which was made possible by deprotection of aspartic and glutamic acid residues during Mtt removal, as discussed in depth above in the Methods section. Internal cyclization competed with lysine acylation with the carboxylic acid library; the amount of cyclized side product detected varied based on acylation kinetics of a given library compound. This was further verified by RP-HPLC-MS analysis of selected library peptides (Figure 26). Despite the highly variable purity across the peptide library, it was possible to screen the library for phosphorylation by target and off-target enzymes because universal detection with mass spectrometry allowed global monitoring of both desired product and side-product modification by the enzymes. Re-synthesis of the peptide library with a more orthogonal protecting group such as ivDde at the modification site would greatly improve purities across the library and possibly reveal more or different lead compounds for enzymes of interest; however, such an effort was not within the scope of this project.

#### **4.3.7 Library Screening and Data Analysis**

The peptide library was screened for on-target phosphorylation by EGFR and for off-target phosphorylation by Syk, a non-receptor tyrosine kinase known to phosphorylate the starting EGFR peptide with similar kinetics to EGFR. Syk is a drug target in immune cell-derived cancers and autoimmune diseases such as rheumatoid arthritis. MALDI-TOF MS spectra of each library peptide were obtained

before and after phosphorylation by EGFR and Syk. The fraction of phosphorylated peptide was estimated using MATLAB as discussed in detail above in the Methods section. The results of those screens are summarized in Figure 27. The fraction of phosphorylated peptide is given normalized to the acetylated control 6FAM-LEDDYEDK(Ac)Nle-CONH-(CH<sub>2</sub>)<sub>2</sub>-SH. In the case of EGFR, no peptide was phosphorylated better than the control (normalized value >1). However, the screen against Syk produced a hit peptide “4-10F” (Figure 28) that was phosphorylated approximated 3-fold better than the control and was not detectably phosphorylated by EGFR. While EGFR was the primary target of this screen, the potential hit substrate for Syk was pursued for its potential utility in single-cell analysis of response to clinical Syk inhibition.

#### **4.3.8 Library Hit Re-Synthesis and Validation**

High-yield synthesis of the compound 4-10F proved to be highly challenging. Despite replacement of lysine(Mtt) for lysine(ivDde) at the library site, allowing fully orthogonal and selective deprotection at that site using hydrazine, persistent formation of a racemic side product was detected during quantitative CE-LIF validation. The side product is of the same mass as the desired peptide and is resolvable from the desired peptide by CE but not by analytical RP-HPLC, strongly suggesting that the side product is a  $\beta$ -isoform resulting from aspartimide formation. In the presence of acid or base, protected aspartic acid residues can cyclize with the

amide nitrogen of the peptide backbone, eliminating the side-chain protecting group. Nucleophilic opening of the aspartimide ring yields a mixture of  $\alpha$ - and  $\beta$ -peptide (Figure 29) that is indistinguishable by mass and difficult to separate using standard liquid chromatography. Aspartimide formation is among the best documented side reactions in solid phase synthesis of peptides and is commonly mediated by the piperidine deprotection reagent used to remove N-terminal Fmoc-residues.<sup>41,42</sup> Repeated piperidine treatments over the course of a synthesis lead to increases in aspartimide formation at each step and make total suppression of aspartimide formation extremely challenging.<sup>43</sup> Several strategies are available for reduction of aspartimide formation in susceptible sequences. The most common is to supplement the piperidine deprotection solution with a weak organic acid, most often HOBT, which partially suppresses deprotonation of the peptide backbone amide<sup>44</sup>. Alternatively, piperidine can be replaced by a weaker base such as piperazine<sup>45</sup>, but suppression of aspartimide formation is balanced by slow deprotection kinetics and possible related side reactions due to long treatment times. Piperidine treatment times can be minimized by using microwave synthesis, where deprotection times can be as short as 1-2 min. This approach is not suitable for all sequences, as microwave synthesis is known to facilitate other common side reactions in SPPS. Protection of the aspartic acid side chain with bulky group such as OMpe<sup>41</sup> has been shown to discourage aspartimide formation, but OMpe-protected aspartic acid is

prohibitively expensive for many syntheses, particularly those on a large scale. Protection of the backbone amide with an Hmb group has been shown to completely eliminate aspartimide formation in most cases<sup>46</sup>, but Hmb-protected reagents must be purchased as expensive dipeptides that are available in a limited range of compositions. Minimizing contaminating water in synthesis reagents is particularly important for labs located in humid climates, as excess water facilitates aspartimide formation. Ensuring that frozen reagents are fully at room temperature before opening and sequentially washing resin with solvents of decreasing hydrophilicity to exclude excess water will greatly aid in managing aspartimide formation in susceptible sequences. Several of these approaches were implemented to mitigate aspartimide formation in the synthesis of 4-10F. A combination of careful control of water content during synthesis, HOBt supplementation during deprotection, use of OMpe-protected aspartic acid, and microwave synthesis was necessary to meaningfully decrease the amount of  $\beta$ -peptide observed by CE-LIF. However, despite extensive purification, an unidentified side product appearing only in CE-LIF analyses of enzyme assays of the peptide continues to complicate quantitative validation of the peptide and limits its present utility for further biochemical studies. Future work should include a more detailed structural characterization of the peptide and its modification by EGFR and Syk, including nano-LC-MS of assay samples. A potential route to achieve superior purification of the reporter at large

scale would be supercritical fluid chromatography, a technique popular in industrial labs for conducting preparative scale high efficiency, highly selective separations of structurally similar compounds that is increasingly being adapted to biomolecule separations and can be implemented using standard HPLC pumps.<sup>47,48</sup>

CE-LIF analysis of 4-10F phosphorylation by recombinant Syk (Figure 30) shows formation of a time- and ATP-dependent peak that likely is due to phosphorylated 4-10F. However, an impurity closely migrating with the major peak which is indistinguishable by MS interferes with straightforward interpretation of the data. This impurity is probably  $\beta$ -peptide whose formation could not be completely suppressed despite all previous efforts. Further efforts to purify 4-10F to remove  $\beta$ -peptide will be necessary, as discussed above.

#### **4.4 Conclusions and Future Directions**

This work has laid the groundwork for a future combinatorial campaign to generate optimized EGFR substrate(s) by optimizing synthetic parameters and validating screening platforms. Future campaigns should follow the optimized synthetic guidelines for peptide scaffold preparation and should explore positions N-terminal to the phosphorylation site as possible modification sites. Future campaigns should also expand selectivity screening to include more receptor and non-receptor tyrosine kinases.

The library screening platform could be improved further by drawing

inspiration from the work of the Mrksich lab at Northwestern University.<sup>49,50</sup>

Researchers in that lab have developed a specialized MALDI technique in which analytes are immobilized to a gold MALDI substrate using a gold-thiol self-assembled monolayer (SAM) technique. The analytes are immobilized in a background of triethylene glycol-modified alkanethiols which prevent adsorption of biomolecules and allows assays to be performed directly on the MALDI substrate. Spots can be washed prior to ionization to remove salts, biomolecules, and other interferents to maximize ionization efficiency and assay sensitivity. This technique, known as SAMDI, would be a particularly interesting addition to the peptide library screening workflow, as the library peptides already are functionalized at the C-terminus with a free thiol that could be used to immobilize the peptide on the MALDI substrate. Figure 31 outlines a proposed workflow for such a detection scheme. A PEG spacer could be introduced between the cystamine resin linkage and the C-terminal amino acid of the library peptide to promote even, dense packing of the SAM and to allow the enzyme interaction sites within the peptide to extend beyond the background triethylene glycol SAM for easy access to the enzyme active site. Use of SAMDI as a screening platform would also enable screening in complex biological matrices such as cell lysates, where peptide selectivity could be evaluated using selective stimulators and inhibitors of enzyme activity, including susceptibility to proteolysis and dephosphorylation.



Generation of selective substrates for tyrosine kinases remains a daunting challenge that is unlikely to be completely solved by any one approach. A toolbox of well-characterized complementary strategies to improve substrate selectivity is highly desirable. Because peptide substrates have traditionally played a limited role in analyses of enzyme activity in complex matrices, work to increase substrate selectivity has been target focused rather than geared toward comprehensive suites of solutions. The combinatorial optimization approach outlined in this work is highly versatile and modular and is not limited to the tyrosine kinases used here. For these reasons, this technology has the potential to comprise an important part of a comprehensive selectivity optimization toolbox

## REFERENCES

- (1) Pinna, L. A.; Ruzzene, M. *Biochim. Biophys. Acta* **1996**, *1314*, 191–225.
- (2) Ubersax, J. A.; Ferrell, J. E., Jr. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 530–541.
- (3) Miller, W. T. *Acc. Chem. Res.* **2003**, *36*, 393–400.
- (4) Kreegipuu, A.; Blom, N.; Brunak, S.; Järv, J. *FEBS Lett.* **1998**, *430*, 45–50.
- (5) Lemmon, M. A.; Schlessinger, J. *Cell* **2010**, *141*, 1117–1134.
- (6) Baselga, J. *Science* **2006**, *312*, 1175–1178.
- (7) Blume-Jensen, P.; Hunter, T. *Nature* **2001**, *411*, 355–365.
- (8) Eichholtz, T.; de Bont, D. B.; de Widt, J.; Liskamp, R. M.; Ploegh, H. L. *J. Biol. Chem.* **1993**, *268*, 1982–1986.
- (9) Litman, P.; Ohne, O.; Ben-Yaakov, S.; Shemesh-Darvish, L.; Yechezkel, T.; Salitra, Y.; Rubnov, S.; Cohen, I.; Senderowitz, H.; Kidron, D.; Livnah, O.; Levitzki, A.; Livnah, N. *Biochemistry* **2007**, *46*, 4716–4724.
- (10) Delehanty, J. B.; Bradburne, C. E.; Boeneman, K.; Susumu, K.; Farrell, D.; Mei, B. C.; Blanco-Canosa, J. B.; Dawson, G.; Dawson, P. E.; Mattoussi, H.; Medintz, I. L. *Integrative Biology* **2010**, *2*, 265–277.
- (11) Zhang, X.-X.; Eden, H. S.; Chen, X. *Journal of Controlled Release* **2012**, *159*, 2–13.
- (12) Koren, E.; Torchilin, V. P. *Trends in Molecular Medicine* **2012**, *18*, 385–393.
- (13) Stewart, K. M.; Horton, K. L.; Kelley, S. O. *Org. Biomol. Chem.* **2008**, *6*, 2242–2255.
- (14) Dang, C. V.; Lee, W. M. *J. Biol. Chem.* **1989**, *264*, 18019–18023.
- (15) Diamond, S. L.; Subramanian, A.; Ranganathan, P. *Nature Biotechnology* **1999**, *17*, 873–877.
- (16) Goldfarb, D. S.; Gariépy, J.; Schoolnik, G.; Kornberg, R. D. *Nature* **1986**, *322*,

641–644.

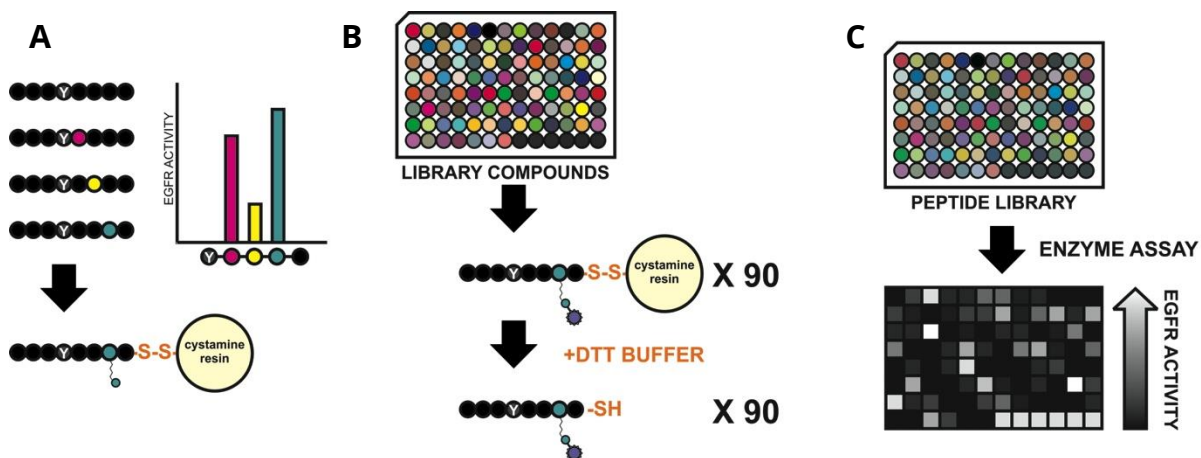
- (17) Horton, K. L.; Stewart, K. M.; Fonseca, S. B.; Guo, Q.; Kelley, S. O. *Chemistry & Biology* **2008**, *15*, 375–382.
- (18) Yousif, L. F.; Stewart, K. M.; Horton, K. L.; Kelley, S. O. *ChemBioChem* **2009**, *10*, 2081–2088.
- (19) Yousif, L. F.; Stewart, K. M.; Kelley, S. O. *ChemBioChem* **2009**, *10*, 1939–1950.
- (20) Chamberlain, G. R.; Tulumello, D. V.; Kelley, S. O. *ACS Chem. Biol.* **2013**, *8*, 1389–1395.
- (21) Souhayer, J. S.; Wang, Y.; Li, H.; Cheung, S.-H.; Rossi, F. M.; Stanbridge, E. J.; Sims, C. E.; Allbritton, N. L. *Biochemistry* **2004**, *43*, 8528–8540.
- (22) Nelson, A. R.; Borland, L.; Allbritton, N. L.; Sims, C. E. *Biochemistry* **2007**, *46*, 14771–14781.
- (23) Lemmon, M. A.; Schlessinger, J.; Ferguson, K. M. *Cold Spring Harbor perspectives in biology* **2014**, *6*, a020768.
- (24) Antczak, C.; Bermingham, A.; Calder, P.; Malkov, D.; Song, K.; Fetter, J.; Djaballah, H. *Assay Drug Dev. Technol.* **2012**, *10*, 24–36.
- (25) Xiang, X.; Sun, J.; Wu, J.; He, H.-T.; Wang, Y.; Zhu, C. *Cel. Mol. Bioeng.* **2011**, *4*, 670–677.
- (26) Morris, M. C. *Biochim. Biophys. Acta* **2013**, *1834*, 1387–1395.
- (27) González-Vera, J. A. *Chem. Soc. Rev.* **2012**, *41*, 1652–1664.
- (28) Fernandes, N.; Bailey, D. E.; VanVranken, D. L.; Allbritton, N. L. *ACS Chem. Biol.* **2007**, *2*, 665–673.
- (29) Priestman, M. A.; Lawrence, D. S. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2010**, *1804*, 547–558.
- (30) *Combinatorial Peptide Library Protocols*; Cabilly, S., Ed.; Humana Press: Totowa, New Jersey, 1997; Vol. 87.

- (31) Adams, J. A. *Chem. Rev.* **2001**, *101*, 2271–2290.
- (32) Lawrence, D. S.; Wang, Q. *ChemBioChem* **2007**, *8*, 373–378.
- (33) Proctor, A.; Wang, Q.; Lawrence, D. S.; Allbritton, N. L. *Anal. Chem.* **2012**, *84*, 7195–7202.
- (34) Donella-Deana, A.; Ruzza, P.; Cesaro, L.; Brunati, A. M.; Calderan, A.; Borin, G.; Pinna, L. A. *FEBS Lett.* **2002**, *523*, 48–52.
- (35) Lee, J. H.; Nandy, S. K.; Lawrence, D. S. *J. Am. Chem. Soc.* **2004**, *126*, 3394–3395.
- (36) Engel, K.; Sasaki, T.; Wang, Q.; Kuriyan, J. *Biochem. J.* **2013**, *453*, 337–344.
- (37) Guyer, C. A.; Woltjer, R. L.; Coker, K. J.; Staros, J. V. *Arch. Biochem. Biophys.* **1994**, *312*, 573–578.
- (38) Phillips, R. M. Development of a Novel Assay of Protein Tyrosine Phosphatase Activity in Single Cells Using Capillary Electrophoresis. Ph.D. Thesis, The University of North Carolina at Chapel Hill, Chapel Hill, NC, 2013.
- (39) Phillips, R. M.; Bair, E.; Lawrence, D. S.; Sims, C. E.; al, E. *Anal. Chem.* **2013**, *85*, 6136–6142.
- (40) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. *Nature Biotechnology* **2008**, *26*, 127–132.
- (41) Mergler, M.; Dick, F.; Sax, B.; Weiler, P.; Vorherr, T. *J. Peptide Sci.* **2003**, *9*, 36–46.
- (42) Lauer, J. L.; Fields, C. G.; Fields, G. B. *Lett Pept Sci* **1995**, *1*, 197–205.
- (43) Ruczyński, J.; Lewandowska, B.; Mucha, P.; Rekowski, P. *J. Peptide Sci.* **2008**, *14*, 335–341.

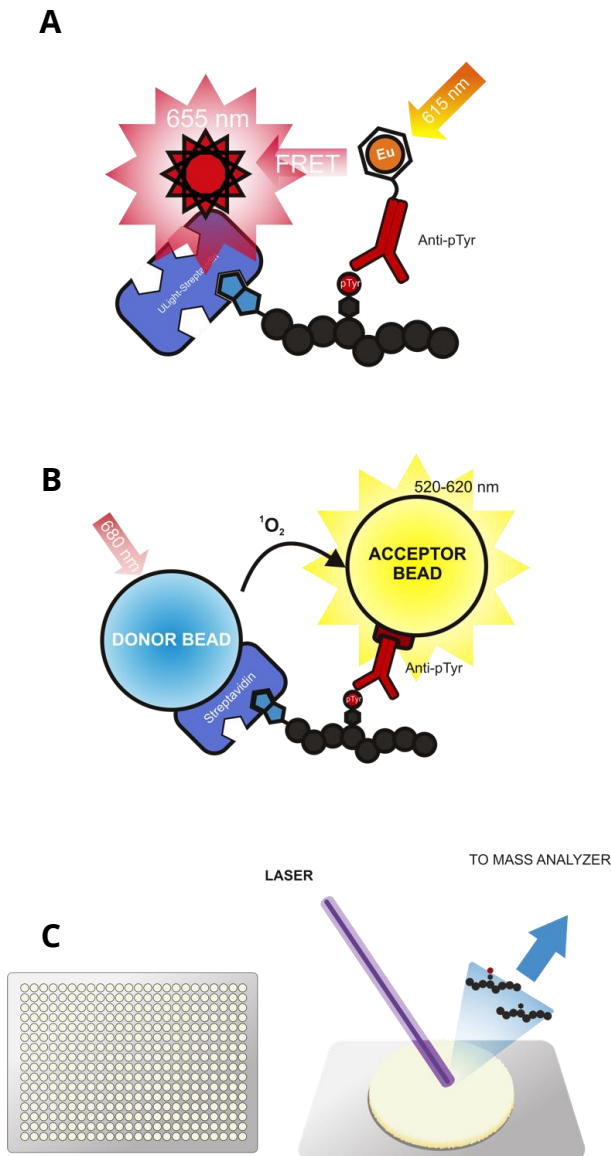
- (44) Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W. *Org. Lett.* **2012**, *14*, 5218–5221.
- (45) Wade, J. D.; Mathieu, M. N.; Macris, M.; Tregear, G. W. *Int J Pept Res Ther* **2000**, *7*, 107–112.
- (46) Mergler, M.; Dick, F. J. *Peptide Sci.* **2005**, *11*, 650–657.
- (47) Mukhopadhyay, R. *Anal. Chem.* **2008**, *80*, 3091–3094.
- (48) Taylor, L. T. *Anal. Chem.* **2010**, *82*, 4925–4935.
- (49) Kuo, H.-Y.; DeLuca, T. A.; Miller, W. M.; Mrksich, M. *Anal. Chem.* **2013**, *85*, 10635–10642.
- (50) Ban, L.; Pettit, N.; Li, L.; Stuparu, A. D.; Cai, L.; Chen, W.; Guan, W.; Han, W.; Wang, P. G.; Mrksich, M. *Nat Chem Biol* **2012**, *8*, 769–773.

**Table 4 Protein Tyrosine Kinase Consensus Sequences; red: anionic; blue: cationic; purple: polar uncharged; grey: hydrophobic; bold: phosphosite; X: any amino acid.**

<b>Kinase</b>	<b>Consensus Sequence</b>	<b>Reference</b>
EGFR	X- <b>E-E-E</b> -Y-F-X-X-X	Ubersax, <i>et al.</i> 2007
SRC	X- <b>E-E-I</b> -Y- <b>E</b> /G-X-F-X	Ubersax, <i>et al.</i> 2007
ABL	X-X-X-I/V/L- <b>Y</b> -X-X-P/F-X	Ubersax, <i>et al.</i> 2007
SYK	<b>E-D-E-D</b> -Y- <b>E-S</b> -V-G	Deng, <i>et al.</i> 2014
IRK	X-X-X-X- <b>Y</b> -M-M-M	Ubersax, <i>et al.</i> 2007
CSK	<b>D-E-Q</b> -I- <b>Y-W-Q</b> -I-A	Deng, <i>et al.</i> 2014
LYN	<b>D-E-D</b> -I- <b>Y</b> -G-V-L-P	Deng, <i>et al.</i> 2014
JAK	<b>D-P-Q-E</b> -Y-I-P-L-P	Deng, <i>et al.</i> 2014
BMX	<b>E-E-P</b> -I- <b>Y</b> -I-I-V-P	Deng, <i>et al.</i> 2014
FRK	<b>R-D</b> -V-I- <b>Y</b> -I-I-I-P	Deng, <i>et al.</i> 2014
PYK	<b>H-P-I-I</b> - <b>Y-pT</b> -I-I- <b>E</b>	Deng, <i>et al.</i> 2014
FES	<b>D-E-E-pY</b> - <b>Y-E</b> -I-I-A	Deng, <i>et al.</i> 2014
ACK	<b>R-A-F-I</b> - <b>Y</b> -A-I-I-P	Deng, <i>et al.</i> 2014

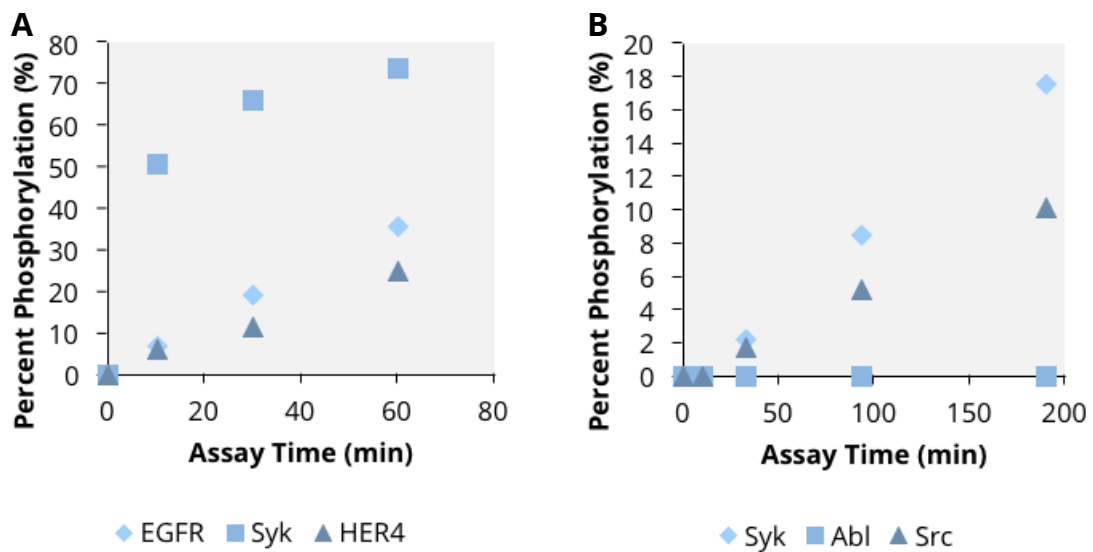


**Figure 17. Library screening workflow.** A) The optimal modification site within the starting peptide is selected by monitoring the effect of substituting acetylated lysine at various positions and comparing the measured EGFR activity with respect to each peptide. B) The library scaffold peptide containing the chosen modification site is acylated with a library of 90 unique carboxylic acids in a 96-well synthesis plate, followed by cleavage of the 90 library peptides in DTT kinase buffer. C) The 90-member peptide library is screened for phosphorylation by EGFR using *in vitro* kinase assays.

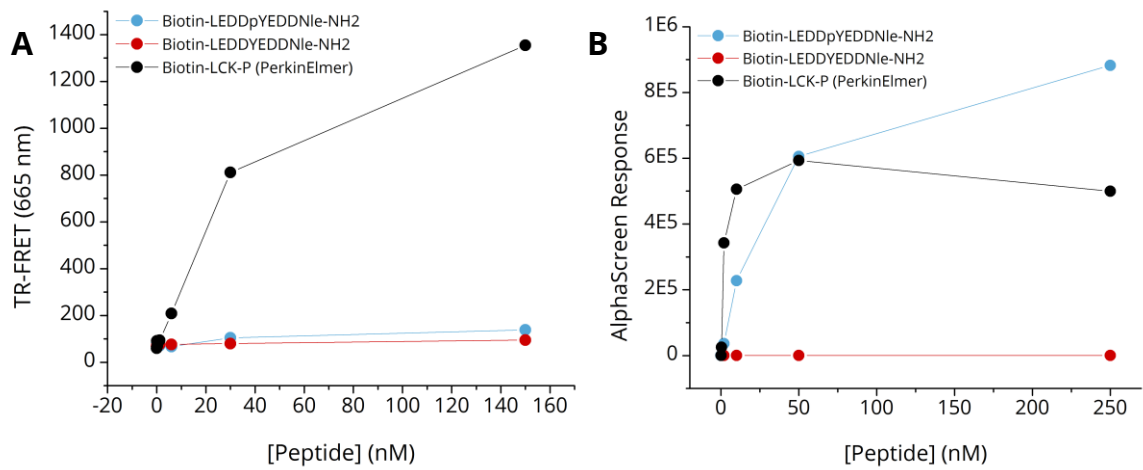


**Figure 18. Summary of library screening platforms.** A) LANCE TR-FRET: Eu chelate-labeled anti-phosphotyrosine antibody binds phosphorylated peptide with N-terminal biotin label, followed by binding of *ULight* dye labeled streptavidin. Excitation at 615 nm stimulates FRET and emission by *ULight* dye at 655 nm. B) AlphaScreen: Donor bead-labeled streptavidin and Acceptor bead labeled anti-phosphotyrosine bind phosphorylated biotinylated peptide substrates. Excitation of the Donor beads at 650 nm stimulates singlet oxygen emission, which stimulates broadband emission between 520-620 nm. C) MALDI-TOF MS: Co-crystallization of phosphorylated and non-phosphorylated peptides with UV-absorbing matrix is followed by desorption of matrix and analytes by irradiation with a pulsed UV laser. Desorbed molecules ionize and are directed to the TOF mass analyzer, where phosphorylated and non-phosphorylated peptide are detected.

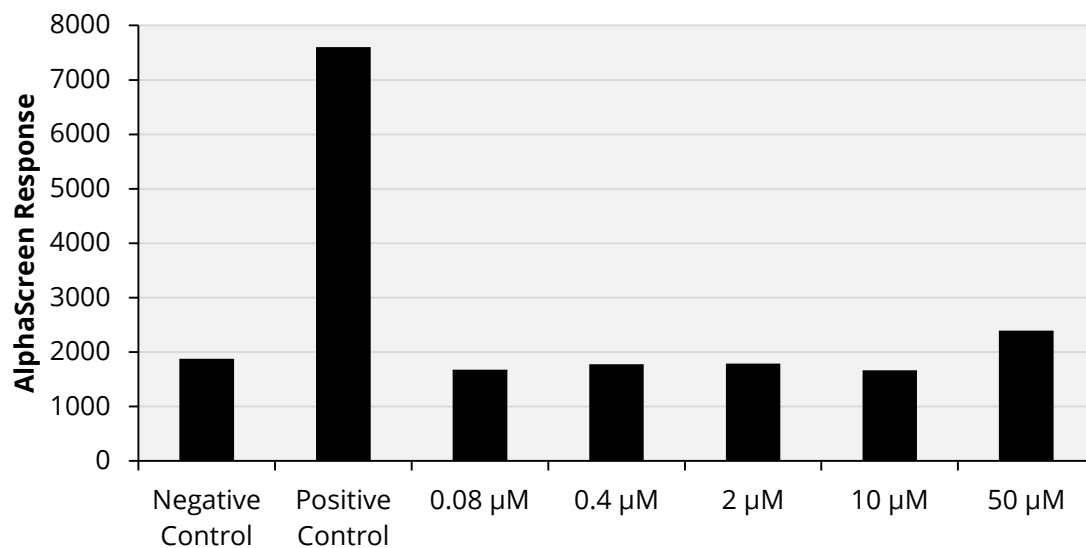




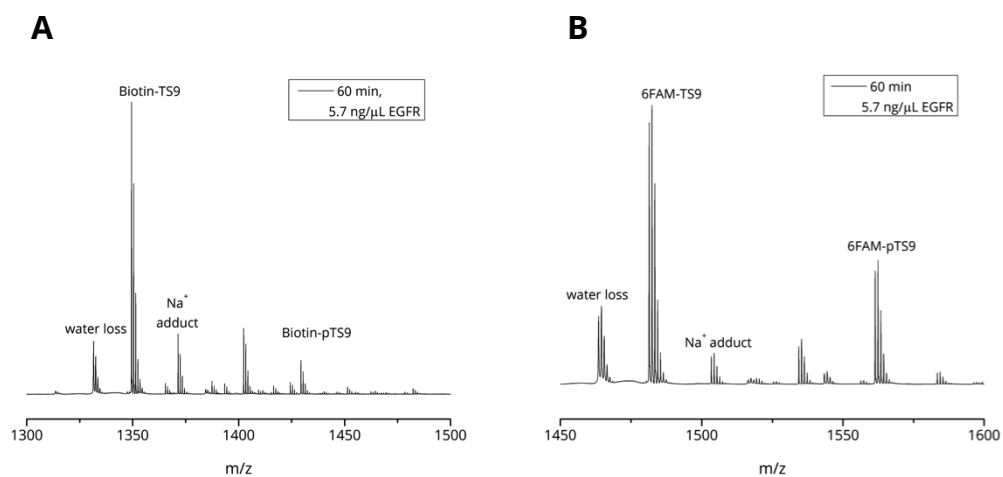
**Figure 19. Time courses of 6FAM-TyrSub-9 phosphorylation by a panel of tyrosine kinases.** A) 6FAM-TyrSub-9 (10  $\mu$ M) was incubated with recombinant EGFR (4 ng/ $\mu$ L), Syk (2 ng/ $\mu$ L), or HER4 (10 ng/ $\mu$ L) for the indicated times and percent phosphorylation quantified by CE-LIF. B) 6FAM-TyrSub-9 (5  $\mu$ M) was incubated with Syk (0.5 ng/ $\mu$ L), Abl (1.6 ng/ $\mu$ L), or Src (1.3 ng/ $\mu$ L) for the indicated times and percent phosphorylation quantified by CE-LIF. N = 1.



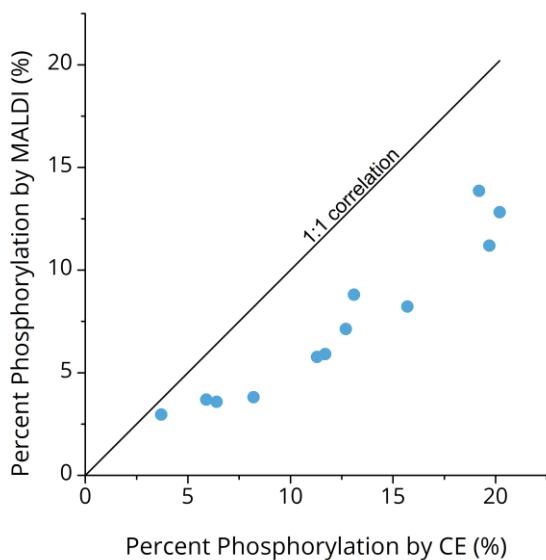
**Figure 20. Comparison of phosphorylated Biotin-TyrSub-9 detection by LANCE TR-FRET (A) and AlphaScreen (B) under optimized conditions. AlphaScreen, but not LANCE, detects phospho-Biotin-TyrSub-9 comparably to the commercial control. N = 1.**



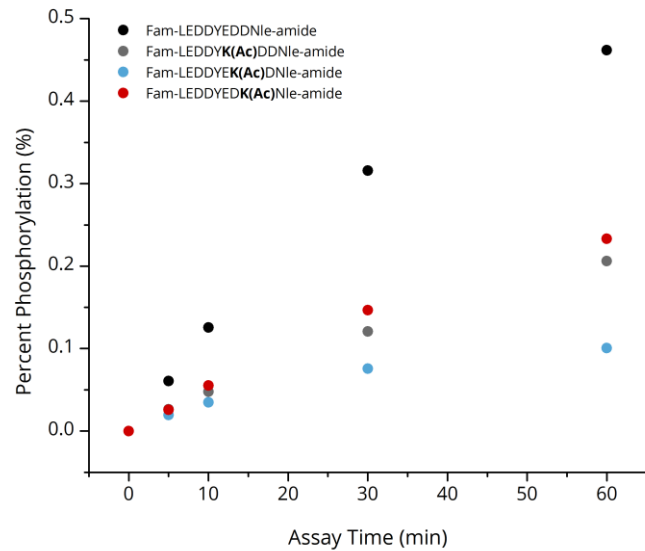
**Figure 21. AlphaScreen detection of *in vitro* phosphorylation of Biotin-TyrSub-9 by EGFR.** A small increase in phosphorylation is detectable by AlphaScreen, but sensitivity to small changes in phosphorylation is very poor and very little phosphorylated peptide is detected overall. N = 1.



**Figure 22. MALDI-TOF MS analysis of *in vitro* EGFR assays with Biotin-TyrSub-9 (Biotin-TS9, A) and 6FAM-TyrSub-9 (6FAM-TS9, B). The biotinylated peptide is phosphorylated significantly more poorly than the 6FAM-labeled peptide.**



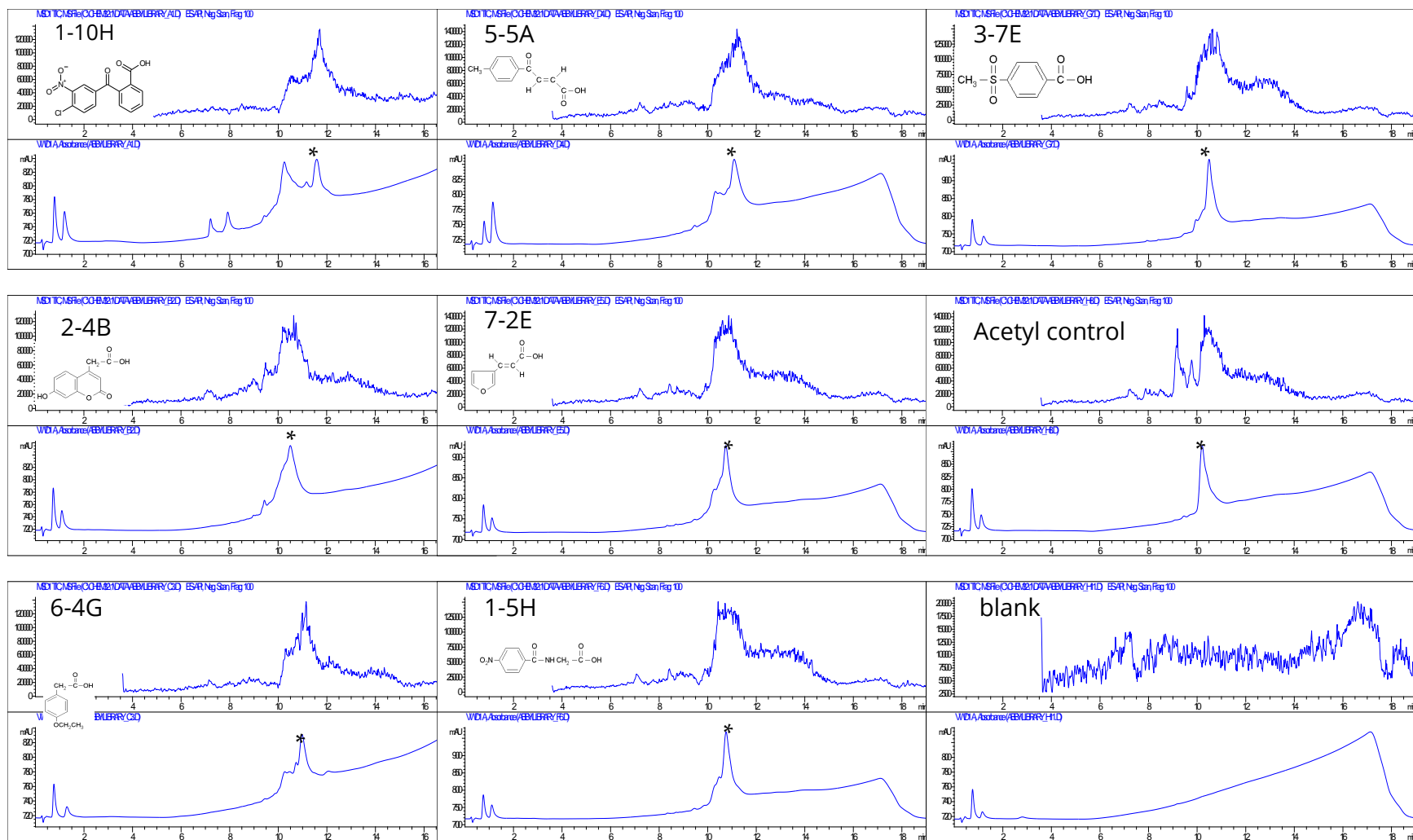
**Figure 23. Effect of ionization efficiency differential on measurement bias in MALDI-TOF MS.** Assays of EGFR-catalyzed phosphorylation of 6FAM-TyrSub-9 were analyzed by MALDI-TOF MS (y-axis) and quantitative CE-LIF (x-axis). MALDI-TOF MS underestimates the percentage of phosphorylated peptide in every sample in a well-defined linear manner. N = 1



**Figure 24. Site scan to determine optimal library optimization site within the starting EGFR peptide.** The starting EGFR peptide TyrSub-9 (FAM-LEDDYEDDNle-amide) and three acetylated lysine (K(Ac)) substituted peptides were assayed for phosphorylation by recombinant EGFR. Percent phosphorylation was quantified by CE-LIF. N = 1.

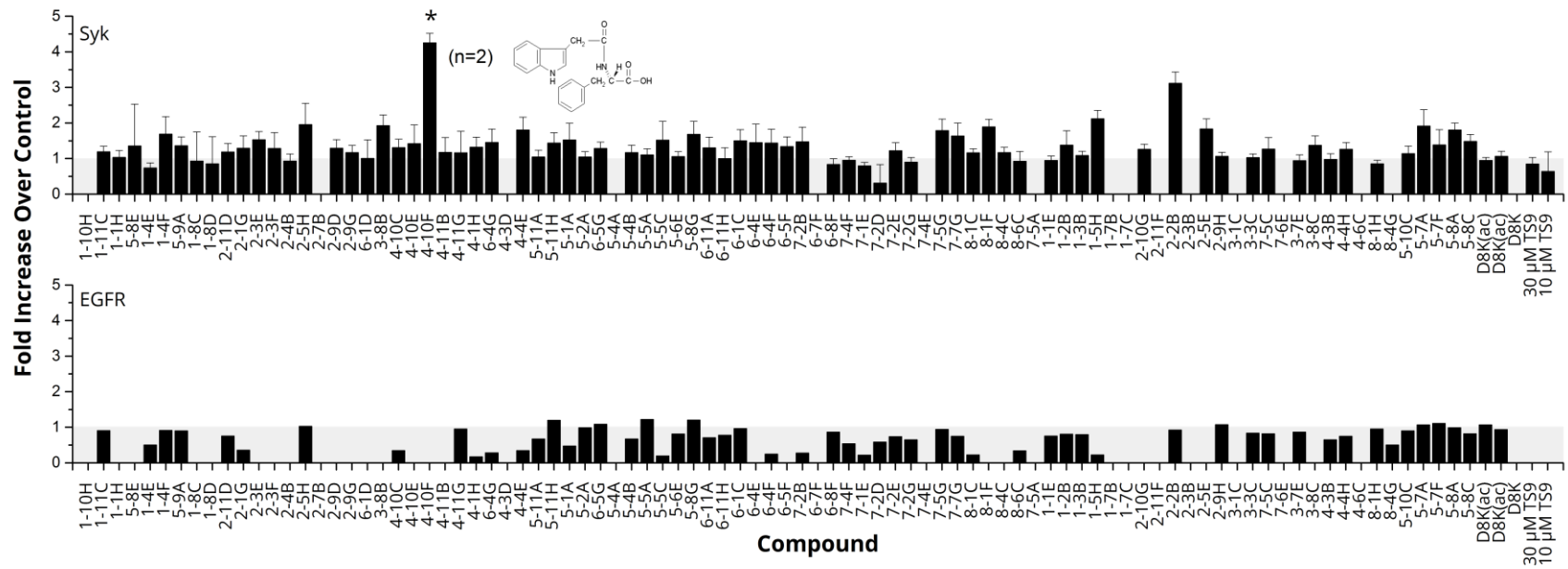
Spot	%Intensity	%Yield	Spot	%Intensity	%Yield	Spot	%Intensity	%Yield	Spot	%Intensity	%Yield	Key
B1	1.0	3.2	C1	22.4	61.6	D1	7.0	28.8	E1	38.9	80.3	0
B2	6.6	41.7	C2	18.7	46.6	D2	7.2	38.8	E2	29.7	80.9	25
B3	11.0	38.9	C3	24.1	55.6	D3	7.6	28.9	E3	7.1	25.2	50
B4	8.2	19.2	C4	4.5	23.4	D4	8.1	31.5	E4	11.5	35.7	75
B5	58.2	86.2	C5	2.8	10.2	D5	25.8	61.8	E5	12.1	43.7	100
B6	15.9	51.9	C6	62.1	88.1	D6	11.3	44.7	E6	5.7	25.3	
B7	15.7	74.3	C7	15.7	67.3	D7	1.0	3.9	E7	21.9	74.1	
B8	14.1	44.5	C8	8.5	42.5	D8	14.5	31.0	E8	12.2	42.1	
B9	5.7	23.7	C9	9.6	30.1	D9	18.5	62.2	E9	13.2	46.9	
B10	23.1	54.2	C10	18.6	65.7	D10	10.2	39.8	E10	7.6	34.7	
B11	8.1	50.2	C11	5.5	20.6	D11	4.2	24.5	E11	8.7	31.0	
B12	12.4	31.5	C12	10.8	48.8	D12	4.0	10.4	E12	17.8	59.2	
B13	8.1	28.9	C13	13.9	47.8	D13	11.0	45.1	E13	4.7	16.8	
B14	6.4	32.6	C14	17.5	52.5	D14	0.2	0.4	E14	10.9	51.0	
B15	23.3	60.7	C15	10.0	46.5	D15	9.2	37.7	E15	14.2	53.3	
B16	3.1	16.2	C16	3.6	14.1	D16	29.6	77.9	E16	23.5	63.2	
B17	12.7	45.3	C17	4.9	20.5	D17	17.8	64.9	E17	20.4	58.6	
B18	8.2	31.1	C18	7.2	30.6	D18	9.2	48.8	E18	7.4	27.4	
B19	8.7	34.0	C19	24.9	63.2	D19	7.7	33.8	E19	61.4	70.3	
B20	13.6	51.1	C20	7.9	29.2	D20	14.8	39.5	E20	62.8	70.0	
B21	8.4	33.2	C21	10.9	39.3	D21	4.9	20.7	E21	26.2	50.0	
B22	3.9	20.1	C22	6.3	26.0	D22	6.2	22.1	E22	0.0	0.0	
B23	4.7	14.0	C23	9.2	50.5	D23	14.6	42.8				
B24	6.3	30.0	C24	2.0	5.4	D24	0.8	5.1				

**Figure 25. Results of the MALDI-TOF screen of the EGFR peptide library.** The percent intensity is equal to the signal collected corresponding to the expected mass for each compound divided by the total signal measured from all species. The percent yield is equal to the signal corresponding to the expected mass divided by the signal corresponding to the unreacted library scaffold. The values are color coded by intensity according to the key provided. N = 1.

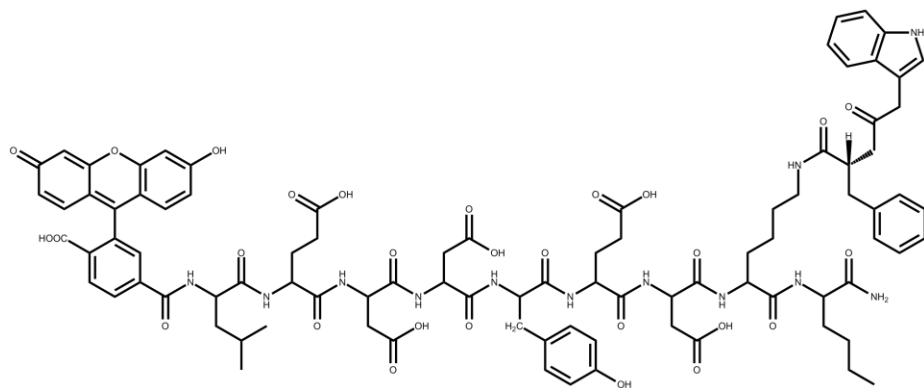


**Figure 26. RP-HPLC-MS analysis of selected library compounds.** The compound code and structure are overlaid on each chromatogram. Stars (\*) indicate peaks corresponding to correctly synthesized products, determined by ESI-MS. For each compound, the top chromatogram is a plot of total ion count vs. time, acquired on an ESI-single quadrupole MS in negative ion mode. The bottom chromatogram is a plot of absorbance at 240 nm vs. time.

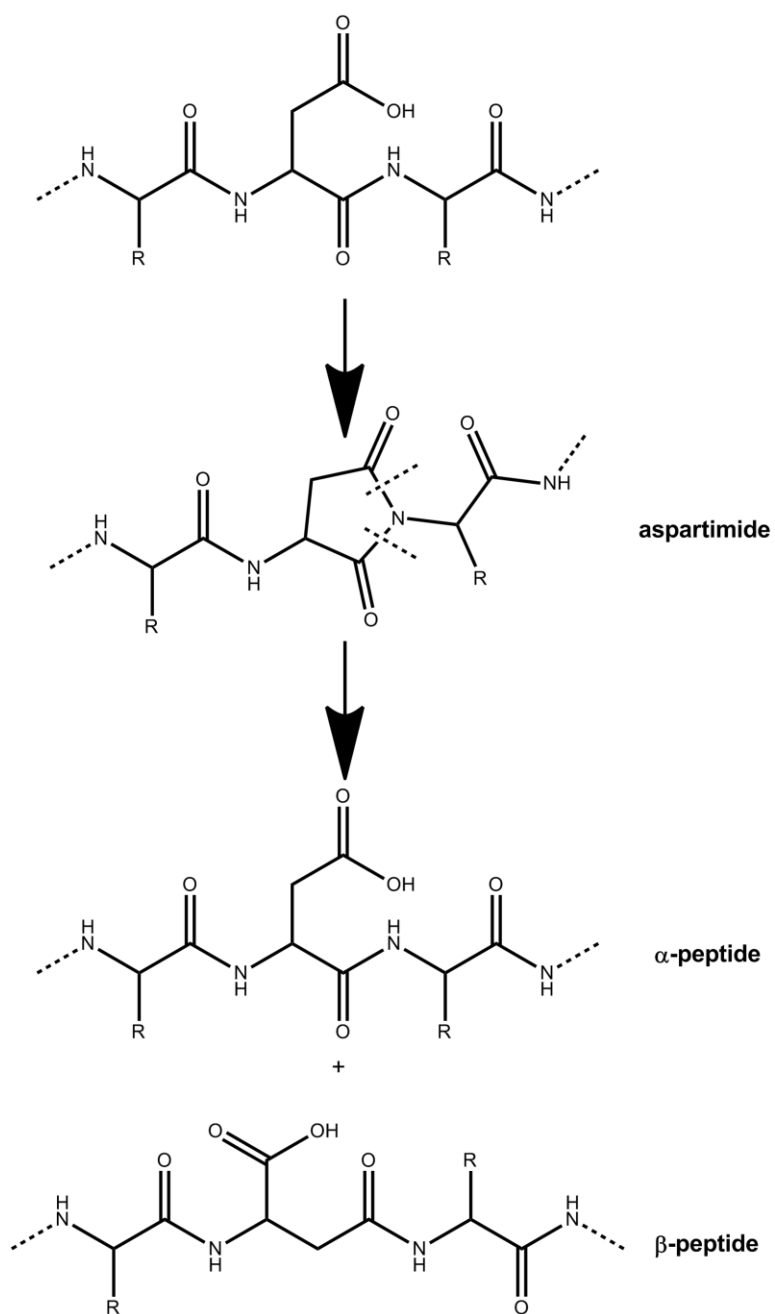




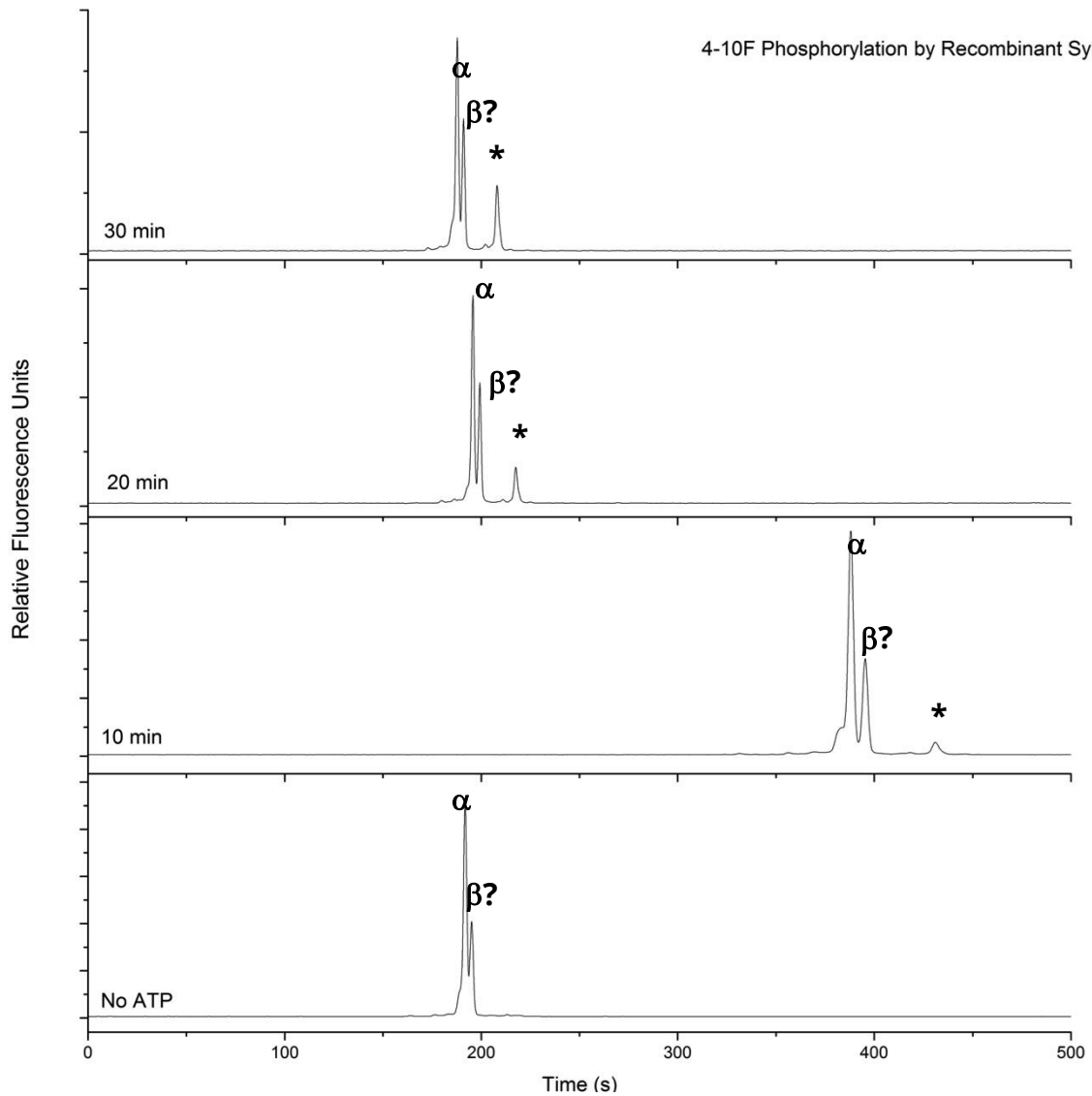
**Figure 27. Library compound screening by MALDI-TOF MS.** Top graph: The library peptides are in general well-phosphorylated by Syk, with several peptides phosphorylated better than the acetylated lysine control peptide. Bottom graph: Acylation at the library modification site is tolerated poorly by EGFR across all compounds. The primary hit for Syk (4-10F, structure of acylating reagent shown) is phosphorylated 4X better than the control and is not detectably phosphorylated by EGFR. N = 3 except where otherwise indicated.



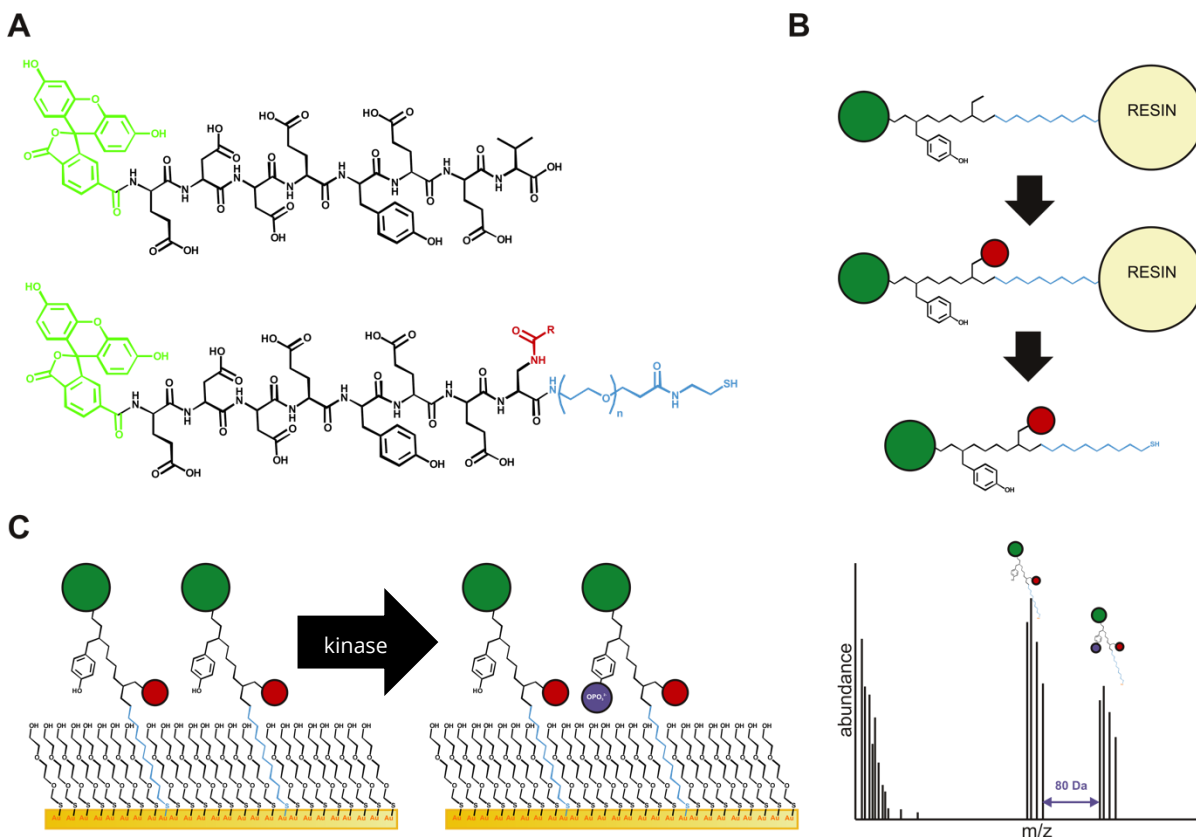
**Figure 28. Structure of the Syk screen hit 4-10F.**



**Figure 29. Mechanism of aspartimide formation and resultant racemization.**



**Figure 30. Phosphorylation of 4-10F by recombinant Syk.** Stacked traces show formation of a new peak (\*) that is time- and ATP-dependent.



**Figure 31. Proposed workflow for modified library screen with SAMDI. A)** Structures of 6FAM-TyrSub-9 (top) and the library scaffold with a C-terminal PEG spacer between the peptide and the resin linkage. **B)** Simplified view of peptide library synthesis using PEG. **C)**

## APPENDIX A. PRACTICAL SOLID PHASE PEPTIDE SYNTHESIS

Solid phase synthesis of short peptides (2-15 residues) using Fmoc chemistry is a mature, extensively studied technique that nevertheless requires a considerable level of experience to practice successfully and efficiently. The aim of this section is to provide a series of both general and specialized synthesis protocols that, borne out of trial and error, consultation with collaborators and the literature, and experience, have served the author well over the course of this work. It is the author's hope that these protocols will offer practical support to any researcher embarking on the journey of solid phase peptide synthesis.

### *Overview of Fmoc Solid Phase Peptide Synthesis*

Peptide synthesis at its most basic involves sequential formation of peptide bonds (Figure 32) between consecutive amino acids in the peptide sequence. In solution phase synthesis, selective bond formation between amino acids is non-trivial. For many years, synthetic challenges limited the capabilities and popularity of peptide synthesis. Merrifield's 1963 solid phase synthesis of a tetrapeptide introduced a novel and, as history has shown, revolutionary concept to the field of peptide synthesis.<sup>1</sup> Solution phase synthesis of even short amino acid polymers is a laborious, lengthy, costly, and inefficient task due to the necessity of isolating and characterizing the many peptide intermediates. Merrifield's introduction of a

heterogeneous synthetic medium reduced the problem of isolating intermediates to a simple wash step. Later, the introduction of the fluorenylmethyl carbonyl (Fmoc) amino protecting group expanded the utility and accessibility of solid phase peptide synthesis (SPPS) through milder, safer reaction conditions and online spectroscopic reaction monitoring.<sup>2</sup> Fmoc SPPS is now the standard peptide synthesis approach in industry and academic labs around the world.

Fmoc SPPS builds on Merrifield's original approach of tethering nascent peptide chains to porous solid supports by their carboxy termini, constructing the peptide from C- to N- from the surface of the support. The carboxy-support linker, usually acid labile, performs the dual functions of protecting the peptide C-terminus and of conferring the desired C-terminal functionality to the peptide upon cleavage. The fundamental advantage of Fmoc chemistry is the orthogonal protection of the amino groups with the base labile Fmoc group, allowing the full peptide chain to be constructed under basic conditions that have no effect on the peptide-support bond.

### *Protocol 1. Manual Synthesis of a General Tyrosine Kinase Substrate*

The following protocol will guide the reader through a solid phase synthesis of the peptide Ac-DNEYFYV-NH<sub>2</sub>,<sup>3</sup> a general tyrosine kinase substrate derived from the protein FLT3. Unless otherwise noted, all reagents were purchased from Fisher Scientific (Waltham, MA) or Sigma Aldrich (St. Louis, MO), with the exception of Fmoc-amino acids, which were purchased from EMD Millipore (Billerica, MA).

#### Supplies

NovaSyn® TGR Amide resin (NovaBiochem, EMD Millipore, Billerica, MA) or other suitable amide resin with moderate loading capacity

Fmoc-Asp(OtBu)-OH

Fmoc-Asn(Trt)-OH

Fmoc-Glu(OtBu)-OH

Fmoc-Phe-OH

Fmoc-Val-OH

Fmoc-Tyr(OtBu)-OH

HCTU

0.4 M 4-Methylmorpholine in DMF (**ACT**)

20 (v/v)% Piperidine in DMF + 0.1 M HOBt (**DEP**)

Acetic anhydride

Trifluoroacetic acid



Triisopropyl silane

Solvents: DMF, DCM, MeOH, IPA

Ninhydrin (Kaiser) test kit (AnaSpec, Fremont, CA)

### Protocol 1.1 Peptide Synthesis on the Solid Support

1. Allow resin and all reagents to reach room temperature before use to minimize contamination from moisture.
2. Weigh desired quantity of resin and transfer to fritted glass synthesis vessel or fritted solvent-resistant polymer synthesis syringe. Rinse resin with each of the following: MeOH (3X), IPA (3X), DCM (3X), DMF (6X). Allow resin to swell in several mL of DMF with gentle agitation for 20-30 min to ensure maximum reagent access to resin active sites.
3. Drain the resin and, if necessary, treat 2X10 min with several mL of DEP solution to deblock Fmoc-protected resin active sites. Following DEP treatment, resin should be rinsed a minimum of 6X with fresh DMF to ensure complete removal of piperidine before proceeding to the next step.
4. Prepare resin loading solution: 5 eq Fmoc-Val-OH + 4.5 eq HCTU in the minimum volume of ACT needed to dissolve the reagents and fully suspend resin in a slurry. Whenever HCTU is used to activate Fmoc-amino acids for coupling to resin, the Fmoc-amino acid should always be maintained in slight excess to avoid side product formation.

5. Add coupling solution to resin and allow to react with gentle agitation for 1 h at room temperature.
6. Rinse resin well with DMF and check for complete coupling using the Ninhydrin test (see Protocol 1.2). If the Ninhydrin test is positive, repeat steps 4 and 5 until a negative Ninhydrin test is achieved.
7. Treat rinsed resin 2X10 min with DEP under gentle agitation, then rinse thoroughly to fully remove piperidine.
8. Repeat steps 4-7 for each subsequent residue, building the peptide from C- to N-terminus.
9. After coupling of the N-terminal aspartic acid and Fmoc-removal at that site, rinse the resin thoroughly with DMF. Prepare the acetylation solution by mixing ~1 mL acetic anhydride with ~1 mL ACT and add immediately to the resin. React 1 h with gentle agitation, then rinse extensively with DMF to remove all trace acetic anhydride. Test for complete acetylation using the Ninhydrin test, and repeat as necessary.

#### Protocol 1.2 Monitoring Coupling Completion with the Ninhydrin Test

1. In a small test tube, combine ~20  $\mu$ L of each of three Ninhydrin test reagents. The solution should be a very pale yellow.
2. Using a Pasteur pipette or wide-bore glass capillary, transfer approximately 0.5 mg of peptide-bound resin beads to the tube.

3. Incubate the test tube in a sand bath between 90-110 °C for 5 min. If no free primary amines are present, the solution will remain pale yellow to brownish. If free primary amines are present, indicating incomplete reaction, the beads will turn dark blue to purple. Depending upon the amount of resin present, the solution may also turn dark purple-blue.

*Note: The Ninhydrin (Kaiser) test will only detect primary amines. Other tests are available to check reaction completion at secondary or tertiary amines.*

### Protocol 1.3 Peptide Cleavage from the Resin and Isolation

1. Rinse peptide-bound resin 6X in each of the following: DMF, IPA (or MeOH), DCM. Dry the resin briefly under vacuum or with a gentle jet of N<sub>2</sub>(g) to remove residual solvent.
2. Prepare the cleavage cocktail: 95% trifluoroacetic acid with 2.5% triisopropyl silane and 2.5% water added to scavenge cleaved side chain protecting groups. Treat resin 2.5-3 h in the cleavage cocktail with gentle agitation to cleave peptide from the resin and to remove OtBu and Trt peptide side-chain protecting groups.
3. Transfer peptide-TFA solution to a solvent-resistant centrifuge tube. Apply a gentle stream of N<sub>2</sub>(g) to agitate the surface of the solution and facilitate TFA evaporation. Continue until the solution is blown down to less than 1 mL.

4. Add 10-15 mL ice cold diethyl ether to the peptide solution. A precipitate should immediately form, indicating that the peptide has been successfully isolated from the TFA. Collect the peptide precipitate via centrifugation at 4 °C and decant ether. It is good practice to reserve the supernatant, removing the ether via evaporation, until the precipitate is successfully characterized as the desired product. Allow the peptide precipitate to dry in the hood overnight, covered with a Kim wipe to prevent contamination by dust.
5. Reconstitute the peptide for characterization and purification in the appropriate solvent. DMSO/water, 50 mM Tris-HCl (pH 7.5), and 50/50 acetonitrile/water are common choices, depending upon the subsequent characterization step.

## *Protocol 2. Labeling Primary Amines with Carboxyfluorescein*

### Reagents

5- or 6-Carboxyfluorescein (AnaSpec, Fremont, CA)

Diisopropylcarbodiimide (DIC)

HOBt or Oxyma

N-Methylpyrrolidone (NMP)

Solvents: DMF, MeOH, DCM

### Protocol

1. Deprotect the resin-bound primary amine to be labeled using the appropriate method. Rinse the resin 6X with DMF.
2. Prepare a solution of 3 eq carboxyfluorescein, 3.4 eq DIC, and 3.4 eq Oxyma in NMP. Allow to preactivate 5-10 min.
3. Add labeling solution to the resin-bound amine and allow to react 12-18 hours with gentle agitation.
4. Rinse the resin well with DMF and check for complete coupling using the Ninhydrin test. Repeat coupling as necessary.
5. Before cleavage, rinse thoroughly with DMF and treat resin 1X 5-10 min with 20% (v/v) piperidine in DMF with 0.1 M HOBt to remove any fluorescein esters.

6. Rinse the completed product extensively with DMF, MeOH, and DCM to remove residual fluorescein and proceed to cleavage step.

*Protocol 3. Microwave Synthesis of a Difficult Aspartic Acid-Rich Tyrosine Kinase Activity Reporter*

This protocol will guide the reader through the difficult, side reaction-prone synthesis of the tyrosine kinase activity reporter 6FAM-LEDDYEDDNle-NH<sub>2</sub> using an automated microwave synthesizer. This aspartic acid-rich sequence is prone to racemization through aspartimide formation,<sup>4,9</sup> resulting in the formation of undesirable beta-peptide side products that are extremely difficult to separate from the desired product by HPLC. Aspartimide formation is promoted by the presence of excess moisture and by prolonged exposure to reactive bases or acids, particularly piperidine. Microwave synthesis significantly decreases the reaction times necessary for completion of each synthetic step, thus minimizing exposure times to piperidine while allowing all reactions to reach completion. Addition of 0.1 M HOBt to the DEP solution has been shown to reduce piperidine-mediated aspartimide formation.<sup>10,11</sup>

Reagents

CLEAR amide resin (Peptides International, Louisville, KY)

250 mM HBTU in DMF

1 M Diisopropylethylamine (DIEA) in DMF (ACT)

20 (v/v)% Piperidine + 0.1 M HOBt in DMF (DEP)

0.1 M solutions of Fmoc-amino acids with appropriate side chain protection, in DMF

- E, Y: OtBu
- D: OMpe

Solvents: DMF, MeOH, IPA, DCM

### Protocol

1. Allow resin and all reagents to reach room temperature before use. This is to minimize unnecessary exposure to excess moisture, which promotes aspartimide formation.
2. Rinse resin to remove residual water: 3X MeOH, 3X IPA, 3X DCM, 6X DMF.
3. Prepare reagents as above and set up the synthesizer as per the manufacturer's instructions. The use of Fmoc-Asp(OMpe)-OH, rather than the more common Fmoc-Asp(OtBu)-OH, greatly minimizes aspartimide formation due to the increased bulkiness of the OMpe protecting group. Because this reagent can be cost-prohibitive, it should only be used in very aspartimide-prone sequences and can be coupled manually to reduce reagent consumption.
4. Run the synthesis with the following reaction times:
5. Label, cleave, isolate, and purify peptide as indicated above in protocols 1.3 and 1.4.

*Protocol 4. Synthesis of a Sterically Constrained Tyrosine Kinase Activity Reporter*

*Containing L-Htc (7-(S)-hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid)*

This protocol will guide the reader through the synthesis of the sterically constrained tyrosine kinase activity reporter 6FAM-EDDE-Htc-EEV-NH<sub>2</sub>, a phosphatase-resistant EGFR Kinase substrate.

Reagents

CLEAR amide (Peptides International, Louisville, KY) or NovaPEG amide resin (NovaBiochem, EMD Millipore, Billerica, MA)

Fmoc-amino acids

HCTU

0.4 M NMM in DMF

20% (v/v) Piperidine in DMF + 0.1 M HOBt

DIC (AnaSpec, Fremont, CA)

Oxyma (NovaBiochem, EMD Millipore, Billerica, MA)

DIEA

2-Chlorotriyl chloride (NovaBiochem, EMD Millipore, Billerica, MA)

Fmoc-7-S-hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid

6-FAM (AnaSpec, Fremont, CA)

TFA

TIS



#### Protocol 4.1 Synthesis of Htc-Containing Peptide

1. Allow resin to reach room temperature, rinse 3X MeOH, 3X IPA, 3X DCM, 6X DMF, and swell 20-30 min DMF.
2. Manually load resin with Fmoc-Val-OH, and couple the subsequent 2 glutamic acids using HCTU and NMM in DMF to catalyze coupling.  
  
Deprotect the peptide resin with 20 (v/v)% piperidine in DMF with 0.1 M HOBt, and rinse thoroughly.
3. Couple 3 eq Fmoc-Htc(OH)-OH overnight to the deprotected peptide resin using 3.4 eq DIC and 3.4 eq Oxyma in DMF, following a 10 min preactivation. Rinse thoroughly and check for complete coupling using the Ninhydrin test. Repeat coupling as necessary.
4. This step should be performed in a glass vessel. Rinse resin 6-10X with DCM and drain. Prepare the tritylation solution: 2-chlorotrityl chloride (50 eq or 100 mg, whichever is larger) dissolved in 1-5 mL DCM, depending on resin mass, with 1000 eq DIEA. Add to peptide resin and react 2 hours. The resin should change color from pale yellow-tan to darker yellow-brown. The 2-chlorotrityl group will be cleaved with the other side chain protecting groups in 95% TFA, but can also be removed earlier in 1-5% TFA in DCM if desired. To test for successful tritylation, rinse the resin thoroughly in DCM, then remove a few

- beads of resin and treat with TFA. If the trityl group was successfully added to the peptide-resin, the TFA will turn yellow and a gas will form.
5. Deprotect the resin using a 2X20 min treatment with 20 (v/v)% piperidine in DMF with 0.1 M HOBt. If desired, perform a secondary amine test (protocol 4.2) to establish a baseline for positive response.
  6. Couple Fmoc-Glu(OtBu)-OH to the peptide resin overnight using the DIC/Oxyma/DMF system. Check for complete coupling using the secondary amine test. Repeat coupling as necessary.
  7. Synthesize the rest of the peptide using standard protocols (HCTU/NMM/DMF), label and cleave as outlined above.

#### Protocol 4.2 Secondary Amine Test

##### Reagents

2% acetaldehyde in DMF

2% chloranil in DMF

##### Protocol

1. Isolate a few beads of resin using the open end of a Mel-Temp capillary.
2. Transfer beads to a small test tube containing a 1:1 mixture of the chloranil and acetaldehyde solutions.

3. Allow to incubate 5 min at room temperature, then inspect the color of the beads. Dark blue or green beads indicate the presence of secondary amine. Colorless or yellow beads indicate no detected secondary amine.

## REFERENCES

- (1) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- (2) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis a Practical Approach* 2000.
- (3) Kerman, K.; Song, H.; Duncan, J. S.; Litchfield, D. W.; Kraatz, H.-B. *Anal. Chem.* **2008**, *80*, 9395–9401.
- (4) Mergler, M.; Dick, F.; Sax, B.; Weiler, P.; Vorherr, T. *J. Peptide Sci.* **2003**, *9*, 36–46.
- (5) Mergler, M.; Dick, F.; Sax, B.; St helin, C.; Vorherr, T. *J. Peptide Sci.* **2003**, *9*, 518–526.
- (6) Mergler, M.; Dick, F. *J. Peptide Sci.* **2005**, *11*, 650–657.
- (7) Lauer, J. L.; Fields, C. G.; Fields, G. B. *Lett Pept Sci* **1995**, *1*, 197–205.
- (8) Ruczyński, J.; Lewandowska, B.; Mucha, P.; Rekowski, P. *J. Peptide Sci.* **2008**, *14*, 335–341.
- (9) Karlström, A.; Undén, A. *Tetrahedron Letters* **1996**.
- (10) Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W. *Org. Lett.* **2012**, *14*, 5218–5221.
- (11) Wade, J. D.; Mathieu, M. N.; Macris, M.; Tregear, G. W. *Int J Pept Res Ther* **2000**, *7*, 107–112.

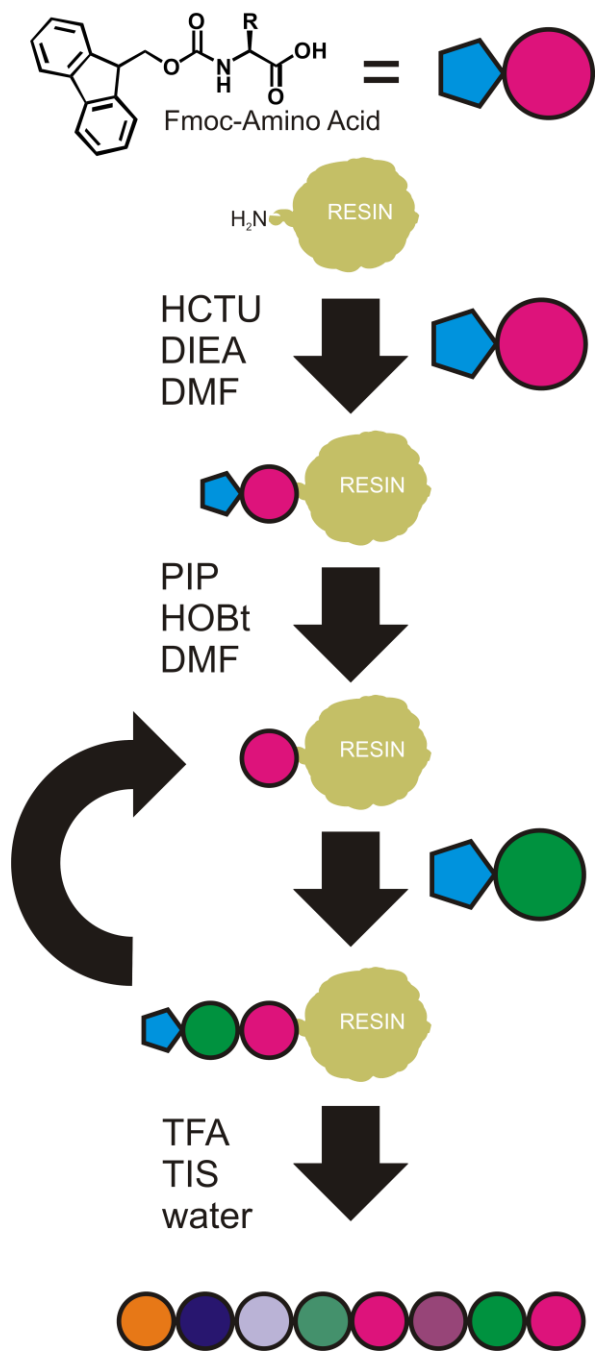


Figure 32. Solid Phase Peptide Synthesis (SPPS) workflow

## APPENDIX B. MATLAB CODE

### Collaboration with Dr. Ryan Phillips

```
function[output] = msbatchp(mass_tol)

% mode = input('MS mode positive or negative?');
if exist('mode')~= 1
    mode = menu('Select MS Mode','Negative','Positive');
end

% scaffoldin = input('What is the scaffold molecular weight in Da?');

%If no mass tolerance is provided, ask for it.
if nargin() == 0
    mass_tol = input('Choose a mass tolerance in Da: ');
    thresh_in = input('Choose a thresholding value (fraction max value):');
end

%Select guide file and data directory
[f,p] = uigetfile('.txt','Select guide file');
filename_guide = strcat(p,f);
dir_data = uigetdir('c:\Documents and Settings\Abby Turner\My
Documents\MATLAB\Data','Select data directory');
cd(dir_data);

%Import guide spot names and MW info
guide_raw = importdata(filename_guide);
spotnames = guide_raw.textdata;
guideMW = guide_raw.data;
iscorrect = zeros(1,length(guideMW));
pct_yield = zeros(1,length(guideMW));
pct_phospho = zeros(1,length(guideMW));

for i = 1:length(guideMW) %opens a loop for opening and analyzing MASCOT
files
    searchterm = strcat('*_',spotnames{i},'_*'); %spot names corresponding to
samples analyzed (MALDI plate location)
    [n,~] = size(dir(searchterm)); %opens all files within designated
directory that contain spot names (MASCOT files)
    if n~=0 %creates file names and opens all files with appropriate format
filename
        data_filename_raw = dir(searchterm);
        data_filename = data_filename_raw.name;
        data_raw = importdata(data_filename,'\t',1);

%extracts columns of data for analysis
MW = data_raw.data(:,1); %MW column from MASCOT files
Intensity = data_raw.data(:,2); %Corresponding intensities for each MW
guideMW(i); %Pulls in guide MW (list of correct target MW per spot)
```

```

threshold = thresh_in*(max(Intensity));

if mode == 1 %Based on mode, chooses which will be prominent molecular
ion.
    MW_molecion(i) = guideMW(i)-1;
    MW_adduct(i) = guideMW(i)-1;
elseif mode == 2
    MW_molecion(i) = guideMW(i)+1;
    MW_adduct(i) = guideMW(i);
end

Intensity = Intensity.*(Intensity>threshold);

%find any traces of correct target in sample
MWmask_MW = ((abs((MW)-MW_molecion(i)))<mass_tol);
MWmask_waterloss = (((abs((MW)-MW_molecion(i)-18)))<mass_tol);
MWmask_sodium = ((abs((MW)-(MW_adduct(i)+22.99)))<mass_tol);
MWmask_potassium = ((abs((MW)-(MW_adduct(i)+39.1)))<mass_tol);
MWmask = MWmask_MW|MWmask_waterloss|MWmask_sodium|MWmask_potassium;

%finds any traces of phospho target in sample
phosphomask_MW = ((abs((MW)-(MW_molecion(i)+80)))<mass_tol);
phosphomask_waterloss = ((abs((MW)-(MW_molecion(i)+80-18)))<mass_tol);
phosphomask_sodium = ((abs((MW)-(MW_adduct(i)+80+22.99)))<mass_tol);
phosphomask_potassium = ((abs((MW)-(MW_adduct(i)+80+39.1)))<mass_tol);
phosphomask =
phosphomask_MW|phosphomask_waterloss|phosphomask_sodium|phosphomask_potass
ium;

%applies masks to data and sums results
maskedMW = MWmask.*MW;
maskedInt = MWmask.*Intensity;
phosphoInt = phosphomask.*Intensity;

if sum(maskedMW)>=1;
    pct_yield(i) = sum(maskedInt)/sum(Intensity)*100; %returns percent yield
    pct_phospho(i) =
(sum(phosphoInt) / (sum(phosphoInt)+sum(maskedInt))) *100; %returns percent
phosphorylation

else
    pct_yield(i) = 0;

end

end

end
end

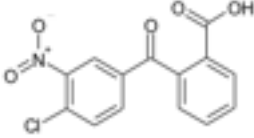
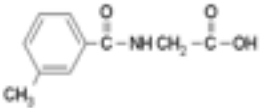
filename_export = strcat(input('Designate a filename for export:
','s'),'.txt');
fid = fopen(filename_export,'w');
header = {'Spot' '%Yield' '%Phospho'};

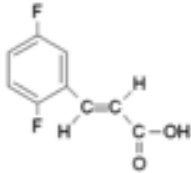
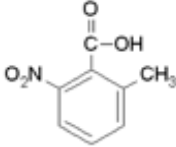
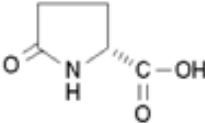
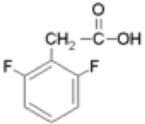
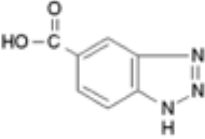
```

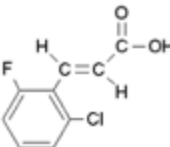
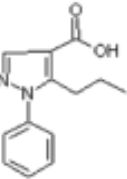
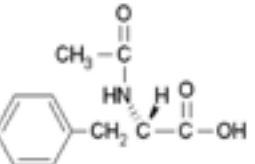
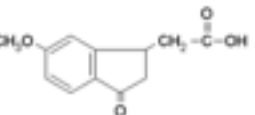
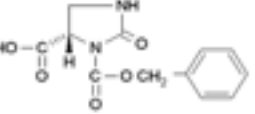
```
fprintf(fid, '%s\t', header{1,1:end-1});  
fprintf(fid, '%s\r\n', header{1,end});  
  
for i = 1:length(spotnames)  
    fprintf(fid, '%s\t', spotnames{i});  
    % fprintf(fid, '%i\t', iscorrect(i));  
    % fprintf(fid, '%i\t', pct_intensity(i));  
    fprintf(fid, '%i\t', pct_yield(i));  
    % fprintf(fid, '%i\r\n', pct_scaffold(i));  
    fprintf(fid, '%i\r\n', pct_phospho(i));  
end  
  
fclose(fid);  
  
end
```

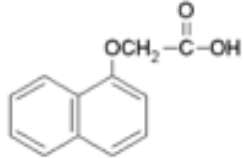
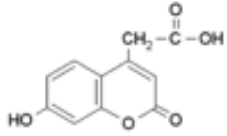
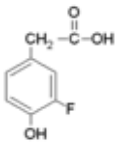
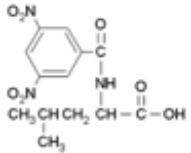
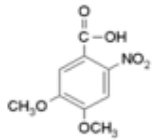


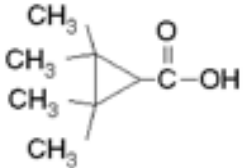
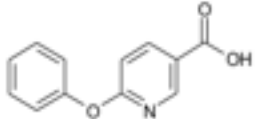
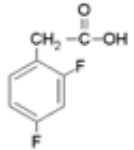
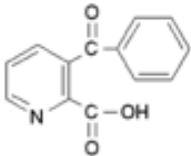
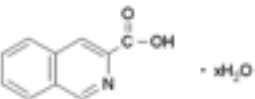
APPENDIX C. CARBOXYLIC ACID LIBRARY

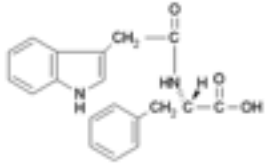
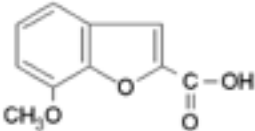
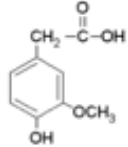
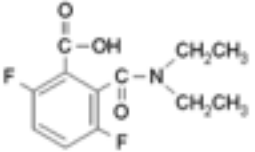
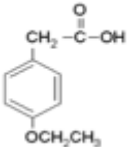
Well #	Cmpd #	Chemical Name	Chemical Structure	CAS No.	CAT No.	FW	Solubility (DMF)
1	1-10H	2-(4-chloro-3-nitrobenzoyl)benzoic acid		85-54-1		305.67	good
2	1-11C	3-Methylhippuric acid		32,801-4		193.2	excellent

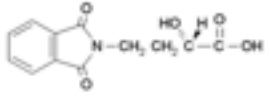
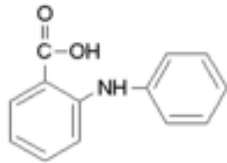
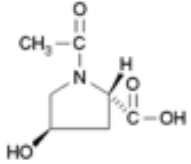
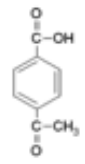
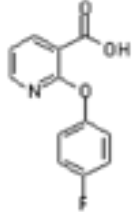
3	<b>1-1H</b>	Trans-2,5-Difluorocinnamic acid		29,034-3		184.14	excellent
4	<b>5-8E</b>	2-Methyl-6-nitrobenzoic acid		15,139-4		181.15	good
5	<b>1-4E</b>	(R)-(+)-2-Pyrrolidone-5-carboxylic acid			42,261-4	129.12	good
6	<b>1-4F</b>	2,6-Difluorophenylacetic acid			26,448-2	172.13	good
7	<b>5-9A</b>	Benzotriazole-5-carboxylic acid		30,423-9		163.14	poor (but went in)

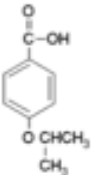
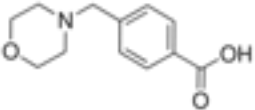
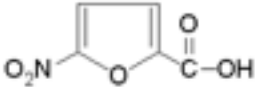
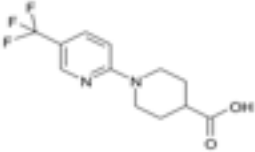
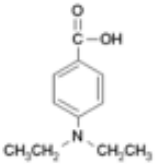
8	1-8C	Trans-2-Chloro-6-fluorocinnamic acid		41,945-1		200.6	excellent
9	1-8D	1-phenyl-5-propyl-1H-pyrazole-4-carboxylic acid (Maybridge)		116344-17-3	230	230.26636	good
10	2-11D	N-Acetyl-L-phenylalanine		85,745-9		207.23	good
11	2-1G	5-Methoxy-1-indanone-3-acetic acid		24467-92-3 #255282-1G		220.23	good
12	2-3E	(S)-(-)-2-Oxo-1,5-imidazolidinedicarboxylic acid 1-benzyl ester		59760-01-9	#392308-1G	264.24	good

13	<b>2-3F</b>	(1-Naphthoxy)acetic acid		2976-75-2 #255416-1G		202.21	good
14	<b>2-4B</b>	7-Hydroxycoumarin-4-acetic acid		6950-82-9 #55157-500MG		220.18	good
15	<b>2-5H</b>	3-Fluoro-4-hydroxyphenylacetic acid		22,451-0		170.14	good
16	<b>2-7B</b>	N-(3,5)-Dinitrobenzoyl)-DL-leucine		74928-54-4	#251526-1G	325.28	good
17	<b>2-9D</b>	4,5-Dimethoxy-2-nitrobenzoic acid		4998-07-6 #329231-1G		227.18	good

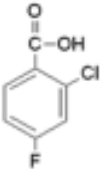
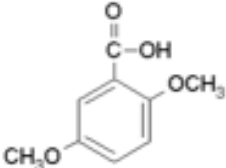
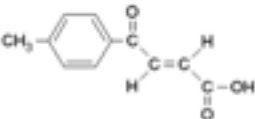
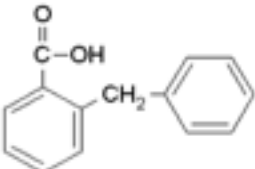
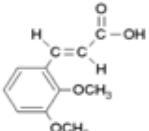
18	<b>2-9G</b>	2,2,3,3-tetramethyl-cyclopropanecarboxylic acid		30,156-6		142.2	excellent
19	<b>6-1D</b>	6-phenoxy nicotinic acid		51362-38-0	CC19601-DA Fisher	215.21	good
20	<b>3-8B</b>	2,4-Difluorophenylacetic acid		26,447-4		172.13	excellent
21	<b>4-10C</b>	3-Benzoyl-2-pyridine-carboxylic acid		64362-32-9 #422576		227.22	good
22	<b>4-10E</b>	3-Isoquinolinecarboxylic acid hydrate		33,854-0		173.17	good

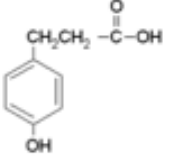
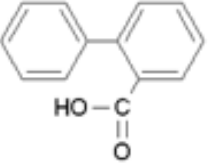
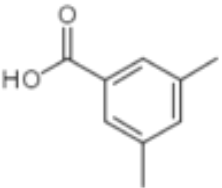
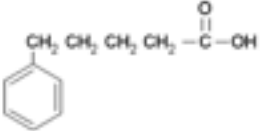
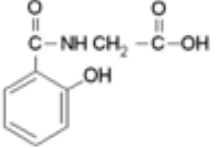
23	<b>4-10F</b>	N-(3-Indolylacetyl)-L-phenylalanine		57105-50-7	#345954-1G	322.37	okay
24	<b>4-11B</b>	7-Methoxy-2-benzofuran-carboxylic acid		35,772-3		192.17	good
25	<b>4-11G</b>	Homovanillic acid		14,364-2		182.18	good
26	<b>4-1H</b>	N,N-Diethyl-3,6-difluorophthalamic acid		131401-56-4	#378291-1G	257.24	okay
27	<b>6-4G</b>	4-Ethoxyphenylacetic acid		4919-33-9	A12811-2	180.2	excellent

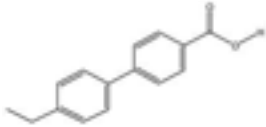
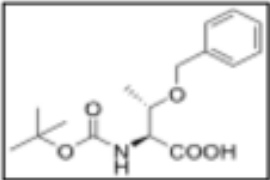
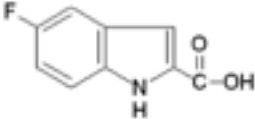
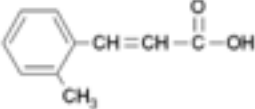
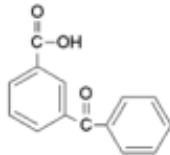
28	4-3D	(S)-(+)- $\alpha$ -Hydroxy-1,3-dioxo-2-isoindolinebutyric acid		40,965-0		249.23	good
29	4-4E	N-Phenylanthranilic acid		14,450-9		213.24	low
30	5-11A	Trans-1-Acetyl-4-hydroxy-L-proline		44,156-2		173.17	good
31	5-11H	4-Acetylbenzoic acid		17,745-8		164.16	okay
32	5-1A	2-(4-Fluorophenoxy)nicotinic acid		54629-13-9		233.198823	good

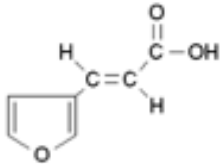
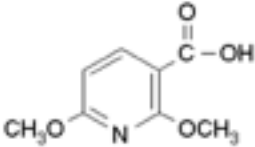
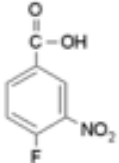
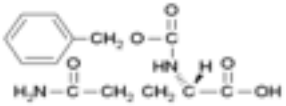
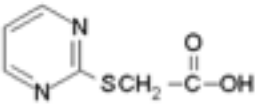
33	5-2A	4-Isopropoxybenzoic acid		36,638-2		180.2	good
34	6-5G	4-Morpholinomethyl benzoic acid		89-86-1	AD10940-1	221.26	good
35	5-4A	5-Nitro-2-furoic acid		15,571-3		157.08	good
36	5-4B	1-[5-(trifluoromethyl)-2-pyridyl]piperidine-4-carboxylic acid		406476-31-1		274.24	excellent
37	5-5A	4-(Diethylamino)benzoic acid		23,735-3		193.25	excellent

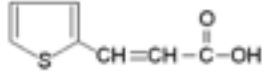
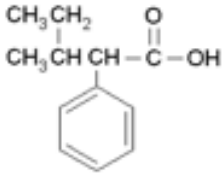
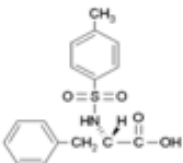
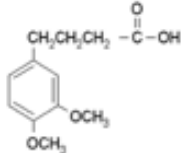
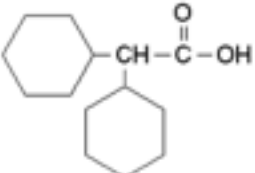


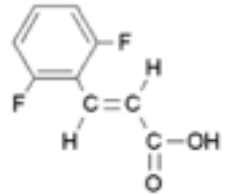
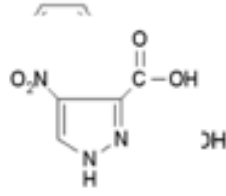
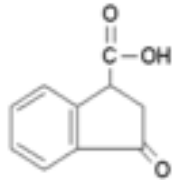
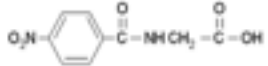
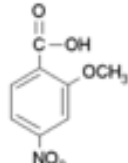
38	5-5C	2-Chloro-4-fluorobenzoic acid		29,783-6		174.56	good
39	5-6E	2,5-Dimethoxybenzoic acid		25,796-6		182.18	good
40	5-8G	Trans-3-(4-Methylbenzoyl)-acrylic acid		34,601-2		190.2	good
41	6-11A	o-Phenyl-o-toluic acid		612-35-1	AP3665-7	212.25	excellent
42	6-11H	Trans-2,3-Dimethoxycinnamic acid		7345-82-6	A27548-4	208.21	good

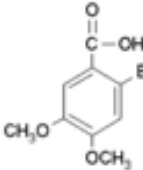
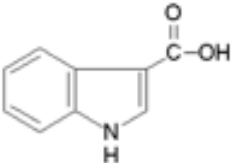
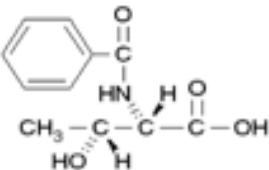
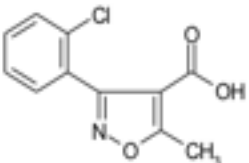
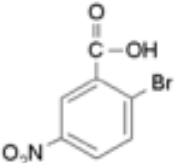
43	<b>6-1C</b>	3-(4-Hydroxyphenyl)propionic acid		501-97-3	AH5240-6	166.18	okay
44	<b>6-4E</b>	2-Biphenylcarboxylic acid		947-84-2	AB3470-2	198.22	good
45	<b>6-4F</b>	3,5-Dimethylbenzoic acid		499-06-9	D149608	150.18	good
46	<b>6-5F</b>	5-Phenylvaleric acid		2270-20-4	AP3760-2	178.23	good
47	<b>7-2B</b>	2-hydroxyhippuric acid		487-54-7	A13406-6	195.17	excellent

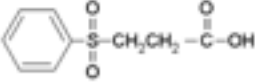
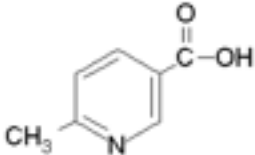
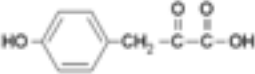
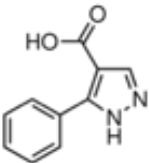
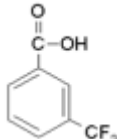
48	6-7F	4-Ethylbiphenyl-4'-carboxylic acid (4(4-Ethylphenyl)benzoic acid)		5731-13-5	AC24048-0050 Acros	226.48	good
49	6-8F	4-Propoxybenzoic acid		5438-19-7	A36639-0	180.2	excellent
50	7-4F	5-fluoroindole-2-carboxylic acid		399-76-8	A26512-8	179.45	good
51	7-1E	2-Methylcinnamic acid		2373-76-4	A43310-1	162.19	good
52	7-2D	3-benzoylbenzoic acid		579-18-0	A26179-3	226.23	excellent

53	7-2E	Trans-3-Furanacrylic acid		81311-95-7	A33638-6	138.12	excellent
54	7-2G	2,6-dimethoxynicotinic acid		16727-43-8	A37271-4	183.16	good
55	7-4E	4-fluoro-3-nitrobenzoic acid		453-71-4	A32904-5	185.11	good
56	7-5G	Carbobenzyloxy-L-glutamine		2650-64-8	A16264-7	280.28	good
57	7-7G	(2-pyrimidylthio)acetic acid		88768-45-0	A27553-0	170.19	excellent

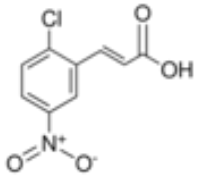
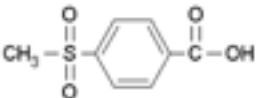
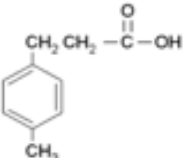
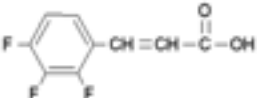
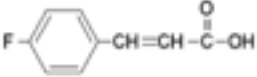
58	8-1C	3-(2-Thienyl)acrylic acid		15690-25-2	A13058-3	154.19	good
59	8-1F	3-Methyl-2-phenylvaleric acid		7782-37-8	A24790-1	192.26	good
60	8-4C	N-p-Tosyl-L-phenylalanine		13505-32-3	A39384-3	319.38	excellent
61	8-6C	4-(3,4-Dimethoxyphenyl)-butyric acid		13575-74-1	A27395-3	224.26	good
62	7-5A	Dicyclohexylacetic acid		52034-92-1	A33384-0	224.35	good

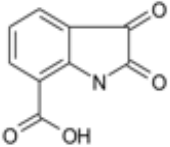
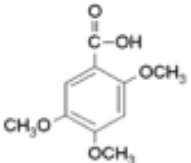
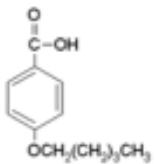
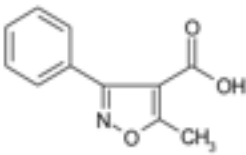
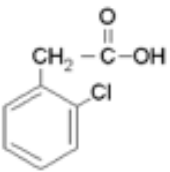
63	<b>1-1E</b>	Trans-2,6-Difluorocinnamic acid		29,035-1		184.14	excellent
64	<b>1-2B</b>	4-Nitro-3-pyrazolecarboxylic acid		41,484-0		157.09	okay
65	<b>1-3B</b>	3-Oxo-1-indancarboxylic acid		41,077-2		176.17	excellent
66	<b>1-5H</b>	4-Nitrohippuric acid		23,343-9		224.17	okay
67	<b>1-7B</b>	2-Methoxy-4-nitrobenzoic acid		42,291-6		197.15	good

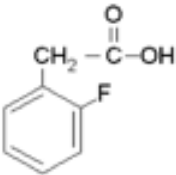
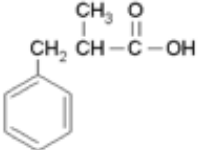
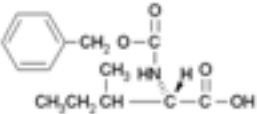
68	<b>1-7C</b>	2-Bromo-4,5-dimethoxybenzoic acid		44,107-4		261.08	excellent
69	<b>2-10G</b>	Indole-3-carboxylic acid		28,473-4		161.16	good
70	<b>2-11F</b>	N-Benzoyl-L-threonine		27696-01-1 #281093-1G		223.23	good
71	<b>2-2B</b>	3-(2-chlorophenyl)-5-methylisoxazole-4-carboxylic acid		23598-72-3		237.64	good
72	<b>2-3B</b>	2-Bromo-5-nitrobenzoic acid		943-14-6 #381845-1G		246.02	okay

73	2-5E	3-(Phenylsulfonyl)propionic acid		25,157-7		214.24	okay
74	2-9H	6-Methylnicotinic acid		#284750-1G		137.14	poor (but went in)
75	3-1C	4-Hydroxyphenylpyruvic acid		11,428-6		180.16	good
76	3-3C	5 phenyl-1-H-pyrazole-4-carboxylic acid		633747-1g		188.18	okay
77	7-5C	A,a,a-Trifluoro-m-toluic acid		454-92-2	A18834-4	190.12	good



78	7-6E	2-chloro-5-nitrocinnamic acid		36015-19-7	A30065-9	227.61	excellent
79	3-7E	4-(Methylsulfonyl)benzoic acid		13,641-7		200.21	okay
80	3-8C	3-(p-Tolyl)propionic acid		11,826-5		164.2	good
81	4-3B	2,3,4-Trifluorocinnamic acid		41,838-2		202.13	good
82	4-4H	4-Fluorocinnamic acid		459-32-5 #222720-1G		166.15	good

83	4-6C	2,3-dioxindoline-7-carboxylic acid		25128-35-2		191.14	poor (but went in)
84	8-1H	2,4,5-Trimethoxybenzoic acid		490-64-2	A13889-4	212.2	good
85	8-4G	4-Pentyloxybenzoic acid		15872-41-0	A20994-5	208.26	excellent
86	5-10C	5-methyl-3-phenylisoxazole-4-carboxylic acid		1136-45-4		203.19	excellent
87	5-7A	2-Chlorophenylacetic acid		19,063-2		170.6	good

88	5-7F	2-Fluorophenylacetic acid		20,894-9		154.14	excellent
89	5-8A	$\alpha$ -Methylhydrocinnamic acid		39,152-2		164.2	poor (but went in)
90	5-8C	N-Carbobenzyloxy-L-isoleucine		40,853-0		265.31	good