Development of Tyrosine Kinase Peptide Biosensors and Methods for Detection

Andrew Michael Lipchik
Purdue University

Follow this and additional works at: http://docs.lib.purdue.edu/open_access_dissertations
Part of the Biochemistry Commons, and the Bioinformatics Commons

Recommended Citation

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
This is to certify that the thesis/dissertation prepared

By Andrew Michael Lipchik

Entitled
Development of Tyrosine Kinase Peptide Biosensors and Methods of Detection

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Laurie L. Parker
Chair

Robert Geahlen

Joseph Irudayaraj

Tony Hazbun

To the best of my knowledge and as understood by the student in the Research Integrity and Copyright Disclaimer (Graduate School Form 20), this thesis/dissertation adheres to the provisions of Purdue University’s “Policy on Integrity in Research” and the use of copyrighted material.

Approved by Major Professor(s): Laurie Parker

Approved by: Richard Borch
Head of the Graduate Program 11/25/2013
DEVELOPMENT OF TYROSINE KINASE PEPTIDE BIOSENSORS AND METHODS FOR DETECTION

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Andrew Michael Lipchik

In Partial Fulfillment of the
Requirements for the Degree

of

Doctor of Philosophy

December 2013

Purdue University

West Lafayette, Indiana
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Laurie Parker, for her support over the last five years. Under her guidance, I have grown tremendously as a scientist. Laurie, you have set an excellent example of what it is to be a scientist and a mentor. I sincerely appreciate your trust in me to pursue the ideas presented here. I also would like to thank you for providing the opportunities and cultivating a learning environment where it’s fun and exciting to come to the lab each day.

I would like to thank the members of my thesis committee for their support during my time at Purdue: Dr. Robert Geahlen, Dr. Joseph Irudayaraj, Dr. Donald Bergstrom and Dr. Tony Hazbun. I would like to thank each of you for your time, feedback, and contributions to my work and growth as a scientist. Dr. Geahlen and Dr. Irudayaraj it has been a pleasure to have the opportunity to collaborate with both of you during my time at Purdue. Dr. Hazbun thank you so much for filling Dr. Bergstrom’s position and your helpful comments and feedback on my work.

My labmates in the Parker lab, past and present, thank you for developing a great work environment. You have made the lab a supportive, collaborative, friendly, and fun place to spend a significant portion of the last five years. I am grateful for your help and feedback over the years when reviewing manuscript drafts, fellowship applications and
feedback during group meetings. I wish the best of luck to all of you in all your future
endeavors.

Special thanks to Michael, Minervo, Amber and Vas; your work at the various
stages of this project has greatly contributed to the work presented here. Also, I hope that
you have learned as much from me as I have learned from you about being a mentor. The
opportunity to mentor each of you has shaped my aspirations as a scientist and the future
career I would like to pursue. I wish all of you the best of luck in your future scientific
careers.

I would like to thank my previous mentors and chemistry teachers who inspired
my interest in chemistry and to pursue a graduate education in chemistry: my high school
chemistry teacher, Mrs. Parow, my undergraduate research advisor at Xavier University,
Dr. Craig M. Davis and my advance organic professor Dr. Richard J. Mullins.

My family, to Mom, Dad, Adam and Grandma – I cannot begin to describe how
much your support, encouragement, and love have meant to me especially over the last
five years. Mom and Dad, you have always encouraged me to be creative, to pursue my
passions and hold nothing back. Adam, thank you for your support and light-heartness.
Grandma, thank you for your wisdom and your candor about life.

To Cindy – your unconditional love, support, and friendship means the world to
me. Thank you for spending those long nights in the lab with me while I worked. This
thesis is as much mine as it is yours. I could not have done this without out and I look
forward to seeing what else we accomplish together in the future. I love you.
# TABLE OF CONTENTS

| LIST OF TABLES | ix |
| LIST OF FIGURES | x |
| LIST OF ABBREVIATIONS | xiii |
| ABSTRACT | xiv |
| CHAPTER 1. INTRODUCTION | 1 |
  | 1.1 Protein Kinases | 1 |
  | 1.2 Kinase Substrate Recognition | 1 |
  | 1.3 Methods for Determining Kinase Substrate Sequence Specificity | 8 |
    | 1.3.1 Genetically Encoded Peptide Libraries | 9 |
      | 1.3.1.1 Phage Display Libraries | 10 |
      | 1.3.1.2 mRNA Fusion Peptide Libraries | 11 |
      | 1.3.1.3 cDNA Display Libraries | 12 |
      | 1.3.1.4 DNA/PNA Programmed Synthetic Libraries | 14 |
    | 1.3.2 Synthetic Peptide Libraries | 15 |
      | 1.3.2.1 Oriented Peptide Libraries | 16 |
      | 1.3.2.2 Positional Scanning Peptide Libraries | 17 |
      | 1.3.2.3 One-Bead-One-Peptide (OBOP) Libraries | 18 |
1.3.3  Computational Methods for the Prediction of Kinase Specificity .......................... 21
  1.3.3.1  Position-Specific Scoring Matrices (PSSM) ................................................... 22
  1.3.3.2  High Order Machine Learning Models .......................................................... 23

1.4  The Role of Protein Kinases in Disease ................................................................. 25

1.5  Peptide Biosensors for Monitoring Kinase Activity .................................................. 27
  1.5.1  Traditional Molecular Biology Techniques ......................................................... 27
  1.5.2  Mass Spectrometry Detection .............................................................................. 29
  1.5.3  Fluorescence Detection Techniques .................................................................... 31
    1.5.3.1  Genetically Encoded FRET-Based Biosensors ............................................... 32
    1.5.3.2  Nanoparticle-Conjugates ............................................................................... 33
    1.5.3.3  Environmentally-Sensitive Fluorescence ......................................................... 35
    1.5.3.4  Chelation-Enhanced Fluorescence .................................................................. 37
    1.5.3.5  Biologically Available Metal Chelation .......................................................... 37
    1.5.3.6  Calcium-Based Chelation-Enhanced Fluorescence Peptide Biosensors ........... 38
    1.5.3.7  Magnesium-Based Chelation Enhanced Fluorescence Peptide Biosensors .......... 38
    1.5.3.8  Lanathanide-Based Chelation-Enhanced Luminescence .................................. 40
    1.5.3.9  EF-Hand Based Sensors ............................................................................... 41
    1.5.3.10  Macrocyclic Terbium Chelate Sensors ......................................................... 43
    1.5.3.11  Chelator Based ............................................................................................. 43

1.6  Dissertation Objective ............................................................................................... 45

1.7  References .................................................................................................................. 48
CHAPTER 2. A PEPTIDE-BASED BIOSENSOR ASSAY TO DETECT INTRACELLULAR SYK KINASE ACTIVATION AND INHIBITION

2.1 Abstract

2.2 Introduction

2.3 Methods and Materials

2.4 Results

2.4.1 Development of a Syk Biosensor

2.4.2 Detection of Dose Dependent Activation and Inhibition of Syk in Intact Cells

2.4.3 Time-Dependence of Syk activity following activation

2.4.4 Determination of the SAStide Biosensor Specificity in Intact Cells

2.4.5 Detection of Syk Activity and Inhibition in Primary Mouse Splenic B-cells

2.5 Discussion

2.6 References

CHAPTER 3. TIME-RESOLVED LUMINESCENCE DETECTION OF SPLEEN TYROSINE KINASE ACTIVITY THROUGH TERBIUM SENSITIZATION

3.1 Abstract

3.2 Introduction

3.3 Materials and Methods

3.4 Results

3.4.1 Time-resolved Luminescence Measurements Increase Signal to Noise
3.4.2 Physical Characterization of SAStide-Lanthanide Binding and Luminescence ................................................................. 105

3.4.3 Quantitative Detection of Syk Kinase Activity Using Tb\(^{3+}\) Sensitization... 108

3.5 Discussion............................................................................................................................................................................. 113

3.6 References............................................................................................................................................................................... 115

CHAPTER 4. PIPELINE TO DEVELOP PHOSPHORYLATION-DEPENDENT TERBIUM SENSITIZING TYROSINE KINASE BIOSENSORS................................. 120

4.1 Abstract..................................................................................................................................................................................... 120

4.2 Introduction............................................................................................................................................................................. 121

4.3 Methods ............................................................................................................................................................................... 126

4.4 Results..................................................................................................................................................................................... 133

4.4.1 Design and Development of Kinase Specific Substrates For Terbium-Based Detection of Phosphorylation................................................................. 133

4.4.2 Identification and Validation of Kinase Specific Residues ......................... 138

4.4.3 Identification and Characterization of Abl, Jak2 and Src-family Kinase Biosensors ............................................................................................................. 146

4.4.4 Tb\(^{3+}\) Luminescence Characterization of KINATEST-ID Identified Biosensors ................................................................................................................................. 150

4.4.5 In vitro Detection of Nonreceptor Tyrosine Kinase Activity ..................... 152

4.4.6 Application of AbASTide for High Throughput Screening of Small Molecule Inhibitions for Chronic Myeloid Leukemia (CML) .................................................................................................................. 155

4.5 Discussion............................................................................................................................................................................... 158
| Page |
---|---|
| 4.6 References | 164 |

CHAPTER 5. SIMULTANEOUS MULTIPLEXED DETECTION OF LYN AND SYK TYROSINE KINASE ACTIVITIES USING TIME-RESOLVED LANTHANIDE-BASED RESONANCE ENERGY TRANSFER (TR-LRET) ...................... 171

5.1 Abstract .................................................................................................................. 171
5.2 Introduction ............................................................................................................. 171
5.3 Methods .................................................................................................................. 173
5.4 Results .................................................................................................................... 176
5.5 Discussion ............................................................................................................... 184
5.6 References ............................................................................................................. 185

VITA ............................................................................................................................. 189

PUBLICATIONS ........................................................................................................... 191
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3-1 Photophysical properties of the pSAStide-terbium and SASTide-terbium complexes.</td>
<td>108</td>
</tr>
<tr>
<td>Table 4-1 Kinetic constants for kinase specific peptide biosensors</td>
<td>150</td>
</tr>
<tr>
<td>Table 4-2 Biosensor characterization of physical and photophysical properties</td>
<td>152</td>
</tr>
<tr>
<td>Table 4-3 Characterization of biosensor high-throughput screening parameters</td>
<td>153</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1-1 Representative schematic of MAPK substrate recognition through distal catalytic domain interactions</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1-2 Protein-protein interactions determine yeast MAPKK function</td>
<td>7</td>
</tr>
<tr>
<td>Figure 1-3 <em>In vitro</em> screening and selection of genetically encoded peptide libraries</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1-4 Workflow of PNA-encoded peptide library for the determination of Abl kinase substrate specificity</td>
<td>15</td>
</tr>
<tr>
<td>Figure 1-5 Synthesis of one-bead-one peptide ladder library</td>
<td>19</td>
</tr>
<tr>
<td>Figure 1-6 Representations of higher order machine-learning models</td>
<td>23</td>
</tr>
<tr>
<td>Figure 1-7 Number of identified phosphorylation sites for each kinase annotated in the Phosphosite Plus data set</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1-8 ELISA-based detection of peptide biosensor phosphorylation</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1-9 Genetically encoded FRET-based biosensor</td>
<td>33</td>
</tr>
<tr>
<td>Figure 1-10 Quantum dot detection of kinase activity</td>
<td>35</td>
</tr>
<tr>
<td>Figure 1-11 Environmentally-sensitive fluorescent peptide biosensors</td>
<td>37</td>
</tr>
<tr>
<td>Figure 1-12 Biologically available metal based CHEF biosensors</td>
<td>40</td>
</tr>
<tr>
<td>Figure 1-13 Lanthanide-based CHEF biosensors</td>
<td>44</td>
</tr>
<tr>
<td>Figure 2-1 Syk peptide biosensor</td>
<td>69</td>
</tr>
<tr>
<td>Figure 2-2 Phosphorylation of the artificial peptide substrate for Syk (SASTide) <em>in vitro</em></td>
<td>72</td>
</tr>
</tbody>
</table>
Figure 2-3 Detection of stimulant dose-dependent intracellular Syk activity ................. 73
Figure 2-4 Detection of Syk inhibitor dose-response ............................................. 75
Figure 2-5 Detection of time-dependent Syk activity following cell stimulation .......... 76
Figure 2-6 Intracellular specificity of the artificial peptide substrate for Syk ............ 79
Figure 2-7 Detection of Syk activity and inhibition in mouse primary splenic B-cells .... 81
Figure 3-1 Time-resolved luminescence detection of Syk kinase activity using SAStide ......................................................................................................................... 100
Figure 3-2 Detection of phosphorylation of SAStide through terbium luminescence .... 106
Figure 3-4 Calibration and detection of phosphorylation of SAStide ......................... 110
Figure 3-5 Detection of dose-dependent Syk inhibition ......................................... 112
Figure 4-1 Design and development of phosphorylation-dependent enhanced Tb$^{3+}$ luminescence tyrosine kinase peptide biosensors ......................................................... 137
Figure 4-2 Analysis of Amino Acid Preference and Substitution Effects on Substrate Specificity ............................................................................................................... 145
Figure 4-3 Identification, validation, and characterization of kinase specific biosensors using KINATEST-ID ........................................................................................................ 149
Figure 4-4 Quantitative time-resolved phosphorylation-enhanced Tb$^{3+}$ luminescence detection of nonreceptor tyrosine kinase activity and inhibition ......................... 154
Figure 4-5 A high-throughput chemical screen using AbASTide biosensor identifies new inhibitors of Abl tyrosine kinase ...................................................................................... 158
Figure 5-1 Multiplexed detection using time-resolved lanthanide-based resonance energy transfer (TR-LRET) and fluorophore conjugated peptide biosensors ............. 177
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 5-2 Time-Resolved Lanthanide-based Resonance Energy Transfer (TR-LRET) detection of phosphorylation-dependent signals and cross-interference</td>
<td>180</td>
</tr>
<tr>
<td>Figure 5-4 TR-LRET quantitative calibration of biosensor phosphorylation</td>
<td>181</td>
</tr>
<tr>
<td>Figure 5-5 Simultaneous multiplexed <em>in vitro</em> detection of Syk and Lyn kinase activities</td>
<td>183</td>
</tr>
</tbody>
</table>
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbAStide</td>
<td>Abl Artificial Substrate Peptide</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>D$_2$O</td>
<td>Deutrium oxide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>E</td>
<td>The percentage of energy transferred</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetraacetic acid</td>
</tr>
<tr>
<td>Fmoc</td>
<td>Fluoren-9-ylmethoxycarbonyl</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-(2-hydroxyethyl)peperazine-N’-ethanesulfonic acid</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>JAStide</td>
<td>Jak2 Artificial Substrate peptide</td>
</tr>
<tr>
<td>$K_D$</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>$K_m$</td>
<td>Concentration of substrate at half the maximum velocity</td>
</tr>
<tr>
<td>$k_{cat}$</td>
<td>Catalytic efficiency of the enzyme to modify the substrate</td>
</tr>
<tr>
<td>KASStide</td>
<td>Kinase Artificial Substrate peptide</td>
</tr>
<tr>
<td>KINATEST-ID</td>
<td>Kinase Terbium Emission Sensor Identification</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Wavelength</td>
</tr>
<tr>
<td>$\lambda_{em}$</td>
<td>Emission wavelength</td>
</tr>
<tr>
<td>$\lambda_{ex}$</td>
<td>Excitation wavelength</td>
</tr>
<tr>
<td>LRET</td>
<td>Lanthande-Based Resonance Energy Transfer</td>
</tr>
<tr>
<td>$q$</td>
<td>Number of water molecules coordinated to a lanthanide ion</td>
</tr>
<tr>
<td>$R$</td>
<td>Distance</td>
</tr>
<tr>
<td>$R_0$</td>
<td>The Foster distance</td>
</tr>
<tr>
<td>RFU</td>
<td>Relative fluorescence units</td>
</tr>
<tr>
<td>SAStide</td>
<td>Syk Artifical Substrate Peptide</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard Error of the Mean</td>
</tr>
<tr>
<td>SFASStide</td>
<td>Src-family Artificial Substrate Peptide</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Lifetime</td>
</tr>
<tr>
<td>$\tau_D$</td>
<td>Lifetime of the donor</td>
</tr>
<tr>
<td>$\tau_A$</td>
<td>Lifetime of the acceptor</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>Maximum catalytic turnover velocity of the enzyme</td>
</tr>
</tbody>
</table>
CHAPTER 1. INTRODUCTION

1.1 Protein kinases

Intracellular phosphorylation catalyzed by protein kinases leads to the integration of information from the extracellular environment to the cell nucleus. Phosphorylation affects every basic cellular process including metabolism, growth and differentiation. Protein kinases catalyze the transfer of the γ-phosphate from adenosine triphosphate (ATP) to the amino acid residues serine (Ser), threonine (Thr) and tyrosine (Tyr) in eukaryotes. Kinases account for ~2% of the human genome and include at least 500 genes, which makes them one of the largest families of genes in eukaryotes. (1) Protein kinases can be classified into two subfamilies based on their ability to phosphorylate Tyr or Ser and Thr residues. These classifications can be further divided into receptor or nonreceptor based on localization of the kinases. It has been estimated that 30% of all proteins are phosphorylated on at least one residue and 100,000 phosphorylation sites have been identified in the proteome to date. The number of sites phosphorylated by each kinase varies greatly and this variation can be accounted for by the mechanisms responsible for kinase substrate recognition.

1.2 Kinase substrate recognition

Deciphering the kinases that are responsible for phosphorylation of the various sites in the proteome is a difficult task. Protein kinases contain structurally similar catalytic
domains comprised of an N-terminal lobe composed of β-sheets and a C-terminal lobe containing α-helices. Between the two lobes is a catalytic cleft which is responsible for binding ATP and the substrate sequence as well as catalyzing phosphotransfer from ATP to the hydroxyl group of the phosphorylatable residue. (2–4) While all kinases share these similar structural folds, they differ in their recognition of the phosphorylation site residues and surrounding sequence. The depth of the catalytic cleft promotes discrimination between Ser/Thr and Tyr residues for their respective kinases. Tyr kinases have a deeper cleft compared to Ser/Thr kinases to accommodate the longer and larger side chain of tyrosine. The ability of the catalytic cleft to identify the phosphorylatable residue makes a significant contribution to substrate recognition through its affect on the binding energy of the substrate to the kinase through this residue. This has been observed from the large difference in binding energy between substrate peptides and pseudosubstrates, in which the phosphorylatable residue is replaced with an unphosphorylatable structural mimic (i.e. substitution of tyrosine to phenylalanine, or serine to alanine). The PKA substrate Kemptide (LRRASLG) has a $K_m = 16 \mu M$, which is a surrogate for binding affinity, $K_d$; whereas the pseudosubstrate LRRAALG has a $K_i = 490 \mu M$, which is the dissociation constant of the inhibitor (pseudosubstrate)-PKA complex, a 30-fold decrease in affinity for the peptide. (5, 6) Likewise the Src substrate YIYGSFK has a $K_m = 55 \mu M$, while the tyrosine to phenylalanine substituted pseudosubstrate YIFGSFK displays significantly weaker interaction (by an order of magnitude) with a $K_i = 575 \mu M$. (7, 8) These results demonstrate that the phosphorylatable residue (or phosphorylation site) contributes a substantial amount of
binding energy to the kinase-substrate interaction, facilitating substrate recognition and
differentiation between Ser/Thr and Tyr kinases.

Kinase catalytic domains also differ based on the hydrophobicity and charge of
surface residues surrounding the catalytic cleft. These residues interact with the amino
carboxylates in the immediate vicinity of the phosphorylation site through electrostatic
interactions, hydrogen bonding, and van der Waals’ interactions contributing
significantly to the kinase-substrate recognition. The effects of these interactions on
kinase substrate recognition have been demonstrated by examining the effect of varying
substrate sequence lengths on kinase recognition. A peptide library containing a range of
sequences from dipeptides to hexapeptides was used to examine the effects of varying the
number of flanking residues as well as the charge and hydrophobicity using glutamic acid
and alanine. The $K_m$ values of the majority of these peptides were rather large (1-16 mM)
demonstrating a weak interaction between the substrate and the kinase. However, the
sequences containing five acidic flanking residues on either side of the phosphorylation
site (EEEEY or YEEEE) exhibited significantly reduced $K_m = 130 \mu M$. (9) These
observations suggest that a sufficient amount of interaction through an appropriate
number of contacts is required between the kinase and substrate to promote
phosphorylation of the sequence.

Additionally, the results described above suggest that the chemical properties of
the substrate cooperate with the residues in the kinase surrounding the catalytic cleft to
affect substrate recognition. The effects of the residues residing in the catalytic cleft on
substrate recognition have been demonstrated using a structural homology model of the
catalytic domains of Abl and Src, based on the crystal structure of insulin receptor
tyrosine kinase (IRK) bound to a peptide substrate. (10, 11) Substitutions of residues in the kinase and the substrate were made based on IRK substrate interaction, and peptide substrates and mutant kinases tested resulted in distinct changes in substrate kinetics. For Abl, the optimal substrate sequences contain the motif Ile-Tyr-Ala-X-Pro; substitutions of the substrate sequence at the critical positions -1, +1 and +3 resulted in significant changes in the $K_m$ and $V_{max}$ of the substrates. However, substitutions made to the catalytic domain residues did not change substrate residue preference and resulted in only modest changes in the $K_m$ values for the substrate. Significant changes were observed in the catalytic efficiency ($V_{max}/K_m$) of the mutant kinases due to elevated $V_{max}$ values. These changes could arise due to several factors, including the possibility of substituted residues not being directly involved in the interaction with the substrate. Recognition could involve multiple residues and it may be that a single substitution is not sufficient to alter the interaction. Another possibility is that the conformation of the catalytic domain is altered due to the substitution. Without a crystal structure of the kinase-substrate complex, it is difficult to fully assess and define the kinase domain residues involved in recognition of a substrate. Nevertheless, these models demonstrate the importance of kinase-substrate recognition determinants in both the catalytic domain and the substrate itself for maintaining kinase substrate specificity.

Another example of the contribution of the substrate sequence to kinase specificity was shown by substrate substitution effects on the autophosphorylation site in the A-loop of anaplastic lymphoma kinase (ALK). (12) Isoleucine at the -1 position was found to be an important determinant for ALK substrate recognition, as a substitution at this position resulted in a 10-fold decrease in phosphorylation compared to the wild-type sequence.
More importantly, the basic residues at the +6 and +7 positions resulting in 33-fold and 5-fold decrease in substrate phosphorylation, respectively. Substitutions at the -3, -2 and +1 did not have the same marked effect, resulting in just 2-fold decreases in phosphorylation. These distal interactions were demonstrated to facilitate substrate release rather than promoting other behaviors that could reduce phosphorylation, such as substrate inhibition. The substrate specificity of ALK found in this study exhibited that substrate specificity through kinase-substrate interactions can also occur with portions of the kinase outside of the catalytic cleft and substrate residues distal to the phosphorylation site.

Distal kinase-substrate interactions are able to increase the affinity of the substrate for the kinase, thus increasing the specificity of the substrate for the kinase. One of the most well studied examples is the control of mitogen activated pathway kinase (MAPK) substrate specificity through binding to the D and DEF domains (Figure 1-1). The D-domain with the consensus sequence (Arg/Lys)$_{2-3}$-(X)$_{1-6}$-$\Phi$-$X$-$\Phi$ is 50-100 residues away from the phosphorylation site, and increases the affinity of the substrate by binding to a complementary docking site on the opposite side of the kinase relative to the catalytic site. This docking site is composed of a highly acidic patch and a hydrophobic pocket that are complementary to the consensus sequence. (13) Many MAPK substrates also contain the DEF domain, which is characterized by the consensus sequence Phe-$X$-Phe-Pro. This sequence facilitates binding to a region adjacent to the catalytic cleft. (14) The application of these motifs individually and together has been shown to significantly enhance the specificity of peptide substrates not just for MAP kinases but also between
family members Erk, p38 and JNK by facilitating lower $K_M$ values, higher $V_{max}$ values and overall greater catalytic efficiency. \(15, 16\)

Figure 1-1 Representative schematic of MAPK substrate recognition through distal catalytic domain interactions. The kinase recognition sequence within the substrate interacts with the MAPK through the catalytic cleft.

The contribution of distal interactions between substrates and kinases has also been shown to extend beyond the kinase domain to noncatalytic domains of the kinase. For Src-family kinases (SFKs) the SH2 and SH3 domains facilitate important protein-protein interactions and have been demonstrated to assist in kinase recognition of substrates. This concept was demonstrated \textit{in vitro} by examining the effect of adding SH3 or SH2 domain ligands to an Hck peptide substrate. \(17, 18\) The addition of the protein-protein interaction domain ligands resulted in significantly lower $K_m$ values (10-fold) and increased the efficiency of phosphorylation of the substrate. This organization of kinase specificity through noncatalytic domains has been observed \textit{in vivo} as a mechanism to allowed highly related kinases have distinct functional roles. The mitogen activated pathway kinase kinases (MAPKKs) share closely related kinase domains, yet have very distinct functions in yeast. \(19\) By introducing the required interaction motif of the MAPKK Ste7 into other MAPKK family members, these other MAPKKs could be made
to perform the same functions (Figure 1-2). These results demonstrate the importance of protein-protein interacts in governing kinase substrate specificity.

Figure 1-2 Protein-protein interactions determine yeast MAPKK function. A) MAPK signaling cascade in yeast is facilitated through protein-protein interaction resulting in Fus3 phosphorylation by Ste7 B) Localization of other MAPKKs to the scaffold protein is not sufficient to recapitulate Ste7 function C) The Ste7-Fus3 interaction motif is necessary and sufficient for Mkk2 and Pbs2 to recapitulate the same function of Ste7.

The recombination of noncatalytic domains to alter kinase specificity and function is not limited to the same family or even closely related families of kinases. The exchange of SFK Hck’s SH2 domain with acidic motif recognition PDZ domain, not found in tyrosine kinases, results in redirecting the kinase specificity, observed by alterations in proteome wide phosphotyrosine. (20) This manipulation of the Hck modular structure resulted in rewiring of the Cas migratory pathway. Without the SH2 domain Hck could no longer interact with or phosphorylate Cas. Only when a PDZ domain ligand was engineered into Cas were the Hck interaction and phosphorylation of Cas reconstituted. This serves as a key demonstration that the noncatalytic domain interaction can redefine substrate specificity and rewire signaling pathways.

The rewiring of signaling networks through noncatalytic domain interaction has extended beyond substitution of protein-protein interaction domains to the use of scaffold
proteins. In one example, an artificial synthetic adaptor protein was used to facilitate a novel protein-protein interaction between the SH3 domain of SFK Hck and the α–helix of the hDM2 protein, resulting in phosphorylation of hDM2. (21) This further demonstrates the importance of protein-protein interaction in kinase substrate recognition. Together with the results described above for SH2 and SH3 domains, these results show that protein-protein interactions can play a dominant role in defining kinase substrate specificity, and new kinase functions can arise from recombining catalytic domains with interaction domains or facilitating novel interactions.

To summarize, kinase substrate specificity is determined by a number of factors, including the structure of the catalytic domain, the residues of the substrate sequence, distal interaction of the substrate sequence within the kinase domain, and noncatalytic domain interactions with other proteins. The ability to define the specificity of the catalytic domains as well as noncatalytic domains would provide insight as to how each kinase recognizes substrate proteins and the substrate sequence within the protein. These insights could then be applied to decipher the vast number of phosphorylation sites identified to date and in the future.

1.3 Methods for determining kinase substrate sequence specificity

The determination of the favorable residues in the substrate for each kinase at each position has proven to be a daunting task, as the sizes of combinatorial peptide libraries considering the key residues in the immediate vicinity of the phosphorylation site (-4 to +4 positions) comprise more than 2.5 billion sequences. Several approaches have been developed to use peptide libraries to empirically determine the substrate sequence features responsible for kinase specificity. These libraries can broadly be
categorized as encoded libraries, such as phage displays, or synthetic libraries, including oriented peptide libraries and one-bead-one-peptide libraries. Both of these categories of methods are based on generating large sequence libraries for kinase substrates, subjecting them to kinase reactions, selecting and identifying the phosphorylated library members, and then determining the features responsible for kinase recognition based on the patterns within sequences observed to be phosphorylated.

1.3.1 Genetically encoded peptide libraries

Nature is able to generate vastly diverse collections of functional biomolecules, including peptides, proteins and antibodies, by subjecting libraries of nucleic acids (genomes) to cycles of translation, selection, amplification and diversification. Genetically encoded peptide libraries mimic this biological process in vitro by exploiting the link between the properties of the library such as binding affinity or substrate specificity (the phenotype) and the genetic information encoded (the genotype). Translation of the encoded genetic constructs under appropriate conditions produces a library of peptides fused to a protein, RNA, or DNA. The library can be subjected to in vitro selection for desired properties, such as binding affinity or substrate specificity. The library members that contain the desired properties are enriched, the genetic information is amplified using PCR, and hits are identified using (for example) high-throughput sequencing (Figure 1-3). The use of genetic material to encode the library offers several key advantages for screening. The information content of DNA is sufficiently high that a 21-mer nucleotide could encode $1.28 \times 10^9$ sequences. Also the ability to amplify sub-femtomole quantities of DNA by PCR provides extremely sensitive and specific quantitative detection. As opposed to other screening methods such as small molecule
inhibitor screening libraries where each library member is examined individually, an entire encoded library can be screened simultaneously in a single experiment. This concept has been applied to examine protein-ligand interactions, antibody epitope binding, and protease substrate specificity. (22-24) In the next section, the application of genetically encoded libraries of peptide sequences and selection for the determination of kinase substrate specificity is reviewed.

![Diagram of in vitro screening and selection of genetically encoded peptide libraries](image)

**Figure 1-3** *In vitro* screening and selection of genetically encoded peptide libraries

1.3.1.1 **Phage display libraries**

Phage display libraries are constructed by using bacteriophages to translate the encoded peptide library as peptide sequences fused to minor or major coat proteins to be displayed on the outer coat of a filamentous phage with the encapsulated DNA within. The peptides can be fused to the most abundant coat protein pVIII, which allows for multiple copies to be displayed; however, instability of the phage coat is a concern with this approach. The use of the single copy pIII coat protein located at the tip of the phage
has been shown to more amenable, since it does not cause as much instability in the phage coat. Once the library is expressed it can be incubated with a purified kinase. Selection of substrate peptides can be achieved using affinity enrichment by immunoprecipitation or metal affinity. Identification of substrate peptides can also be performed using radioactive labeling with $^{32}$P-ATP during the kinase assay. Following selection, phages can be genetically sequenced to identify their displayed peptide substrate sequence. (25)

The use of phage display suffers from some limitations as a result of the library expression and display. Phage display libraries require an in vivo step to facilitate expression; the DNA library must be transformed into the bacteria resulting in limitation of the library sizes to $10^9$-$10^{10}$ sequences. Additional concerns arise from the need for the in vivo step with the use of phage display, including poor expression in the host, failure to process peptide-protein fusion to the phage surface, toxicity of the gene product, and maintenance of phage coat integrity. Together these issues make it difficult to ensure the entire sequence library is expressed. Nevertheless, this approach has been productive; for example, phage display was applied for the identification of substrate motifs for tyrosine kinases Syk, Src, Lyn and Blk. The following consensus motifs were identified for the kinases: Src, EXLYWXFX; Blk, XXIYDXLP; Lyn, DXGIEXLP; Syk, XXDYEXXX. (26)

1.3.1.2 mRNA fusion peptide libraries

mRNA display is an alternative cell free technique that eliminates the limitations associated with the in vivo step of phage display libraries. In mRNA display, the encoded
peptide library is fused directly to the mRNA by the formation of a covalent linkage between the mRNA template and the expressed peptide through 3’-puromycin. Puromycin mimics an amino-acyl tRNA, resulting in covalent binding to the nascent peptide following the peptidyl transferase activity of the ribosome. The coupling efficiency through is puromycin is low (10-40%), reducing the number of expressed sequence that can be identified. Despite this low efficiency the sequence complexity achieved is greater than phage display by 10,000 fold. (27)

A mRNA-peptide fusion library designed to display a tyrosine residue surrounded by five random amino acids was applied to interrogate v-Abl tyrosine kinase specificity. Following exposure to v-Abl, phosphorylated peptides were isolated through rounds of immunoprecipitation with the phosphotyrosine-specific antibody 4G10. Amplification of the mRNA by PCR allowed for DNA microarray identification of sequences. (28) v-Abl was found to select substrates with the motif I/L/V-Y-X₃₋₄–P/F which was in agreement with the results obtained by several other methods. Despite the success of this approach, the application of mRNA display is limited due to the instability of mRNA. This constrains the types of assays and selection conditions available to probe the library. The stability of the mRNA template can be improved by reverse transcription to form a RNA/DNA hybrid; however, this only provides a modest improvement.

1.3.1.3 cDNA display libraries

Using DNA fused to the peptide rather than mRNA provides greater stability for the library, allowing for the use of a variety of selection conditions in which RNA would otherwise degrade. The increased stability also allows for selection assays that require
longer experimental times, such as off-rate selections for affinity reagents such as antibodies and aptamers, as well as enzyme inhibitors. Also, the use of DNA allows for directed evolution of the selected sequences from the library due to the compatibility with PCR-based mutagenesis, which can’t be performed with mRNA display libraries. The construction of a cDNA display library is accomplished by first expressing the library as mRNA-peptide fusions as described above using a covalent puromycin linkage. Next, a DNA primer is crosslinked to the RNA by psoralen and the RNA is reverse transcribed to produce a single stranded DNA (ssDNA)/single strand RNA (ssRNA) hybrid. The RNA is removed from the ssDNA/ssRNA by RNase treatment and the single stranded cDNA-peptide fusion is transcribed to form a double stranded DNA (dsDNA)-peptide fusion. The dsDNA has been shown to be stable in cell lysates for over two hours making them ideal for cell-based selections, which provide more biologically-relevant systems than purified kinases. (29)

The substrate specificity of Abl was assessed using cDNA fusion by encoding a combinatorial mutation library of the previously identified optimal Abl substrate, Abltide, EGEAIYAAPFA. (30) This approach allowed for the entire sequence space to be examined through all iterations of this sequence to access single substitutions quantitatively. It was ultimately determined that Abltide was the optimal substrate for Abl, but a more detailed analysis of Abl specificity was also performed. This analysis revealed that Abl substrate sequence preference also included the motif E-A/L/P/G-I-Y-A/W/S/E/Q-X-P, where X is any amino acid.
1.3.1.4 DNA/PNA programmed synthetic libraries

The translation of the biosynthetic approaches of encoded peptide libraries to organic synthesis of peptide using DNA has many advantages over traditional peptide chemistry. The use of DNA encoded peptide libraries is an attractive synthetic approach for the development of a combinatorial library because the DNA not only facilitates construction of the library but also tags the sequence. Several distinct synthetic approaches have been developed to produce peptide libraries including DNA-templated, DNA-routed and YocotReactor syntheses. (31-33) In each of these methods, each unique DNA/PNA triplet contained in the genetic tag encodes a specific amino acid monomer. The synthesized libraries can be subjected to in vitro selection to identify individual compounds possessing the desired property through amplification and sequencing of the DNA template.

This approach was used to examine the substrate specificity of Abl, VEGFR2, and ErbB2/Her2 with a PNA encoded library containing 10,000 fluorescently labeled peptides with four variable positions in the sequence Phe-Gln-AA<sub>4</sub>-AA<sub>3</sub>-Tyr-AA<sub>2</sub>-AA<sub>1</sub>-Ile-Lys. (34) The library was exposed to the kinases and then hybridized to a DNA microarray. The array was probed for phosphorylation using the antiphosphotyrosine antibody 4G10 followed by a Cy3 conjugated secondary antibody (Figure 1-4). The substrate specificity profiles determined for Abl, VEGFR2 and ErbB2/Her2 were F-E-S/E-E-Y-E/S, K-E/K/S-Y-E/K/S-D-Pro and K/E/S/S/K/D-Ala-Y-D-Pro/S/D-Val-D-Pro/S respectively.

This approach provides several advantages for determination of kinase substrate specificity. This assay allows for radioactive isotope-free detection and rapid
identification of individual sequences with high sensitivity and accurate profiling, and also gives relative quantification of phosphorylation. Relative quantification is accomplished through normalization of the fluorescent phosphorylation signal to the fluorescent peptide signal, allowing for variation in hybridization and synthesis yields to be taken into account. The use of the DNA array permits rapid identification of the peptide sequences due to the hybridization to specific locations on the microarray. PNA encoded libraries also provide the additional advantage of allowing for the incorporation of non-natural amino acids, such as the D-isomers, into the peptide sequences.

Figure 1-4 Workflow of PNA-encoded peptide library for the determination of Abl kinase substrate specificity. The PNA library is translated through peptide synthesis. The library is then exposed to Abl kinase in an in vitro kinase assay. The library is hybridized to a DNA microarray and probed for phosphorylation.

1.3.2 Synthetic peptide libraries

Synthetic peptide libraries provide the most robust way to have absolute control over the sequence space available in the kinase screen. However, challenges arise due to
the time and cost required to synthesize large peptide libraries that are capable of exhaustively examining the entire sequence space. Also, without the genetically encoded tag present for identification, determining the individual sequences is a technical challenge. Several approaches have successfully used synthetic peptide libraries and sufficiently addressed the technical challenges of sequence identification to determine kinase substrate specificity.

1.3.2.1 Oriented Peptide libraries

Oriented peptide libraries are typically designed to cover the entire combinatorial sequence space (more than 2.5 billion sequences). The library of peptides is synthesized exhausting all possible combinations of amino acids, in which each residue is screened at each position. The entire library is exposed to the kinase of interest and the phosphorylated peptides are enriched using antibody-based or metal affinity chromatography. Due to the large number of sequences selected as substrates (~1%, 25 million sequences) it is usually impossible to identify each sequence individually; however, since each sequence is the same length, Edman sequencing can determine the abundance of each amino acid at each position in selected pools of sequences. This approach allows for universal determination of kinase substrate specificity and has been applied to 13 Tyr kinases and 18 Ser/Thr kinases. (35, 36) However, due to the selection of only phosphorylated sequences and the population-based determination of specificity, any potential information concerning interpositional dependence (in which certain amino acids at a given position influence the interactions of those at other positions) and/or non-permissive residues (residues that actively decrease phosphorylation as opposed to either
promoting or being neutral for phosphorylation) is lost. Additionally, these libraries are unable to determine substrate interactions outside of the -5 and +4 (which may influence specificity) without resulting in an exponential increase in the library size. Finally, the time and cost to synthesize this type of library and determine amino acid abundance can make this approach somewhat cumbersome for determination of kinase substrate specificity.

1.3.2.2 Positional scanning peptide libraries

To improve on the limitations of oriented peptide libraries, the synthetic library approach was modified by partitioning the entire sequence space into 198 sublibraries. Each library contains a fixed amino acid at one position and random residues at all other positions. Each sublibrary is exposed individually to the kinase in solution with $^{32}$P-labeled ATP, and then captured on streptavidin-coated nitrocellulose paper by a C-terminal biotin tag. The level of phosphorylation is typically assessed based on incorporation of radioactive phosphate. Rather than simply determining the abundance of each amino acid at each position represented overall, this method enables the effect of an individual amino acid at each position to be empirically determined. This approach also increases the throughput by eliminating Edman’s sequencing. This method has been successfully applied to determine the substrate specificity of 61 yeast kinases as well as at least 179 human kinases. (37-39). Like OPL, since only one amino acid is fixed and known at a given position, PSPL do not provide information about the individual sequences that were phosphorylated; therefore, any potential contextual information,
including the possibility of nonpremissive residues, cooperativity, and/or interdependent positions, is lost.

1.3.2.3 One-bead-one-peptide (OBOP) libraries

Solid supported synthetic combinatorial peptide libraries have also been applied for the determination of kinase substrate specificity. These peptide libraries are synthesized using solid-phase split-and-pool peptide synthesis such that in principle, only a single peptide sequence is conjugated to each individual bead of the solid support. The on-bead library of peptides is then exposed to the kinase of interest and substrate sequences can be detected by a variety of methods including radioactive labeling with \(^{32}\)P-labeled ATP followed by autoradiography. The identities of the phosphopeptides then must be determined by Edman sequencing. (7) The universal application of this method has been limited due to the requirement for Edman sequencing for determination of substrates.

An alternative approach to allow for higher throughput identification of individual peptide sequences is to apply mass spectrometry (MS) analysis following selection of the phosphorylated peptides. To facilitate MS sequence identification the OBOP library is synthesized to contain a sequence ladder by terminating the peptide during each amino acid coupling using competitive acetylation. (Figure 1-5A) (40) The ladder provides a mass fingerprint in the MS spectra facilitating analyses of matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) data and increasing the throughput of the sequence identification compared to Edman sequencing. (Figure 1-5B) Following selection of phosphorylated sequences using alkaline
phosphatase-conjugated antiphosphotyrosine antibody, the mass ladder can be liberated from the solid support via a UV-cleavable linker and analyzed by MALDI-TOF MS. This method has been applied for determination of Src, Zap-70 and Brk substrate specificity.

Figure 1-5 Synthesis of one-bead-one peptide ladder library. The library is synthesized with a photolabile linker (PLL) and spacer group (BEBE). The peptide sequence is synthesized with the Fmoc-amino acid and 10% (mol/mol) acetic acid resulting in acetylation and termination of the peptide sequence. This process is repeated until the peptide sequence is completed. B) Experimental workflow using one-bead-one-peptide ladder library for determination of kinase substrate specificity. The library is exposed to the kinase of interest and phosphorylated sequences are identified based on alkaline phosphate fluorescent. The positive beads are isolated and the peptide ladder is liberated from the bead by UV irradiation. The substrate sequence is then determined by MALDI-TOF analysis of the peptide ladder.

This method can be adapted for the identification of Ser/Thr kinase substrate specificity using chemical modifications of the phosphorylated Ser/Thr residues. Specific detection of Ser/Thr kinase substrates leverages the availability of β-protons and the presence of a good leaving group, phosphate. Under alkaline conditions phosphoserine and threonine undergo β-elimination resulting in the formation of dehydroalanine and β-methyldehydroalanine. (41) These residues provide an α, β-unsaturated system that can be chemically modified by the addition of a nucleophile-linked probe, such as an organic dye or affinity tag such as biotin, via a Michael addition. However, a notable limitation of
this approach is the β-elimination reaction can give rise to false positive selection of sequences, which could alter the interpretations of the specificity profile.

Recently, an alternative approach for the determination of kinase specificity that overcomes the limitations of the OBOP methods discussed has been developed using thiophosphorylation to label substrate sequences. The kinase reactions are performed in the presence of adenosine 5’-O-(3-thio)triphosphate, which resulted in the labeling the substrate sequences with a thiophosphate. (42) Phosphorylated peptides are selected using a pyridylidithio-containing fluorescent probe that results in a disulfide exchange covalently attaching the organic dye to the peptide. The identity of the positive sequences was determined by partial Edman degradation mass spectrometry (PED/MS), where the sequences are converted to a sequence ladder of progressively shorter fragment peptides by performing multiple cycles of partial Edman degradation similar to the acetylation terminated peptide ladder synthesis. The series of peptides can then be analyzed by MALDI-TOF MS to give the amino acid series for the sequence. This approach has successfully been applied to identify the substrate specificities of the Ser/Thr kinases Pim1 and MKK6 and the Tyr kinase Csk demonstrating the potential for universal application of the method to all types of kinases.

While the OBOP library methods described here are successful, they each still suffer from some limitations. The synthesis of the library requires a complicated method to produce the sequence fragment ladder, which must be optimized. If the whole sequence is synthesized the method still relies on the costly and time-consuming Edman’s degradation. While promising, the use of thiophosphorylation significantly reduces catalytic efficiency of some kinases unless conditions such as metal cofactor are
optimized resulting in lower sensitivity, as well as altering the kinetics of the substrates.

(43)

1.3.3 Computational methods for the prediction of kinase specificity

In the past, short sequence motifs have been used to attempt to describe the pattern of kinase substrate recognition. These motifs are generally insensitive and unspecific. For example, the motif [S/T]-[Q] has been used in the past to describe DNA protein kinase (DNA-PK) kinase substrate recognition. This motif is expected to occur often in the proteome and be recognized by many other Ser/Thr kinases including ataxia telangiectasia mutated (ATM) kinase. This means that without further elaboration, “traditional” substrate motifs to describe kinase substrate specificity offer little information. The identification of more complex and subtle patterns is required to accurately predict kinase substrate specificity with the necessary specificity and sensitivity.

The use of computational approaches provide an alternative method to probe deeper into the information obtained from experimental biology to predict and identify kinase substrates. Prediction of kinase substrate specificity is a binary classification problem as each sequence included in a training set could be classified as phosphorylated (substrate) or non-phosphorylated (non-substrate). The establishment of a method for the prediction of kinase substrate specificity requires 1) having a well-collected and curated data set of positive and negative sequences 2) identification of features to characterize the sequences as one of the two classes and 3) the development of a classifier trained from the known sequences that is capable of making the prediction for new sequences. Over the past decade more than 40 methods have been developed to predict phosphorylation...
sites. These methods differ in their approaches based on a number of attributes associated with them including the algorithm or machine learning technique used, the number of residues considered in a sequence, the incorporation of kinase sequence and/or structural information, focusing the prediction for individual kinases or families, and the training and testing data.

1.3.3.1 Position-specific scoring matrices (PSSM)

Position-specific scoring matrices represent the simplest of the machine learning methods for kinase substrate prediction. PSSMs are matrices that represent each amino acid at every position within the sequence. Each element in the matrix can represent a variety of features describing each amino acid depending the construction of the matrix. The elements can represent the frequency of a given amino acid at a particular position. The values can also represent the weighted contribution of each amino acid to the overall ability to be a substrate. (44) In any case, PSSMs are ideal for expressing the value of a residue at a particular position for promoting or impeding phosphorylation by a particular kinase. However, more complex patterns in the substrate motif such as cooperativity or interdependence of positions are overlooked. Additional features could provide insight into kinase substrate recognition such as secondary and tertiary structural information.
1.3.3.2 High order machine learning models

Artificial neural networks (ANNs) are computational approaches that attempt to mimic the processing that occurs in the human brain. These methods aim at having highly parallel, distributive computation learning to achieve a generalized understanding of the classification. The algorithm accomplishes this task by receiving inputs from a number of features, weighing each input and adding them together. If the total sum of the inputs is above an arbitrary threshold, the output will classify that sequence as a substrate. Weights of the inputs and the definition of the threshold can be determined in an iterative manner by training the algorithm with examples of both classifications (positive and negative). (Figure 1-6A) The application of this method for prediction of kinase substrates can take into account many features of kinase substrate interactions including the peptide sequence and any structural elements associated with the potential substrate and/or the kinase for the prediction. One of the inherent problems with ANNs is that it is difficult to determine why the algorithm makes a particular decision while weighting the features for classification; this leaves the algorithm as a “black box”, which makes it difficult to examine and understand the learning and classification process.
Another higher order model is Support Vector Machines (SVMs), which are kernel methods build on two concepts first to map the feature vectors in a nonlinear way to a high dimensional space, then utilize linear classifiers in the new space. This means that the algorithm manages the inputs and translates them into processed data that can be represented in a high dimensional space (for example, three features produces data represented in a 3-dimensional plot) and then separates the data sets into classes based on a linear function (plane). (Figure 1-6B) This method has been applied to kinase substrate specificity in the cell by using known information about the kinase recognition sequences, protein-protein interactions, and subcellular localization as input features to build the classification model.

One of the shortcomings of machine learning techniques for prediction of kinase substrates is that the model is only as good as the data used to develop the model. While some kinases are extremely well studied (Abl, Src and Erk had 710, 1,047 and 2,512 articles published studying each of them, respectively, in 2012 alone), other kinases are poorly studied (Ack, GRK4 and PRP4 had 4, 11 and 1 articles published studying each of them, respectively, in 2012 and only 64, 116 and 27 articles, respectively, total in Pubmed). These results are also reflected in the number of phosphorylation sites identified for each kinase. Examining the Phosphosite Plus database (www.phosphosite.org), which consists of over 13,000 phosphorylation sites (many from proteomics datasets), 220 of the 518 kinases in the kinome do not have a single identified and verified substrate. (Figure 1-7) This limited or even complete lack of data makes it difficult, and in some cases impossible, to generate an accurate model of kinase substrate specificity regardless of the machine learning approach applied.
Figure 1-7 Number of identified phosphorylation sites for each kinase annotated in the Phosphosite Plus data set.

1.4 The role of protein kinases in disease

As demonstrated above, kinase function is governed by substrate specificity through sequence recognition and protein-protein interactions. Other regulatory mechanisms, including regulation of the on/off state of the kinase by phosphatases (which remove phosphates), subcellular localization, and alterations in expression level or proteasomal targeting for degradation through ubiquitination, also contribute to the control of kinase activity. The synchronization of these processes is critical for maintaining normal cellular function, and deregulation of one these processes can drastically affect the dynamics of a kinase’s activity often leading to overactivation of the kinase. Given the importance phosphorylation plays in most cellular processes, it is not surprising that dysfunction in kinase activity results in numerous disease states including neurological diseases such as Alzheimer’s disease, metabolic disorders including type 2 diabetes and cancer.
A recent review examining 915 kinase mutations that underlie 67 single-gene diseases identified that more than 80% of mutations identified in kinases directly affected the catalytic domain. (45) These mutations result in gain- and/or loss-of-function in the kinase. Gain-of function mutations increase the activity of the kinase through an additional function; whereas loss of function mutations that lead to increases in kinase activity occur through mutational loss of inhibitory regulatory elements. Specific kinases have been seen to be mutated in a variety of diseases states, such as Btk in X-linked agammaglobulinemia, and Akt2 in type 2 diabetes and lipodystrophy, resulting in overactivation of kinase activity. (46, 47) Somatic mutations can also arise due to large chromosomal rearrangements such as translocations. The most well studied example of this occurs in chronic myeloid leukemia (CML), which is characteristically caused by the reciprocal translocation of chromosomes 9q and 22p forming the Philadelphia chromosome. This genetic aberration results in the fusion of the Breakpoint cluster region (BCR) gene and Abl tyrosine kinase gene (Abl) into the BCR/ABL gene. This gene product produces the oncogenic fusion protein Bcr-Abl, which possesses constitutive Abl tyrosine kinase activity. Constitutive activation is conferred through the loss of function of a normal Abl intramolecular inhibition mechanism due to the presence of Bcr on the N-terminus of Abl. Additionally, there is also a gain of function facilitated by the coil-coiled domain of Bcr, which results in trans-autophosphorylation and activation of the Abl kinase domain. (48)

Given the role protein kinases play in these diseases, and the direct link between the genotype of certain disease states and aberrant kinase activity, the development of therapeutic treatments targeting kinases has been a major focus of the pharmaceutical
industry over the past two decades. Despite this effort by the pharmaceutical industry, only 24 small molecule inhibitors have been FDA approved since the approval of the first kinase inhibitor imatinib in 2001. These inhibitors target only 18 different kinases (~5% of the kinome) and are only approved for the treatment of various cancers for which constitutive kinase activity is involved. (49) One of the rate limiting factors in the discovery of new small molecule inhibitors is the lack of specific probes that provide sensitive detection of kinase activity. Therefore there is a need for tools that provide sensitive, selective and high throughput monitoring of kinase activity.

1.5 Peptide biosensors for monitoring kinase activity

The use of peptide substrates for detection of kinase activity is attractive for a number of reasons including the specificity and sensitivity of kinase substrate recognition and interactions (as reviewed above), the compatibility of peptides with multiple chemistries and multiple detection platforms ranging from traditional biochemistry techniques such as radioactivity and ELISA to emerging techniques such as surface enhanced Raman spectroscopy (SERS) and fluorescence lifetime imaging. (50, 51) The following is a review of the current methods used for monitoring kinase activity using peptide substrates.

1.5.1 Traditional molecular biology techniques

The use of traditional molecular biology techniques such as Western blotting and radioactive labeling have been applied for the detection of kinase activity using peptide substrates. Radioactive detection using $^{32}$P-ATP has long been the traditional strategy for detecting peptide phosphorylation, but is unattractive for large scale implementation due
to the waste generated during detection and the specific considerations for disposal as well as the safety concerns.

Antibody-based detection has also been applied for monitoring the phosphorylation status of peptide substrates using traditional Western blotting. The development of new multiwell read-out platform technologies using fluorescence polarization, time-resolved lanthanide-based resonance energy transfer, and chemifluorescent solid-phase kinase assays has drastically increased the throughput for monitoring kinase activity in vitro. (52, 53) (Figure 1-8) However, detection in these formats requires phosphospecific antibodies or generic phosphoantibodies such as 4G10 for phosphotyrosine. Such an antibody does not exist for phosphoserine or threonine; however, recently introduced small molecule and nanoparticle-based phosphate affinity reagents including Phos-tag and pIMAGO show promise in providing more universal phosphate probes. (54, 55) Regardless, the use of generic phosphate affinity reagents also reduces the number of phosphorylation events that can be monitored without confounding the results. This prevents the detection of more than one phosphorylation event (i.e. multiplex detection). Additionally, the number of handling, wash, and labeling steps required with some antibody–based methods reduces the throughput as well as increases the cost per data point.
Figure 1-8 ELISA-based detection of peptide biosensor phosphorylation. Following exposure of the peptide biosensor to the cells, the cells are lysed and added to a Neutravidin™-coated 96-well plate. The substrate is captured from the lysate by the high affinity interaction of biotin and Neutravidin™. Phosphorylated substrates are labeled with the anti-phosphotyrosine antibody, 4G10. The primary antibody is labeled with a horseradish peroxidase-conjugated secondary antibody. Finally, Amplex Red™ is added to the well and oxidized to a fluorescent reporter in the presence of horseradish peroxidase.

1.5.2 Mass Spectrometry detection

Post-translation modifications of proteins have long been studied using mass spectrometry due to the unique mass changes associated with each modification, which enables label-free detection. Peptide phosphorylation can be detected by the shift 80 Da mass shift, and the peptide sequence and phosphorylation site can also be identified simultaneously through its mass fragmentation spectra. The ability to determine individual sequences based on the fragmentation patterns provides robust identification of peptides and allows for multiplexed detection of thousands of peptides in a single assay. Mass spectrometry also provides high sensitivity of detection, which allows for small amounts of biological material (cell lysate or number of cells) to be used. Additionally, MS detection is compatible with online and offline enrichment strategies such as metal chromatography for phosphopeptides, which can further increase the sensitivity of detection. Finally, MS detection can also provide quantitative detection through a number of techniques including the incorporation of isotopically labeled standards. (56) All of
these features make mass spectrometry an attractive detection platform for monitoring kinase activity using peptide biosensors.

A recently developed technique named “Kinase ActivitY Assay for Kinome profiling” (KAYAK) has enabled quantitative, simultaneous detection of phosphorylation rates for 90 different peptides from proteins representing selected signal transduction pathways from cell lysates in a single in vitro reaction. (57) Following quenching of the kinase reaction, a set of the identical stable isotope-labeled phosphopeptides are added to provide a quantitative internal standard. The phosphopeptides are enriched and detected using Orbitrap MS to allow for high sensitivity and mass resolution. This approach allows for phosphorylation signatures to be generated that can be associated with a particular cellular stimulus or inhibitor. A setback of this approach is that it has limited utility to study biology de novo. The peptides used to generate the signature via KAYAK are not specific enough to provide information about which kinases are active, and without a previously well-studied inhibitor or stimulus signature to compare against, it is difficult to discern the significance of the biology. Additionally, using cell lysate to perform the kinase reaction can result in drastic changes in the organization of signaling networks due to removal of the compartmentalization of the cell. This can result in alterations in the kinase activity, since changes in protein-protein interaction can occur. Therefore, measuring kinase activity in intact, live cells is desirable for achieving the most biologically relevant detection.

Recent work in our lab has aimed at circumventing these issues by delivering kinase-specific peptide biosensors into live cells to directly monitor kinase activity. (58) The design of these sensors leverages two degrees of kinase specificity by using both an
artificial kinase specific substrate and a specific protein-protein interaction domain ligand. The application of artificial substrates with multiple reaction monitoring mass spectrometry (MRM-MS) has the advantage that the tryptic fragment is an unnatural sequence providing a unique mass fingerprint that is unconfounded by the background of tryptic peptides from native proteins in the cell lysate. MRM-MS allows for greater sensitivity for detection in the fmol-amol range depending on the specific analytes and instrumentation. MRM detection can achieve this sensitivity because it uses targeted mass analysis, in which the mass analyzer focuses on detection of specific parent and fragment masses (termed “transitions”) while filtering out all other ions. Monitoring two transitions for an Abl-specific substrate provided confident identification of the peptide and quantificative detection with limits of detection (LOD) and limits of quantification (LOQ) of 0.5 fmol from a whole cell lysate.

1.5.3 Fluorescence detection techniques

Fluorescent techniques have become widespread in biological research. Fluorescent proteins, small molecule organic dyes, semiconductor nanoparticles and organometallic chelators have been developed and applied to monitor a wide array of cellular events including protein localization, protein-protein interactions and cation signaling. The application of these strategies has provided continuous real-time detection with high sensitivity and selectivity as well as high spatial and temporal resolution in the cell. The use of peptides for detection of kinase activity has been applied with a number of fluorescent techniques including lifetime measurements and anisotropy. The following is a review of peptide biosensor technologies using steady-state fluorescence and time-resolved measurements for detection of kinase activity.
1.5.3.1 Genetically encoded FRET-based biosensors

The discovery and development of fluorescent proteins including GFP and its engineered variants has led to the development of numerous cellular probes for monitoring virtually every cellular process. Perhaps the most widely used and extensively developed sensors for detection of cellular kinase activity are genetically encoded Förster resonance energy transfer (FRET) biosensors. These biosensors are designed to contain a FRET pair of two terminal fluorescent proteins with a kinase-dependent molecular switch that induces a conformation change upon phosphorylation. (59) The molecular switch is designed to contain a polypeptide substrate recognized by the kinase of interest and an appropriate phospho-recognition domain (i.e. SH2, 14-3-3, WD40 domains). Upon phosphorylation, the phospho-recognition domain binds the kinase substrate, altering the distance between the fluorescent proteins and resulting in detectable FRET signal (Figure 1-9). These sensors allow for real-time continuous monitoring of cellular kinase activity with high spatial and temporal resolution at the single cell level. Recently FRET sensors were applied to monitor the temporal regulation of Erk kinase activity at the single cell level following steady-state EGF stimulation. The ability to continuously monitor Erk activity revealed that Erk is activated in discrete pulses and the duration of the pulse controls S-phase entry.
The development of genetically encoded FRET biosensor suffers from a severe bottleneck due to a number of factors. The identification of sensitive and specific kinase recognition sequences that are simultaneously compatible with phosphobinding domains is a laborious process that has not yet been optimized to provide high throughput development. The sensitivity and amplitude of response of these sensors varies based on FRET pair, phosphobinding domain and conformation change induced by the phosphobinding domain and kinase recognition sequence. This requires time and effort to develop iterations of the sensor to achieve optimal detection. Even with optimization, these sensors display small dynamic ranges due to modest changes in signal (10-60%) making them only moderately quantitative and not ideal for detection of kinase activity compared to other fluorescent methods.

1.5.3.2 Nanoparticle-conjugates

Quantum dots (QDs) are nanometer-size luminescent semiconductors that possess many unique optical properties, making them excellent fluorescent probes. QDs have
high fluorescence quantum yield, robust photostability, broad absorption bands and narrow emission bands, which reduces spectral bleed through. These properties can be tuned by controlling the size, shape and composition of the QD. Also, the surface chemistry of QDs can be modified to allow for number of conjugation chemistries that are compatible with peptides. These conjugation chemistries include Michael-addition reactions, coupling through carbodiimide reactions, and succinimide ester activation. Additionally, traditional metal surface chemistries are also available for conjugation by self-assembly through chemisorption and electrostatic interactions. (61) These properties of QDs have been exploited for the detection of kinase activity through the use of QD-peptide substrate conjugates.

QD-peptide substrate conjugates have been applied for sensitive detection of the activity of the Ser/Thr kinase casein kinase 2 (CK2) using γ-adenosine triphosphate-Atto-590, which fluorescently labels the peptide substrate upon phosphorylation. (62) FRET and electron transfer quenching processes between the QD and the Atto-590 labeled peptide were used to monitor CK2 activity. In this system the QD is the electron donor and Atto-590 fluorophore is the electron acceptor; therefore, phosphorylation of the peptide results in decreasing QD fluorescence and increasing Atto-590 fluorescence. (Figure 1-10A) Quantification of phosphorylation was determined by the percent decrease in the fluorescence of the QD. A similar strategy was used to monitor Abl and Src kinase activity using a CdSe/ZnS QD-peptide conjugate and an AlexaFluor 647 conjugated anti-phosphotyrosine antibody for fluorescent labeling. (63) While these applications demonstrate the potential of QDs for kinase activity detection, the use of ATP-analogs and fluorophore labeled antibody increases the handling steps and prevents
these platforms from being applied for real-time detection or to *in vivo* models, which would require a label-free approach.

A label-free approach for the detection of kinase activity using QD-peptide conjugates exploits effect of surface charge on QD dispersion in solution. QDs are monodisperse in solution due to electrostatic repulsion; however, neutralizing the surface charge abolishes the electrostatic effects resulting in aggregation of the semiconductors. (Figure 1-10B) A cationic substrate for PKA when unphosphorylated was capable of neutralizing the surface charge of the QD leading to aggregation and diminished red-shifted fluorescence. (64) Phosphorylation of the substrate results in the addition of a negative charge to the peptide neutralizing the cationic charge and eliminating aggregation leading to an increase in blue-shift fluorescence emission.

![Figure 1-10 Quantum dot detection of kinase activity using A) phosphorylation-dependent fluorescent labeling and FRET and B) surface charge aggregation.](image)

**1.5.3.3 Environmentally-sensitive Fluorescence**

Environmentally-sensitive fluorophores (ESFs) display differences in their spectroscopic properties depending on the physiochemical conditions of their
environment. These fluorophores include 7-nitrobenz-2-oxa-1,3-diazole (NBD), pyrene, and dapoxyl. ESFs are characterized by large excited-state dipole moments and low quantum yields in polar solvents such as water but become highly fluorescent with blue-shifted emission in non-polar environments such as nonpolar solvents or hydrophobic binding pockets of proteins. The properties of these fluorophores have been exploited for the detection of kinase activity through the incorporation of ESFs into kinase peptide substrates positioned one to five residues from the phosphorylation site. Protein kinase-catalyzed incorporation of phosphate into the sequence in the proximity of the ESF should, in principle, result in a change in the local polarity of the fluorophore resulting in enhanced fluorescence emission. (Figure 1-11A) This approach has been applied to the detection of PKC activity. (65)

An alternative approach using ESFs combines kinase substrate recognition with phosphopeptide recognition domains such as phosphotyrosine-binding SH2 domains, phosphoserine/threonine-binding 14-3-3 and WD40 domains. Once the sequence is phosphorylated the phosphopeptide product is recognized by the binding domain introducing the ESF to the hydrophobic environment of the phosphorecognition domain-binding pocket resulting in a change in polarity and fluorescence. (Figure 1-11B) This strategy has been applied for the detection of Src kinase activity using the Src SH2 domain with dapoxyl; however, this robust result was not obtained with other ESFs. (66) The lack of universally robust detection with ESFs makes it difficult to design peptide-based sensors using this approach.
Figure 1-11 Environmentally-sensitive fluorescent peptide biosensors. A) Phosphorylation alters the local environment of the fluorophore leading to an increase in fluorescence. B) Phosphorylation of the peptide substrate results in SH2 domain binding and a change in the fluorophore environment leading to an increase in fluorescence.

1.5.3.4 Chelation-enhanced Fluorescence

Another approach for fluorescent detection of kinase activity using peptide substrates uses hard metal ions such as zinc, calcium, magnesium or lanthanide ions including terbium and europium to mediate fluorescence or luminescence detection. These sensors display low affinity for the metal ions in the unphosphorylated state and as a result have little to no fluorescence. Introduction of the hard phosphate di-anion increases the binding affinity of the sensors for the metal ions, causing enhanced fluorescent or luminescent signal corresponding to kinase activity.

1.5.3.5 Biologically available metal chelation

One strategy is the use of chelation-sensitive fluorophores containing heteroatoms that can coordinate biologically available divalent metal ions such as Zn$^{2+}$, Mg$^{2+}$ and Ca$^{2+}$. These fluorophores have different electronic structures when coordinated with a metal ion than in the absence of the metal ion, which results in difference in the spectroscopic properties. These changes can be attributed to a lone electron pair of the fluorophore, which quenches fluorescence by an electron transfer mechanism in the absence of the
metal ion; however, the involvement of the electron pair in metal coordination results in increased fluorescence. These sensors are designed to leverage the ability of phosphate to coordinate hard metal ions such as those mentioned above, so once incorporated in proximity to the fluorophores, the phosphate group can aid in coordination to produce the enhanced fluorescence signal.

1.5.3.6 Calcium-based chelation-enhanced fluorescence peptide biosensors

Detection of phosphorylation through calcium chelation has been facilitated by the development of fluorophores that combine the eight coordination positions of tetracarboxylate in 1,2-bis(2-amino-phenoxyethane-N,N,N’,N’)-tetracetic acid (BAPTA) with a xanthene chromophore. These fluorophores displayed strong Ca\(^{2+}\) coordination through the four carboxylic acids with binding constants in the range of 0.37-2.3 \(\mu\text{M}\). Calcium coordination changes the electronic configuration of the fluorophore resulting in 3-40 fold enhanced fluorescein fluorescence. (67) Placing the fluorophore next to the phosphorylation site of a peptide allowed for one of the iminodiacetate arms of the fluorophore to be replaced. (Figure 1-12A) This biosensor has been successfully applied for the detection of cellular PKC activity by exploiting the Ca\(^{2+}\) dependent activation of PKC. (68) However, the vast majority of protein kinases are not Ca\(^{2+}\) dependent making this approach limited to detection of PKC and other Ca\(^{2+}\) dependent kinases.

1.5.3.7 Magnesium-based chelation enhanced fluorescence peptide biosensors

To circumvent the dependence on calcium activation, an alternative approach is to achieve enhanced fluorescence through chelation of metals with consistently high intracellular concentrations such as Mg\(^{2+}\), K\(^{+}\) and Na\(^{+}\). Mg\(^{2+}\) has steady-state intracellular
concentration in the millimolar range making it an ideal metal for chelation-based
detection. The development of sulfonamide-oxime (Sox) fluorophores has utilized Mg for
enhanced fluorescence following phosphorylation for detection of kinase activity. The
Sox residue is separated from the phosphorylation site by a β-turn, which results in a
change in the directionality of the peptide and facilitates coordination between the
phosphorylation site and nonnatural fluorescent amino acids. (Figure 1-12B) Installation
of the β-turn is accomplished through a two amino acid sequence consisting of a proline
and glycine. Phosphorylation typically results in a 4-12 fold increase in fluorescence. (69)
However, the structural requirement for the β-turn hinders the ability of sensor design to
consider all aspects of kinase substrate recognition within the sequence. The development
of the next generation C-Sox, which consists of an alkylated cysteine containing the Sox
chromophore, allows for greater flexibility in sensor design by alleviating the
requirement for a β-turn. (70) The next generation design of Sox-containing biosensors
has focused on kinase recognition, resulting in sensors with comparable catalytic
parameters to those of the parent optimal kinase substrate. (Figure 1-12C) (71)
1.5.3.8 **Lanthanide-based chelation-enhanced luminescence**

The lanthanide metals (Ln^{3+}), elements 57-71, also known as the “Rare Earth” metals, display many unique photophysical properties due to the presence of electrons in the 4f orbitals, which offers several advantages compared to biologically available hard metals. The luminescence emission of lanthanides produces large Stokes shifts (~100-200 nm) between the absorption and the emission compared to organic fluorophores (~20-30 nm). The emission spectra produce very narrow bands (~20 nm wide) and are insensitive to the chemical environment. The f-f electric dipole transitions that would lead to short lifetimes are Laporte-forbidden by parity rules, therefore the Ln^{3+} ions have extremely long luminescence lifetimes (millisecond for Tb^{3+} and Eu^{3+} and microseconds for Sm^{3+} and Dy^{3+}). The long lifetimes make lanthanides well suited for time-resolved measurements. Organic dyes and biological chromophores have fluorescence lifetimes in the nanosecond ranges, therefore time-resolved measurements will allow for these short
lived fluorophores to decay prior to collection of the longer-lived lanthanide-specific signal. This is ideal for a detection method because it allows for highly specific and sensitive signal from the sensor’s phosphorylation. (72)

Another direct result of Laporte-forbidden nature of the f-f transitions is that $\text{Ln}^{3+}$ have extremely low molar extinction coefficients. This results in the need for a chromophore to absorb the incident radiation and transfer it to the $\text{Ln}^{3+}$ to facilitate excitation. This has been accomplished by a variety of UV-absorbing chromophores including coumarin derivatives, tryptophan, and tyrosine. Excitation of the chromophore results in an excited singlet state. This singlet can either fluoresce or undergo intersystem-crossing to an excited triplet state. The triplet state can emit a photon and phosphoresce, or it can undergo energy transfer to and excitation of $\text{Tb}^{3+}$. The excited state of $\text{Tb}^{3+}$ can be either quenched by nonradiative energy transfer or can emit producing luminescence. Quenching can occur through nonradiative energy transfer to O-H oscillators due to the similarity in the energy levels between the excited state of $\text{Tb}^{3+}$ and the fourth overtone of water; therefore, water molecules coordinated to $\text{Tb}^{3+}$ can quench luminescence emission. This makes it important to consider the incorporation of an appropriate chromophore as well as the coordination environment provided by the sensor during the design of probes for the detection of kinase activity using lanthanides. (72)

1.5.3.9 EF-hand based sensors

The $\text{Ln}^{3+}$ are very similar to $\text{Ca}^{2+}$, with $\text{Tb}^{3+}$ and $\text{Eu}^{3+}$ having only slight differences in atomic radius and both elements are hard acids. The calcium-binding EF
hand domains contain five hard base side chains and one main chain carbonyl to facilitate coordination of Ca$^{2+}$. Upon binding, these domains undergo drastic conformational changes into well organized helix-loop-helix structures. In the past Tb$^{3+}$ has been used as a probe to determine Ca$^{2+}$ binding affinity of these domains as well as to survey for potential Ca$^{2+}$ binding sites in proteins. Zondlo and coworkers have leveraged the properties of the EF hand domains to develop kinase responsive sensors using Tb$^{3+}$ luminescence for detection. (73) To accomplish this design, serine (Ser) was used as a phosphorylation inducible mimic of glutamic acid (Glu) due to the similarities in electrostatic and Lewis base properties. The Glu at position 12 of the EF hand domain is critical for Ca$^{2+}$ binding through its bidentate coordination, and thus is conserved amongst EF hand domains. The replacement of this Glu with Ser generates a phosphorylation responsive motif. Unphosphorylated Ser is a poor mimic of Glu and results in poor Tb$^{3+}$ binding and minimal luminescence, while phosphorylated Ser provides the required bidentate coordination through the phosphate group, and has the appropriate side chain length to tightly bind Tb$^{3+}$ and enhance luminescence. (Figure 1-13A) Additionally, the residues on either side of this Glu are poorly conserved across EF hand domains, allowing some elements of kinase substrate recognition to be incorporated at these positions. This design has been applied to develop a sensor for Ser/Thr kinase PKA.

This approach has been modified to apply this sensor design to tyrosine kinases. The size of phosphotyrosine prevents it from mimicking Glu, but length of side chain allows for an alternative structure-based design by shifting the phosphorylation inducible mimic motif by one position to replace Glu. (74) This strategy was applied to the detection of Abl activity and inhibition. Similar to the first generation of Sox-based
chemosensors, the requirements for the conformational change and Tb$^{3+}$ binding somewhat limits the ability of the sensor design to consider all the necessary kinase substrate recognition features, ultimately limiting the catalytic efficiency and sensitivity of the sensors.

1.5.3.10 Macrocyclic terbium chelate sensors

Another approach using lanthanides for detection of kinase activity has employed binuclear Tb$^{3+}$ complexes. Similar to the dinuclear zinc-pyridylmethlamine complex, “phos-tag”, which binds the divalent phosphate anion selectively, Komiyama and coworkers have developed a dual 1,4,7,10-tetrazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) chelator incorporated with Tb$^{3+}$ for selective detection phosphotyrosine. This design allows one water molecule to be coordinated at each Tb$^{3+}$ ion and does not contain a chromophore for sensitization. Once bound to pTyr the water molecules are displaced and the aromatic ring of tyrosine can facilitate sensitization resulting in a 10-fold increase in luminescence. (Figure 1-13B) (75) This approach was shown to be selective in the presence of other phosphate contain molecules including ATP, ADP, GMP, GDP due to the required appropriate chromophore for sensitization being absent or too far away for energy transfer. The application of this technique for detection of Ser/Thr kinases is limited due to the lack of an appropriate chromophore attached to the phosphate containing Ser/Thr residue.

1.5.3.11 Chelator Based

Inspired by the design of the Ca$^{2+}$ chelating BAPTA-xanthene peptide conjugate described above, a similar design was applied for lanthanide-based detection of tyrosine
kinases Src and Abl. The chemosensor was designed to contain an iminodiacetate chelator, which would coordinate a lanthanide ion in the presence of phosphotyrosine contained within the generic Tyr kinase substrate. (76) Sensitization of both Tb$^{3+}$ and Eu$^{3+}$ was achieved through incorporation of the chromophore carbostyril 124 as an “antenna”. (Figure 1-13C) Phosphorylation of the sensor resulted in binding of Tb$^{3+}$ or Eu$^{3+}$ and a 10-fold increase in luminescence. This sensor design could also be applied to Ser/Thr kinases due to the incorporation of the antenna into the sensor. However, the low substrate efficiency resulting in the requirement for the kinase reaction to proceed for 20 hours and the lack of quantitative detection suggest that the sensors require further development to be of wider use.

Figure 1-13 Lanthanide-based CHEF biosensors. A) Phosphorylation inducible EF-hand domain B) Binuclear Tb$^{3+}$ based sensor C) Iminodiacetate chelator biosensor
1.6 Dissertation Objective

As demonstrated above, previous work has successfully established a number of techniques for the identification of optimal kinase specific substrates. These substrates have been applied to develop various peptide biosensors for specific detection of kinase activity through a wide range of methods. However, over the last decade there has been a bottleneck in the identification of kinase specific substrates and thus also in the development of peptide-based biosensors for novel kinases. Herein, strategies for the development of various novel peptide biosensors as well as novel detection methods will be presented.

Spleen tyrosine kinase (Syk) is critically important to B-cell function through the B-cell receptor pathway, but there are few methods that provide direct detection of Syk activity. Chapter 2 will describe the development of a novel modular peptide biosensor for live cell detection of Syk activity. The design of this sensor was inspired by a biosensor previously developed in the lab for detection of intracellular Abl activity. The Syk artificial substrate peptide (SAStide) sequence was identified using a semi-empirical bioinformatic analysis of Syk phosphorylation sites. The specificity of the biosensor will be demonstrated using a small panel of purified kinases as well as control cell lines. The application of this sensor for specific detection of Syk activity and inhibition in response to physiologically relevant stimuli and inhibitors will be demonstrated. The development of an enzyme-linked immunosorbent assay (ELISA)-based chemiluminescent method for detection of SAStide phosphorylation will also be discussed.

Based on the similarity of SAStide with the α-synuclein sequence surrounding Y125, which has been demonstrated to sensitize Tb³⁺ luminescence in a phosphorylation
dependent manner, the use of Tb for detection of Syk activity using SAStide was explored. Chapter 3 will report the use of SAStide to sensitize Tb in a phosphorylation dependent manner as well as characterization of the mechanism of sensitization through the photophysical properties of Tb$^{3+}$. The characterization of this method for detection of Syk activity and the application for monitor activity and inhibition in vitro will also be reported.

Expanding on this initial work of SAStide for phosphorylation-dependent enhanced Tb$^{3+}$ luminescence detection, the development of a pipeline to identify peptide substrates, but also substrates that are not only kinase specific, but that would also sensitize Tb$^{3+}$ in a phosphorylation-dependent manner will be presented in Chapter 4. Although previous methods have successfully identified kinase specific substrates, these methods require costly and laborious analysis and experimentation, which has lead to the aforementioned bottleneck in the development of kinase peptide biosensors for new kinases. New computational approaches have attempted to identify kinase substrate specificity and predict phosphorylation networks de novo from phosphoproteomic data. However, these approaches have yet to be applied to the development of new artificial peptide substrates. Chapter 4 will detail the development of a new algorithm aimed at identifying kinase substrate specificity and key determinants involved in determining specificity. Biochemical validation of the key determinants for kinase specificity as well as the predictive power of the algorithm will be examined. Additionally, the algorithm will also predict the ability of the sequences to sensitize Tb$^{3+}$ luminescence in a phosphorylation-dependent manner. The generation of novel biosensors for Abl, Jak2, and Src-family kinases and their application in kinase assays will also be detailed.
The development of multiple kinase specific phosphorylation-dependent Tb$^{3+}$ sensitizing biosensors inspired further engineering of the sensors to obtain simultaneous detection of the activity of two kinases in a single assay. To achieve this goal, the multiple emission bands of Tb$^{3+}$ will be exploited to excite conjugated fluorophores using the time-resolved lanthanide-based resonance energy transfer (TR-LRET) for multiplex detection. Chapter 5 details the development of this approach using SAStide and the Src-family kinase artificial substrate peptide SFAStide to monitor Lyn and Syk activity simultaneously.
1.7 References


specificity for protein kinase by using an improved one-bead-one-compound

protein kinases: detection methods and combinatorial peptide libraries,
*Biopolymers* 94, 753-762.

42. Trinh, T. B., Xiao, Q., and Pei, D. (2013) Profiling the substrate specificity of
protein kinases by on-bead screening of peptide libraries, *Biochemistry* 52, 5645-
5655.

thiophosphorylation in the presence of competing phosphorylation with MALDI-
TOF-MS detection, *Journal of proteome research* 4, 1863-1866.

44. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: Proteome-
wide prediction of cell signaling interactions using short sequence motifs, *Nucleic
acids research* 31, 3635-3641.

mutations in human disease: interpreting genotype-phenotype relationships,

46. Vorechovsky, I., Zhou, J. N., Vetrie, D., Bentley, D., Bjorkander, J.,


CHAPTER 2. A PEPTIDE-BASED BOSENSOR ASSAY TO DETECT INTRACELLULAR SYK KINASE ACTIVATION AND INHIBITION

2.1 Abstract

Spleen Tyrosine Kinase (Syk) has been implicated in a number of pathologies including cancer and rheumatoid arthritis and thus has been pursued as a novel therapeutic target. Because of the complex relationship between Syk’s auto- and other internal phosphorylation sites, scaffolding proteins, enzymatic activation state and sites of phosphorylation on its known substrates, the role of Syk’s activity in these diseases has not been completely clear. To approach such analyses, we developed a Syk-specific artificial peptide biosensor (SAStide) to use in a cell-based assay for direct detection of intracellular Syk activity and inhibition in response to physiologically relevant stimuli in both laboratory cell lines and primary splenic B cells. This peptide contains a sequence derived from known Syk substrate preference motifs linked to a cell permeable peptide, resulting in a biosensor that is phosphorylated in live cells in a Syk-dependent manner, thus serving as a reporter of Syk catalytic activity in intact cells. Because the assay is compatible with live, primary cells and can report pharmacodynamics for drug action on an intended target, this methodology could be used to facilitate a better understanding of Syk’s function and the effect of its inhibition in disease.
2.2 Introduction

Syk is a 72 kDa non-receptor tyrosine kinase originally isolated from bovine thymus and porcine spleen best known for its role in B lymphocyte development and activation. Loss of Syk expression results in perinatal lethality in mice and an arrest in the development of B cells at the pro-B to pre-B cell and immature to mature B cell transitions. Upon antigen binding to the B cell antigen receptor (BCR), the Src family kinase, Lyn initiates the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) on components of the BCR. Phosphorylation of the ITAMs in turn recruits and activates Syk, inducing its phosphorylation on multiple tyrosines including Y525 and Y526 in its activation loop. Following this activation of Syk, numerous signaling pathways are initiated leading to the activation of downstream transcription factors including NFAT, NFκB and Elk1 and ultimately contributing to the induction of cell proliferation and differentiation.

Dysregulation of the expression or the activity of Syk contributes to various disease states, making it a potential therapeutic target. Syk has been implicated as a factor in rheumatological disorders (such as rheumatoid arthritis) and malignant diseases of myeloid, lymphocytic and even epithelial origin. For instance, Syk was found to be constitutively active in primary blasts from a set of patients with acute myeloid leukemia (AML). Inhibition of Syk decreased the viability of these AML blasts in vitro and reduced the number of these cells infiltrating spleen and bone marrow in a mouse xenograft model. In some chronic lymphocytic leukemia cells (B-CLL), Syk is hyperactive despite exhibiting normal expression levels, and inhibition of Syk or silencing of Syk expression via siRNA decreases cell viability. In another example,
several peripheral T-cell lymphomas (PTCLs) exhibit aberrant expression of Syk. In these cells, siRNA silencing of Syk translation or inhibition of its activity with a kinase inhibitor (R406, Rigel Pharmaceuticals) induces apoptosis and blocks proliferation in cells with elevated Syk Y525/Y526 phosphorylation. These results suggest that Syk could be a novel therapeutic target for the treatment of PTCLs. Conversely, in breast cancer, which has an epithelial origin, Syk appears to have tumor suppressor functions: while Syk is expressed in normal breast epithelia, there is little to no Syk present in more metastatic breast cancer cells. Expression of Syk negatively affects motility and invasion in these carcinomas. Accordingly, to guide both the treatment of these cancers and the development of novel therapeutics, it would be beneficial to selectively measure Syk activity in patient samples.

There are three predominant methods currently in use to measure Syk activity: in vitro kinase assays, luciferase reporter assays of downstream transcription factors and phosphotyrosine antibody-based detection of Syk autophosphorylation or substrate sites. Each of these methods has drawbacks that make them less than optimal for both determination of Syk biology and translation to the clinical setting. In vitro kinase assays measure Syk activity post lysis—however, Syk’s function is integrally dependent on its binding partners. For example, in lysed cells proteins such as c-Cbl that normally modulate the function of the kinase (and which are known to be critical for obtaining biologically-relevant activation for Syk) can become separated from Syk; also, proteins normally found in different subcellular compartments are able to artifactually interact with Syk and alter its activity. Moreover, as a result of changes in Syk’s phosphorylation state via phosphatase activity and autophosphorylation in vitro, Syk’s
activity can change during in vitro kinase assays in ways that may not be relevant to its intracellular activity in a disease context. Transcription factor-driven luciferase reporter assays (another common read-out used for determining Syk activity) are performed in whole cells and maintain the endogenous context of Syk—however, these represent an indirect measure of Syk activity, with numerous proteins in the pathways between Syk and the transcription factors allowing potential disruption of additional components of these cascades. Phosphotyrosine antibody-based methods such as Western blots and Phosphoflow cytometry\textsuperscript{12} use phosphorylation sites in endogenous proteins, such as known Syk-targeted sites and/or Syk autophosphorylation sites, as surrogate reporters of Syk activity. However, because of the complex relationship between Lyn and Syk activation during B-cell receptor signaling, phosphorylation at endogenous sites (on Syk and other proteins) has not been confirmed to be Syk-specific in all cases. Furthermore, Syk is phosphorylated on multiple sites (by itself and other kinases including Lyn) including some that negatively regulate the kinase,\textsuperscript{11a} making it difficult to parse the complex relationship between the phosphorylation state of these native sites and the intrinsic activation of the enzyme. Besides these functional caveats, the fundamental limitation exists that antibodies for every potentially meaningful site of Syk phosphorylation are not available, and their development is subject to the uncertainties inherent in epitope and antibody generation. Therefore, an ideal method to measure Syk activity would be one that specifically monitors the ability of the kinase to catalyze a phosphotransferase reaction in an intact cell.

Peptide-based approaches offer an alternative method that allows for direct detection of kinase activity. Several peptide substrates for Syk have been previously
reported, including some based on endogenous protein sequences and others derived artificially from peptide libraries (through e.g. phage display).\textsuperscript{13} Peptides have been applied to the detection of Syk activity \textit{in vitro} using constrained tyrosine analogs and artificial peptides containing fluorescent amino acids.\textsuperscript{13b,c,14} Using these substrates, Syk activity has been detected in stimulated cell lysates and from recombinant protein, but so far they have not been used in live, intact cells that provide the necessary context for Syk function. We were interested in developing a peptide biosensor substrate for Syk that could be implemented to detect kinase activity in intact cells using a cell-penetrating peptide approach. TAT-labeled peptide substrates have been used for specific and sensitive detection of endogenous Akt activity in live single cells using capillary electrophoresis.\textsuperscript{15} Our previous work applied a similar approach to detect intracellular Abl activity in an over-expression model using a peptide biosensor and matrix-assisted laser desorption/absorption ionization-time of flight mass spectrometry (MALDI-TOF) or Western blotting.\textsuperscript{16} Here we present the development of a Syk specific artificial peptide substrate capable of detecting physiologically relevant Syk activity in live, intact cells following BCR engagement and oxidative stress using an ELISA-based assay.

2.3 Methods and Materials

**Cell culture and biological reagents.** The DG75 and DT40 B cell lines were grown to a density of 0.4 x 10\(^6\) cells/mL in RPMI-1640 medium containing 7.5% FBS, 1 mM sodium pyruvate, 100 IU/mL penicillin, 100 \(\mu\)g/mL streptomycin and 50 \(\mu\)M 2-metcaptoethanol. Additionally, DT40 cell medium contained 1% chicken serum. Anti-chicken IgM and anti-mouse IgM F(ab\(^\prime\))\(_2\) fragments were purchased from Bethyl Laboratories. Anti-human IgM was purchased from Rockland Immunochemicals.
Hydrogen peroxide was purchased from Mallinckrodt. Anti-Syk (N-19) was purchased from Santa Cruz Biotechnology. Anti-tubulin and anti-phosphotyrosine (4G10) were purchased from Millipore. Anti-phospho Syk (Y525/526) was purchased from Cell Signaling Technology.

**Peptide Synthesis and Purification.** The Syk peptide biosensor was synthesized using ‘Fast’ Fmoc solid-phase peptide chemistry with a Prelude Parallel Peptide Synthesizer (Protein Technologies). The synthesized peptides were purified using a C18 reverse-phase column on an Agilent 1200 preparative HPLC system. The peptide was characterized using liquid-chromatography mass spectrometry (LC/MS) on an Accela/LTQ system (Thermo-Finnegan) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/TOF) on a Voyager 4800 instrument (Applied Biosystems).

**In vitro kinase assay.** Recombinant Abl, Src and Lyn enzymes were obtained from a commercial source (Millipore). EGFP-conjugated Syk was isolated from DT40 chicken B cells stably expressing Syk-EGFP. Cells were lysed using a solution containing 1% Nonidet P-40, 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 1 mM sodium orthovanadate, 2 mM NaF and 1X mammalian protease inhibitor cocktail (Sigma). Syk-EGFP was immunoprecipitated using anti-GFP magnetic nanoparticle beads (MBL, Japan). Lysates were incubated with the beads (40 µl per 1x10^7 cells) for 1 h at 4°C. The kinase-bound beads were washed and then used in the in vitro kinase assay. Immobilized Syk-EGFP (55 nM per reaction, from 1.2 µg/µl stock, see supporting information Fig. S1), recombinant Lyn, Abl or Src kinase (0.1 U each) was incubated with kinase reaction buffer (500 µM ATP, 5 mM MnCl₂, HEPES, pH 7.2) containing the
peptide substrate at 25 µM. Aliquots (22 µL) were taken at designated time points and quenched in 0.5 M EDTA, pH 8.5 (22 µL). The quenched sample (1 µL) was diluted into ELISA wash buffer and analyzed as described below. For substrate comparison assays, kinase reaction conditions were as described above except that substrate concentrations were 4 µM, and concentration of enzyme used per reaction was 6 nM. The volume of aliquots diluted in an equal volume of quench buffer was 4 µL (for 8 µl total quenched volume), and the entire quenched amount was diluted into ELISA wash buffer and analyzed as described below. For characterization of SAStide kinetic parameters, substrate concentrations and reaction/quench volumes are given in Table S1.

**ELISA-based fluorescence detection.** Samples were incubated in a 96-well Neutravidin™ coated plate (15 pmol biotin binding capacity per well, Thermo Scientific) in Tris-buffered saline (TBS) containing 0.1% BSA and 0.05% Tween 20 for 1 h at room temperature on a short-radius plate shaker (600 rpm). Following incubation, wells were washed three times with wash buffer (TBS, 0.1% BSA, 0.05% Tween 20), then incubated with mouse anti-phosphotyrosine monoclonal antibody 4G10 (1:5000 dilution in wash buffer, 100 µL per well) for 1 h at room temperature with shaking. Wells were washed three times with wash buffer and incubated with horseradish peroxidase-conjugated goat anti-mouse immunoglobulin G (IgG) secondary antibody (Abcam) (1:1000 dilution in the wash buffer, 100 µL per well) for 1 h at room temperature with shaking. Wells were then washed three times with wash buffer and twice with sodium phosphate buffer (0.05 M, pH 7.5). For chemifluorescence detection, Amplex Red™ reaction buffer (100 µL total volume/well) consisting of Amplex Red™ reagent (50 µL) (Invitrogen), 20 mM H₂O₂
Fluorescence of Amplex Red™ was measured using a Synergy4 multiwell plate reader (Biotek) with an excitation wavelength of 532 nm and emission wavelength of 590 nm.

**Cell-based peptide biosensor assay.** Cells were cultured as described above, harvested and resuspended at a density of 8 x 10⁶ cells/mL (8 mL), then treated with the SAStide biosensor peptide (25 µM) for 15 min prior to stimulation with either anti-IgM antibody (5 µg/mL), H₂O₂ (3.33 mM) or both. Concentrations of peptide lower than 25 µM resulted in low signal to noise in detection of phosphopeptide using the ELISA-based read-out, and no toxicity was observed in the presence of the peptide at 25 µM (similar to what was observed in our previously published work on Abl kinase¹⁶). Aliquots of the cell suspension (1 mL) were harvested, lysed in PhoshoSafe Extraction Reagent (EMD Millipore) containing 167 mM EDTA and freshly prepared protease inhibitor cocktail (Roche) and flash frozen. Half the cell lysate from each sample was used for ELISA-based fluorescence detection and the other for immunoblotting. For dose-response experiments, cells were stimulated with varying concentrations of anti-IgM (2.5-10 µg/mL) or hydrogen peroxide (1-7 mM), harvested at 5 min post-stimulation and processed as described above. For immunoblotting, membranes were blocked in 5% goat serum for 1 h. All primary antibodies were incubated at a dilution of 1:1,000 for 1 h at room temperature and visualized using an HRP-conjugated secondary antibody (Pierce) and ECL reagents (PerkinElmer). Uniformity of the amount of peptide taken up was tested in a representative experiment using Syk-EGFP reconstituted Syk(-/-) DT40 cells, and while there was a very slight (but not statistically significant) trend towards higher peptide amounts over time, no significant difference was seen across conditions.
**Cell-based inhibition assay.** Cells (4 x 10^6 cells/ml) were pre-treated with varying concentrations of piceatannol or dasatinib for 30 min and with the SAStide biosensor peptide (25 µM) for 15 min. Cells were then stimulated with anti-IgM antibody (5 µg/mL) and H_2O_2 (1 mM) and harvested after 5 min as described above.

**Isolation of primary mouse splenic B-cells and primary cell biosensor assay.** B cells were enriched from mouse spleens via “panning” as previously described. Cells (6 ml, 5 x 10^6 cells/mL) were treated with vehicle (DMSO), piceatannol (50 µM) or dasatinib (100 nM) for 1 h and with the SAStide biosensor peptide (25 µM) for 15 min prior to stimulation. The cells were stimulated with anti-IgM F(ab')_2 (5 µg/mL). Cells were harvested at 0, 5, 10 and 15 min following stimulation, lysed and analyzed as described above.

2.4 **Results**

2.4.1 Development of a Syk biosensor

Using a Syk substrate consensus motif suggested recently by Geahlen (EXXDEEDYEXPXEPX) and incorporating information on the preferences at the variable sites of that motif gathered from phosphoproteomic studies,\(^{11a,18}\) we designed a putative Syk peptide substrate, DEEDYEEPDEP, which we named “Syk artificial substrate” or SAStide. We incorporated SAStide along with several functional modules to form a Syk biosensor peptide. These modules include a biotinylated lysine for affinity capture of the substrate, a cell penetrating peptide for delivery of the biosensor into cells and a photocleavable amino acid for release of the substrate from the rest of the biosensor in case mass spectrometry-based analysis is desired for future applications (Figure 2-1).
Figure 2-1 Syk peptide biosensor. The biosensor contains a biotinylated substrate reporter that is phosphorylated by Syk, a photocleavable linker for release of the reporter from the biosensor (for future mass spectrometry applications) and a cell penetrating peptide (TAT peptide) for delivery into live cells. Upon delivery of the biosensor to cells and stimulation of the B cell receptor, the substrate reporter is phosphorylated and can be enriched via the biotin tag and analyzed using an antiphosphotyrosine antibody in an ELISA-style multiwell plate format.

Phosphorylation of the biosensor by Syk was assessed using an *in vitro* kinase assay, incubating the SAStide biosensor with immobilized GFP-tagged Syk kinase (immunoprecipitated from DT40 chicken B cells) in the presence of ATP and MnCl₂ for varying times after which aliquots of the reaction mixtures were removed and kinase activity quenched by adding an excess of EDTA. Phosphorylated biosensor peptide was detected using a fluorescent ELISA-based assay in which the quenched reaction mixture was incubated in a 96-well Neutravidin™ coated-plate to allow for affinity capture of the biotinylated substrate. Each well had a biotin binding capacity of 15 pmol. The total amount of peptide in the quenched reaction mixture applied to each well was 20 pmol, which ensured that each well was saturated with peptide for analysis. The captured peptide was then incubated with an anti-phosphotyrosine primary antibody (4G10) followed by a horseradish peroxidase-conjugated secondary antibody. Chemifluorescent
detection was accomplished by incubating each well with Amplex Red™ reagent and hydrogen peroxide, which gives a fluorescence signal readout that is proportional to the amount of horseradish peroxidase-conjugated antibody in each well, and thus reports the degree of phosphotyrosine present (via the anti-phosphotyrosine antibody) (as described by Wu et al. \(^{19}\)). Relative fluorescence units (RFU) were measured, which provided a proportional representation of the amount of phosphorylated biosensor present. In control experiments to generate a standard curve, phosphorylated SAStide was demonstrated to bind reproducibly to the wells, and exhibited a linear increase in signal up to an amount of phosphopeptide per well of 0.5 pmol (beyond which saturation of the ELISA signal occurred). This validated that the amount of antiphosphotyrosine antibody-related signal was proportionally related to the degree of peptide phosphorylation. A substantial increase in signal over time was observed, demonstrating that SAStide was phosphorylated by Syk \textit{in vitro} (Figure 2-2A). To provide preliminary evidence for the specificity of this substrate for Syk, SAStide was also assayed using Src, Abl and Lyn kinases, none of which produced any significant signal for phosphopeptide (Fig. 2-2A).

We also compared the phosphorylation of SAStide over time with that of two other known substrates of Syk kinase, the natural peptide parent of the previously reported ‘Syktide’ (EDDEYEEV, which in the original work contained an unnatural amino acid in place of the tyrosine that the authors demonstrated conferred specificity for Syk over other kinases e.g. Lyn)\(^{13b}\) and a Syk substrate discovered from a phage display library (EDPDYEWPSA).\(^{13d}\) Each was synthesized with glycine spacers on the C- and N-terminal sides of the substrate portion with a biotinylated lysine included at the C-terminus for capture and ELISA-based detection of phosphorylation signal (sequences are
given in the caption for Fig 2-2). The kinase assay was performed using immunoprecipitated Syk-EGFP (6 nM) with each substrate at 4 µM. We found that SAStide was phosphorylated more rapidly than the previously reported Syktide sequence, but not as rapidly as the phage display-derived sequence (Fig. 2-2B). Performing the assay on SAStide at a range of concentrations (Fig. 2-2C) enabled the characterization of its substrate kinetics with the immunoprecipitated Syk-EGFP enzyme as sigmoidal, indicating some substrate activation through an as-yet unknown mechanism. The kinetic parameters were characterized as $K_{\text{half}} = 86\pm10$ µM (which represents the substrate concentration at half of $V_{\text{max}}$ in a sigmoidal curve and is analogous to $K_m$) and $V_{\text{max}} = 11\pm1$ pmol/min, from which $k_{\text{cat}} = 216$ min$^{-1}$ and $k_{\text{cat}}/K_{\text{half}} = 2.5$ were derived (Fig. 2-2D). Previously reported for Syktide$^{13b}$ were $K_m = 3$ µM, $k_{\text{cat}} = 62.2$ min$^{-1}$ and catalytic efficiency $k_{\text{cat}}/K_m = 20.73$ (the kinetic parameters for the Phage substrate have not been reported). The SAStide catalytic efficiency is overall comparable to Syktide, with the caveat that the assays were performed with different enzyme preparations and under different substrate (ATP) concentrations. In the direct comparison shown in Fig. 2-2B (under conditions that emphasize the effects of $k_{\text{cat}}$ rather than $K_m$ on substrate phosphorylation), SAStide was phosphorylated more efficiently than the parent Syktide. Compared to SAStide, the phage display peptide was more efficiently phosphorylated, but exhibits substrate inhibition at higher concentrations (greater than 20 µM).$^{13d}$ SAStide exhibited no substrate inhibition at concentrations up to 100 µM.
Figure 2-2 Phosphorylation of the artificial peptide substrate for Syk (SASTide) in vitro. A) The SASTide biosensor (25 µM) was incubated with Syk-EGFP (closed circles), Lyn, Abl or Src in an in vitro kinase assay. Aliquots were removed at designated time points and the amount of phosphorylated substrate was measured using ELISA-based detection (given as relative fluorescence units, RFU). B) SASTide (GGDEEDYEEDPEPGEKKGG) and two known Syk peptide substrates, ‘Syktide’ (GGEDDEEEVGGKGG) and ‘Phage,’ (GGEDPDYEWPSAGGKGG), were incubated with immobilized, immunoprecipitated Syk-EGFP kinase (Syk-EGFP) as described in the Materials and Methods. Substrate concentration for each was 4 µM. C) SASTide was assayed with Syk-EGFP at a range of concentrations. D) Initial velocities were calculated from the data in (C) and plotted and fitted to a sigmoidal curve to fit $K_{\text{half}}$, $V_{\text{max}}$, $k_{\text{cat}}$, and $k_{\text{cat}}/K_{\text{half}}$. For all panels, data points represent the average of three replicate experiments and error bars indicate the standard error of the mean.

2.4.2 Detection of Dose Dependent Activation and Inhibition of Syk in intact cells

We next examined the ability of the biosensor to detect dose-dependent activation of Syk in the context of BCR activation and oxidative stress in intact, living Burkitt’s lymphoma DG75 B-cells. DG75 cells incubated with the SASTide biosensor were stimulated by cross-linking the BCR using polyclonal antibody specific for IgM or by treatment with hydrogen peroxide to inhibit protein tyrosine phosphatases and activate Syk. The stimulated cells were harvested and lysed in Phosphosafe extraction buffer with EDTA.
and protease inhibitors (to prevent further kinase activity, dephosphorylation by phosphatases and proteolytic degradation in the sample), and the lysates were incubated in Neutravidin™-coated wells to capture the biosensor peptide. The chemiluminescent ELISA-based assay described above was used to detect phosphorylation of SAStide. Syk activity was analyzed after 5 min and compared to unstimulated cells as a control. No signal above the reagent background was observed in cells not treated with the peptide. BCR stimulation resulted in increased phosphorylation of the biosensor in a dose-dependent manner (Figure 2-3A). Induction of oxidative stress in the B cells via hydrogen peroxide treatment also resulted in dose-dependent increases in phosphorylation of the biosensor (Figure 2-3B). These results show the ability of the peptide biosensor to detect dose-dependent changes in the Syk activity at endogenous levels of expression in live cells.

Figure 2-3 Detection of stimulant dose-dependent intracellular Syk activity. (A) DG75 cells were treated with the SAStide biosensor (25 µM) 15 min prior to stimulation. The cells were stimulated with varying concentrations of anti-IgM (A) or varying concentrations of H₂O₂ (B). Cells were harvested 5 min following stimulation and the amount of phosphorylated biosensor was measured. Experiments were performed in triplicate and the data are reported as fold change compared to the unstimulated control (which exhibited levels of signal similar to background observed in the in vitro assay shown in Fig. 2, ~4000-8000 RFU). Data points represent the average of three measurements and error bars indicate the standard error of the mean.
We also examined the ability of the biosensor to monitor dose-dependent inhibition of Syk using the Syk-specific natural product inhibitor piceatannol and the Src-family kinase inhibitor, dasatinib (currently being explored as a therapeutic Syk inhibitor\textsuperscript{21}). The inhibitors were assayed in a dilution series from 1 mM – 100 pM (Figure 2-4A). We found that dasatinib had greater potency than piceatannol for inhibiting Syk phosphorylation of the biosensor. However, as expected from the literature\textsuperscript{22} and observed from the loss of tubulin and Syk immunoblot signal (Fig. 2-4B, lanes 17 and 18), high concentrations of piceatannol in the presence of BCR activation and oxidative stress were toxic to DG75 cells whereas dasatinib was not. The apparent IC\textsubscript{50} values (the concentration at which SAStide phosphorylation was decreased by 50% compared to the control uninhibited cells) were calculated using the Hill function to be 10.8±9.3 nM and 1.2±1.5 μM for dasatinib and piceatannol, respectively. These results were also consistent with the reduction in the level of Syk tyrosine phosphorylation as detected by Western blot analysis (Figure 2-4B).
Figure 2-4 Detection of Syk inhibitor dose-response. Burkitt’s lymphoma DG75 B cells were treated with varying concentrations of dasatinib (closed squares) or piceatannol (open circles) for 30 min and with the SAStide biosensor (25 µM) for 15 min prior to stimulation. The cells were then stimulated with anti-IgM (5 µg/mL) and hydrogen peroxide (1 mM) for 5 min and harvested. (A) The extent of biosensor phosphorylation was analyzed by ELISA. Data indicate averages +/- SEM of experiments performed in triplicate. (B) The level of total Syk, of Syk phosphorylated on Y525 and Y526 (Syk pY525/526) and of tyrosine-phosphorylated proteins (pTyr) were analyzed by Western blotting of lysates of cells not stimulated (NS), stimulated but not treated with inhibitor (NI) or treated with increasing concentrations of the indicated inhibitor. Tubulin was measured as an internal loading control.

2.4.3 Time-dependence of Syk activity following activation

To examine the time dependence of Syk activity following stimulation through the BCR and/or oxidative stress, DG75 cells were treated as above and Syk biosensor phosphorylation was analyzed every few minutes for the first 15 min following stimulation. BCR stimulation gave a rapid increase followed by steadily maintained phosphorylation of the biosensor (Figures 2-5A and B). As expected, the addition of
hydrogen peroxide following BCR stimulation gave a very robust increase in phosphorylation over the time course due to the amplified and extended BCR signaling.\(^{11a, 12a, 20}\) Oxidative stress alone also resulted in increased phosphorylation of the biosensor, peaking at 5 min and subsequently showing a slight decrease.\(^{20}\) The results of this experiment demonstrate that the SAStide biosensor is able to monitor time-dependent increases in Syk activity in live cells following stimulation.

Figure 2-5 Detection of time-dependent Syk activity following cell stimulation. DG75 B cells were treated with the SAStide biosensor (25 µM) 15 min prior to stimulation. The cells were stimulated with anti-IgM (5 µg/mL) (panel A and closed circles in panel D), H\(_2\)O\(_2\) (3 mM) (panel B and open squares in panel D) or both (5 µg/mL anti-IgM and 3 mM H\(_2\)O\(_2\)) (panel C and open triangles in panel D). (A-C), the expression of Syk, Syk phosphorylated on Y525 and Y526 (Syk pY525/526) and GAPDH (loading control) in cell lysates were analyzed by Western blotting. NS - no stimulation; NB - no biosensor (15 min harvest). (D), cells were harvested at varying time points following stimulation and analyzed for biosensor phosphorylation. Experiments were performed in triplicate. Data are reported as normalized change compared to the unstimulated control; error bars are shown either above or below data points for clarity. Data were analyzed using a repeat measure one-way ANOVA test and a Dunnet post-test. Statistical significance is indicated as follows: \(\ast P<0.05\), \(\ast\ast P<0.01\) and \(\ast\ast\ast P<0.001\).
2.4.4 Determination of the SAStide biosensor specificity in intact cells

While *in vitro* kinase assays (such as those described in Fig. 2-2) can provide useful preliminary comparisons of the activity of a kinase for a given substrate, they cannot necessarily be used to unequivocally demonstrate specificity or selectivity. In particular, using classical Michaelis-Menten kinetic parameters such as $k_{cat}/K_m$ as “specificity constants” to compare the activity of different enzymes on a single substrate is not relevant since it relies on assumptions that only apply to comparisons of the same enzyme and its interaction with different substrates.\(^{23}\) Therefore, we sought to demonstrate the selective phosphorylation of the SAStide biosensor in the context of the complex intracellular environment. The specificity of the biosensor as a substrate for Syk and not other tyrosine kinases in live cells was explored using DT40 chicken B-cells in which the endogenous gene for Syk has been eliminated by homologous gene targeting.\(^{24}\) Syk-deficient DT40 cells were exposed to the biosensor and stimulated using a polyclonal antibody specific for IgM. As a control, cells were treated with the biosensor but were not stimulated. Phosphorylation of the biosensor was then monitored via quenched aliquots every few minutes for a total of 15 min following stimulation (Fig. 2-6A). No significant change in phosphorylation of the biosensor was detected over the time course following stimulation when compared to unstimulated control cells. These results indicated that the specificity of the biosensor was maintained in living cells in the context of IgM engagement and BCR-activated signaling.

Since little tyrosine-phosphorylation was observed in anti-IgM-activated, Syk-deficient cells (Figure 2-6A), BCR-stimulated phosphorylation was amplified by the addition of hydrogen peroxide. As seen with the anti-IgM stimulation alone, no
significant change in the phosphorylation of the biosensor was detected over unstimulated cells (Figures 2.6A). Western blot analysis of tyrosine-phosphorylated proteins demonstrated amplified phosphotyrosine signaling compared to anti-IgM treatment alone, indicating activation of multiple tyrosine kinases and inhibition of tyrosine phosphatases (Figure 2-6E). Similarly, no significant change in the phosphorylation of the biosensor was detected in cells treated only with hydrogen peroxide (Figures 2-6D). This showed that even in the presence of amplified and extended BCR-related and other H₂O₂-related signaling the SAStide biosensor was not appreciably phosphorylated by other highly activated tyrosine kinases in Syk-deficient cells. This experiment serves as a highly relevant specificity control in the intracellular context, given that tyrosine kinase activity was clearly upregulated, yet none of these activated kinases phosphorylated the biosensor peptide.
Figure 2-6 Intracellular specificity of the artificial peptide substrate for Syk. Syk(−/−) DT40 cells or Syk(−/−) DT40 cells reconstituted with stably expressed Syk-EGFP were treated with the SASTide biosensor (25 µM) for 15 min prior to stimulation. The cells were stimulated by treating with anti-IgM (5 µg/mL) (panels C and F, closed circles in graphs A and B), H₂O₂ (3 mM) (panels D and G, open squares in graphs A and B) or both (5 µg/mL anti-IgM and 3 mM H₂O₂) (panels E and H, open triangles in graphs A and B), then harvested at varying time points. (C-H): the phosphorylation of proteins in cell lysates was measured by Western blotting using anti-phosphotyrosine (4G10), anti-Syk, antiphospho-Syk(Y525/526), antiphospho-BLNK(Y84) and anti-tubulin antibodies (loading control). NS - no stimulation; NB - no biosensor (15 min harvest). (A, B): chemifluorescence detection of Syk biosensor phosphorylation. The data are reported as normalized change compared to the unstimulated control. Data were analyzed using a repeat measure one-way ANOVA test and a Dunnet post-test; In graph A, no statistically significant difference was seen for any time point or treatment relative to control. Experiments were performed in triplicate.

To further support specificity, we performed the same set of experiments in Syk(−/−) DT40 cells that had been reconstituted with Syk-EGFP. In the presence of BCR stimulation by IgM, signal indicating phosphorylation of the biosensor peptide was increased approximately 2-fold over control (unstimulated Syk-EGFP-expressing cells)
and decreased slightly over time (Fig. 2-6B). When treated with H$_2$O$_2$ with or without concurrent BCR stimulation, biosensor phosphorylation signal increased more dramatically (to approximately 12-fold over control) (Fig. 6B). Phosphorylation of the biosensor was consistent with that observed for the Syk autophosphorylation site and the known Syk substrate BLNK (Fig. 2-6F-H).

2.4.5 Detection of Syk Activity and Inhibition in Primary Mouse Splenic B-cells

To examine the ability of the biosensor to monitor Syk activity and response to inhibitor treatment in primary cells expressing endogenous levels of Syk, we isolated mouse primary splenic B cells and treated these with the biosensor peptide in the presence of the Syk inhibitors piceatannol or dasatinib. Primary B cells were treated with 10 nM dasatinib, 1 µM piceatannol or vehicle (DMSO) for one h prior to stimulation. The SAStide biosensor peptide was added 15 min prior to stimulation, and the cells were then treated with anti-IgM F(ab')$_2$ (5 µg/mL) and harvested at 0, 5, 10 and 15 min following stimulation. As seen with the DG75 cells, stimulation of Syk activity following BCR engagement resulted in a rapid increase in phosphorylation of the SAStide biosensor (within approximately two minutes—the time required to handle and process an aliquot of cells collected immediately after stimulation) followed by a maintenance of phosphorylation over time as compared to control unstimulated cells, which showed only background levels of phosphorylation-related signal (Fig. 2-7). In the presence of each inhibitor the level of phosphorylation of the biosensor was decreased, giving a signal that was close to background levels. These results showed that the biosensor peptide was capable of detecting the activation and inhibition of endogenous Syk kinase expressed at normal levels in primary cells that exhibit physiologically relevant B cell receptor
signaling, and suggest that elevation of Syk activity is very rapid in response to B cell receptor engagement in these primary cells.

Figure 2-7 Detection of Syk activity and inhibition in mouse primary splenic B-cells. Mouse primary splenic B-cells were treated without (panel A and closed circles in panel D) or with the tyrosine kinase inhibitors dasatinib (100 nM) (B), (D △) or piceatannol (50 µM) (C), (D □) for 30 min prior to stimulation, then treated with the SAStide biosensor (25 µM, 15 min) and stimulated with anti-IgM F(ab′)2 (5 µg/µL). Cells were harvested at various time points following stimulation. (A-C) the expression of Syk and of Syk phosphorylated on Y525 and Y526 (Syk pY525/526) were determined by Western blotting. NS - no stimulation; NB - no biosensor (15 min harvest). (D) The data are reported as normalized change in chemifluorescence signal compared to the unstimulated control. N = 6; error bars indicate SEM.

2.5 Discussion

These studies present the development of a new versatile tool for direct and specific monitoring of intracellular Syk kinase activity in physiological contexts. We showed that we could develop a Syk-specific biosensor peptide by combining a peptide based on a Syk substrate consensus sequence\textsuperscript{11a,18a} with other modular units including a biotinylated lysine for affinity capture of the substrate and a cell penetrating peptide for delivery of the biosensor into cells. Similar to the Abl kinase biosensor we previously reported,\textsuperscript{16} this Syk kinase peptide biosensor (SAStide) did not cause toxicity at the concentration used in these studies, and was able to detect dose-dependent and time-dependent activation and inhibition of endogenous Syk using physiologically relevant stimuli in cultured cell lines.
as well as primary splenic mouse B cells. These results demonstrate the potential for this strategy to be used in a multiwell plate ELISA assay to analyze Syk activity in contexts that could include study of signaling processes in a basic research setting or even potentially monitoring therapeutic response in translational applications.

Crucially, specificity of the SAStide biosensor for detecting Syk activation was comprehensively demonstrated in a complex, biologically-relevant system using Syk-deficient DT40 chicken B cells. Syk plays a key role in B cell signaling, but is partly dependent upon activation of Lyn through antigen binding to the BCR. Stimulation of Syk-deficient cells by cross-linking the heavy chain of the BCR in the presence and absence of H$_2$O$_2$ allowed for the specificity of the biosensor to be assessed in the context of Lyn kinase activation. Phosphorylation of the biosensor was not increased above background levels and did not change over time following BCR engagement in the absence of Syk, even when overall tyrosine kinase signaling was amplified as a result of H$_2$O$_2$ exposure. Syk also plays a major role in oxidative stress signaling, and its activation during oxidative stress is the result of both its own activity (via autophosphorylation) as well as other protein tyrosine kinases. Other kinases also become activated during oxidative stress (as observed via antiphosphotyrosine blotting), yet there was no increase in phosphorylation of the biosensor in the absence of Syk under these conditions. Reconstitution of the Syk deficient cells with Syk-EGFP resulted in phosphorylation of the biosensor peptide under B cell receptor-activating conditions and in the presence of oxidative stress. This suggests that even in a complex cellular environment, our biosensor is selectively phosphorylated by Syk and not by Lyn or other activated tyrosine kinases in
these cells. Together these results demonstrate that our biosensor is specific for the
detection of Syk activity in B cells and B cell model systems.

Another advantage of this strategy is the ability to monitor kinase activation and
inhibition by compounds such as piceatannol and dasatinib in an intact cell. Isolation of a
kinase from the cellular environment can alter its function by removing regulatory
proteins, eliminating alternatively spliced variants, altering post-translational
modifications and/or disrupting subcellular compartmentalization. Additionally, isolation
of the kinase precludes evaluation of the contribution of off-target effects of the drug,
which could potentially affect efficacy (positively or negatively) via inhibition of
upstream signaling in addition to the direct inhibition of the target. While kinase
activation in intact cell contexts has been studied by detecting the phosphorylation of
known endogenous substrate sites, unambiguous detection of these sites requires high-
quality phospho-specific antibodies, and furthermore, not every endogenous site can
serve as a reliable generic ‘marker’ for the kinase activation state. For Syk, monitoring
phosphorylation of its tyrosines through surrogate sites (Y348, Y352, Y525/526) has
been the traditional strategy for determining its activation\textsuperscript{11a,21}—however, this does not
necessarily give an accurate report of Syk activity, since some of these sites are
phosphorylated by upstream kinases and some by Syk autophosphorylation, and
moreover, phosphorylation of some of these sites modifies the function of the kinase
while others do not. While Y525/526 is in the activation loop of Syk and is a major
autophosphorylation site in B cells\textsuperscript{26} and canonically phosphorylation of these sites would
suggest enhancement of kinase activity, the role of this phosphorylation in the functional
activation of Syk remains convoluted. Previous work using a Sox-based fluorogenic,
generic tyrosine kinase substrate from Invitrogen has demonstrated that substrate phosphorylation kinetics by activation loop-phosphorylated Syk kinase domain (i.e. the phospho-Y525/526 form) are similar to those for the unphosphorylated and double phenylalanine mutant (F525/F526) forms.\textsuperscript{14} This suggests that these residues are not accurate markers of intrinsic kinase activation, and that a selective tool that can monitor Syk’s activation state in complex cellular environments independently of these existing markers would be useful for delineating the effects of specific phosphorylation sites and kinase inhibitors on different components of the B cell receptor signaling cascade.

Using the SAStide biosensor, we also detected inhibition of endogenous intracellular Syk activity in a dose-dependent manner by the Src-family kinase inhibitors, dasatinib, and piceatannol. We found that the apparent IC\textsubscript{50} values were 10 nM and 1 \textmu M for dasatinib and piceatannol, respectively. These values are lower than those reported in the literature,\textsuperscript{21,27} however the literature values were determined using assay formats that were very different than the format employed here, and so cannot necessarily be directly compared. When examining Syk activation in the mouse primary B cells in the presence of these inhibitors, we observed levels of biosensor phosphorylation consistent with those observed in EGFP-reconstituted DT40-Syk(-/-) cells—however, according to Western blot analysis, Y525/526 remained phosphorylated under these conditions. Since these cells were isolated from the spleen, this could be related to the role of kinases acting upstream of Syk in signaling pathways.\textsuperscript{28} Along with the previous report that Y525/526 to Phe mutations did not affect the kinase activity, these data also support the idea that monitoring an endogenous phosphorylation site as a surrogate for Syk activity does not necessarily correctly report the activity of the kinase or response to the inhibitor.
treatment—whereas phosphorylation of this SASTide biosensor may be used as a marker for intrinsic activation of the kinase.

Aside from its potential utility in basic research on the function of Syk kinase, the straightforward workflow and compatibility of this biosensor substrate with the multiwell ELISA-style readout might be useful in a translational setting to determine Syk kinase pharmacodynamics in patient B cell populations. Despite the observations from the basic science research that Syk dysfunction plays a role in disease, the outcomes of clinical studies and trials for Syk inhibitors indicate that the benefits of blocking Syk activity in patients are less clear. A recent report of results from an ongoing clinical trial in non-Hodgkins lymphomas (#NCT00446095) of the Syk inhibitor fostamatinib (R788, a prodrug form of R406, Rigel Pharmaceuticals) described a range of effects depending on lymphoma subtype, which included DLBCL, follicular lymphoma, mantle cell lymphoma, MALT lymphoma, marginal zone lymphoma, CLL and SLL.29 The cohort was very small, particularly when dissected by lymphoma subtype, and response rates were low overall, with slightly better response in chronic lymphocytic leukemia than the other types. Notably, in a recent study the R406 form of this drug was found to be one of the least selective kinase inhibitors evaluated, in terms of binding affinity for kinases.30 This highlights the need for pharmacodynamically-relevant assays during the evaluation and development of so-called “specific” kinase inhibitors.

On the immune side, another Syk inhibitor (R112, Rigel Pharmaceuticals), was shown to be effective for relieving allergic rhinitis symptoms when delivered intranasally, suggesting that there could be a therapeutic benefit to targeting Syk (which is upstream of histamine release) for this purpose, and thus may provide an alternative to
antihistamine drugs. For rheumatoid arthritis, an earlier trial found that Fostamatinib provided a significant benefit (over placebo) for rheumatoid arthritis patients receiving methotrexate; however, a more recent follow-up study by the same group found that this inhibitor was not significantly beneficial in another subset of patients who had previously failed biologic therapy (e.g. rituximab or abatacept). As the authors noted, their “study was fraught with seemingly contradictory data.” Many other factors in these patients’ diseases (such as status of other inflammatory markers and ability to assess response metrics in the presence of confounding factors) may have contributed to the overall assessment of responses. Another intriguing possibility is that initial response to biologics requires Syk involvement, and therefore this patient subset is inherently unlikely to respond to an alternative Syk-targeting strategy. This is also consistent with evidence that lower Syk expression in acute myelogenous leukemia is related to a lack of response to another therapeutic antibody conjugated to a toxin (gemtuzumab ozogamicin). Either way, this collection of paradoxical and weak responses indicates that monitoring Syk kinase activity and function in the context of individual patients could aid the interpretation of outcomes for trials of Syk-targeting agents, and maybe someday help guide patient stratification for likelihood of benefit. This will be increasingly important as new, more functionally selective Syk inhibitors (e.g. the recently reported compound P505-15) are brought into the market and through clinical trials.
2.6 References


27. Geahlen, R. L.; McLaughlin, J. L., Piceatannol (3,4,3',5'-tetrahydroxy-trans-stilbene) is a naturally occurring protein-tyrosine kinase inhibitor. *Biochemical and biophysical research communications* 1989, 165 (1), 241-245.


CHAPTER 3. TIME-RESOLVED LUMINESCEENCE DETECTION OF SPLEEN TYROSINE KINASE ACTIVITY THROUGH TERBIUM SENSITIZATION

3.1 Abstract

Disruption of regulatory protein phosphorylation can lead to disease, and is particularly prevalent in cancers. Inhibitors that target deregulated kinases are therefore a major focus of chemotherapeutic development. Achieving sensitivity and specificity in high-throughput compatible kinase assays is key to successful inhibitor development. Here we describe the application of time-resolved luminescence detection to the direct sensing of Syk kinase activity and inhibition using a novel peptide substrate. Chelation and luminescence sensitization of Tb$^{3+}$ allowed the direct detection of peptide phosphorylation without any antibodies or other labeling reagents. Characterizing the Tb$^{3+}$ coordination properties of the phosphorylated vs. unphosphorylated form of the peptide revealed that an inner-sphere water was displaced upon phosphorylation, which likely was responsible for both enhancing the luminescence intensity and also extending the lifetime, which enabled gating of the luminescence signal to improve the dynamic range. Furthermore, a shift in the optimal absorbance maximum for excitation was observed, from 275 nm (for the unphosphorylated tyrosine peptide) to 266 nm (for the phosphorylated tyrosine peptide). Accordingly, time-resolved measurements with
excitation at 266 nm via a monochromator enabled a 16-fold improvement in base signal to noise for distinguishing phosphopeptide from unphosphorylated peptide. This led to a high degree of sensitivity and quantitative reproducibility, demonstrating the amenability of this method to both research laboratory and high-throughput applications.

3.2 Introduction

Protein kinases play a critical role in multiple cellular processes including growth, proliferation and apoptosis. The deregulation of this mechanism can cause an imbalance between proliferation and death leading to a number of disease states including cancer. Kinase inhibitor drug development is often confounded by the complexity of protein kinase signalling pathways, leading to advancement of inhibitors with substantial off-target effects (which may or may not be beneficial to efficacy) into clinical trials. Spleen tyrosine kinase (Syk) has recently become an attractive therapeutic target due to its role in multiple diseases, including rheumatoid arthritis, non-Hodgkin’s lymphomas and chronic lymphocytic leukaemia. However, the development of small molecule inhibitors specifically and selectively targeting Syk has been challenging, with the inhibitor fostamatinib (previously described as ‘specific’ for Syk) found to bind to at least 100 other targets with affinity better than that for Syk. Therefore, sensitive and specific tools for monitoring Syk activity are needed to assist in the development of Syk specific inhibitors through the rapid screening of therapeutic agents as well as for monitoring therapeutic response as a companion diagnostic in the clinic.

Peptide-based approaches can be used for measuring kinase activity either in vitro or in intact cells with a variety of detection strategies. Because they enable the
screening of many samples and conditions at once with high sensitivity using commonly available instrumentation, multi-well plate-based detection methods using fluorescence or luminescence are most compatible with drug discovery and clinical applications.\(^{12-13}\) Phosphorylation-sensitized lanthanide luminescence has been used as a high-throughput-compatible probe for kinase activity. Lanthanides have unique optical properties including large stokes shifts, long luminescent life times and narrow emission bands.\(^{14}\) Additionally, lanthanides are hard acids giving them preferred coordination with negatively charged species such as phosphate, carboxylate and carbonyl groups.\(^{15-16}\) However, lanthanides are poor absorbers and require a sensitizing chromophore to achieve luminescence as well as a coordination environment that excludes water to prevent solvent luminescence quenching.\(^{17}\) Accordingly, highly acidic peptides that contain appropriate chromophores can provide the necessary coordination environment and sensitization for luminescence.

Several approaches have been developed for the detection of tyrosine phosphorylation using terbium (Tb\(^{3+}\)) based on these properties. Phosphorylation of a peptide can dramatically increase the peptide’s affinity for terbium.\(^{18-24}\) Sensitization of luminescence is achieved either through a chromophore label or through the phosphorylated tyrosine itself.\(^{13, 18-20, 22, 25-26}\) Previously reported approaches provide approximately a 2 to 10-fold increase in steady-state luminescence upon phosphorylation, depending on the peptide sequence involved—however, since the unphosphorylated forms of most acidic peptides also bind Tb\(^{3+}\) and exhibit luminescence to some degree, signal to noise and background can be a challenge for distinguishing phosphorylated from unphosphorylated peptide.
Recently, we developed a Syk specific artificial substrate peptide (SAS
tide) consisting of an acidic motif with a central tyrosine capable of detecting intracellular Syk activity in response to physiologically relevant stimuli and pharmacological inhibition\(^{(27)}\).

Based on the similarity of the arrangement of carboxylate groups and the central tyrosine in this sequence with the motif found in the atypical phosphorylation dependent terbium sensitizing peptide of α-synuclein (Fig. 3.2A), we chose to explore the use of this peptide as a probe for the detection of Syk activity through terbium-sensitized luminescence.

Because (as with all fluorophores) lanthanide luminescent lifetimes are sensitive to the chemical environment of the metal ion,\(^{(28)}\) we also tested whether chelation of Tb\(^{3+}\) with the phosphorylated vs. unphosphorylated peptide could be more easily distinguished using time-resolved luminescence detection. In a time-resolved luminescence measurement, the short-lived background luminescence is allowed to decay before measuring the luminescence of the chelated Tb\(^{3+}\) (Figure 3.1). While the overall sensitivity for a time-resolved vs. a steady state measurement is usually lower because of the gating of the highest intensity initial signal, this can be mitigated to some degree by optimizing the integration times. The long luminescent lifetime of terbium enables use of longer integration times, and provides an opportunity to eliminate any interference from autofluorescence and other background within the sample increasing the specificity and signal to noise for detection. Here we describe the testing and characterization of this strategy as a method for detecting Syk kinase activity and inhibition.
Figure 3-1 Time-resolved luminescence detection of Syk kinase activity using SAStide

3.3 Materials and Methods

Peptide Synthesis and Purification. Peptides were synthesized at a 50 µmol scale using a Protein Technologies Prelude Parallel peptide synthesizer on CLEAR-amide resin (Peptides International). Coupling of standard Fmoc (9-fluorenylmethoxy-carbonyl)-protected amino acids (Peptides International) were achieved with HCTU in the presence of NMM in DMF for two 10 min couplings. Fmoc deprotection was achieved in 20% piperidine in DMF for two 2.5 min cycles. Side-chain deprotection and peptide cleavage from the resin was performed in 5 ml cocktail of trifluoroacetic acid (TFA); water: ethane dithiol (EDT): triisopropylsilane (TIS) (94:2.5:2.5:1). Peptides were precipitated and washed three times with cold diethyl ether. The peptides were dissolved in acetonitrile: water: TFA (50:50:0.1), flash frozen and lyophilized. The peptides were purified by preparative reverse-phase HPLC (Agilent Technologies 1200 Series) a using C18 reverse-phase column.
**Luminescence Emission Measurements.** Emission spectra (both steady-state and time-resolved) were collected on a Biotek Synergy4 plate reader equipped with a monochromator at 23°C in black 384-well plates (Greiner FluorTrac 200). For time-resolved measurements, spectra were collected after excitation at either 266 nm or 280 nm (as denoted in specific experiments) with a Xenon flash lamp followed by a delay of 50 µsec. A luminescence scan between 450-650 nm was collected in 1 nm increments with 1 msec collection time and 10 readings per data point. Sensitivity (an instrument parameter similar to gain) was adjusted as necessary and is reported where relevant.

**Job’s Plot.** The molar fraction of the pSAStide biosensor and Tb$^{3+}$ were continuously varied inversely of each other while maintaining a total molar concentration of 16 µM (i.e. 1 µM pSAStide and 15 µM Tb$^{3+}$, 2 µM pSAStide with 14 µM Tb$^{3+}$, 15 µM pSAStide and 1 µM Tb$^{3+}$) for each data point. Luminescent emission spectra were collected as described above and the area under the emission spectra was used as the parameter for quantification of complex formation as luminescent increases with complex formation.

**Binding Affinity.** Tb$^{3+}$ binding to SAStide and pSAStide was measured using Tb$^{3+}$ luminescence sensitized by the central tyrosine or phosphotyrosine residue of SAStide and pSAStide respectively. Tb$^{3+}$ was added to 100 nM of either peptide at final concentrations ranging from 0 to 20 µM. All experiments were carried out in 10 mM HEPES and 100 mM NaCl (pH 7.5) at a volume of 100 µL. After excitation of the samples at 266nm (pSAStide) or 280nm (SAStide), Tb$^{3+}$ luminescence emission spectra between 450 to 650 nm were collected for 1 ms following a 50 µs delay and 30 readings per data point. Background luminescence emission was subtracted from the peptide in the absence of terbium. The area under each spectrum was integrated and used as the metric
for quantification. The data were fit to Eq. 1 by using KaleidaGraph nonlinear curve-fitting software, where $I$ is the Tb$^{3+}$ luminescence at a given concentration, $I_{\text{max}}$ corresponds to the maximum Tb$^{3+}$ emission, $[\text{Tb}^{3+}]_T$ is the total Tb$^{3+}$ concentration, $[\text{P}]_T$ is the total peptide concentration and $K_d$ is the equilibrium dissociation constant.

$$I = I_m \times ([\text{Tb}^{3+}]_T + K_d + [\text{P}]_T) - \sqrt{([\text{Tb}^{3+}]_T + K_d + [\text{P}]_T)^2 - 4([\text{Tb}^{3+}]_T \times [\text{P}]_T)} / (2 \times [\text{P}]_T) \quad (1)$$

**In vitro kinase assay.** EGFP-conjugated Syk was isolated from DT40 chicken B cells stably expressing Syk-EGFP. Cells were lysed using a solution containing 1% Nonidet P-40, 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 1 mM sodium orthovanadate, 2 mM NaF and 1X mammalian protease inhibitor cocktail (Sigma). Syk-EGFP was immunoprecipitated using GFP-Trap_A beads (Chromotek). Lysates were incubated with the beads for 1 h at 4°C. The kinase-bound beads were washed and then used in the *in vitro* kinase assay (0.4 µg/µL). Syk-EGFP was incubated with the kinase reaction buffer (3.4 µg Syk-EGFP, 100 µM ATP, 10 mM MgCl$_2$, 1 µM Na$_3$VO$_4$, leupeptin, aprotinin, 125 ng/µL BSA and 25 mM HEPES pH 7.5, total volume 170 µL) containing SAStide at 37.5 µM at 30°C. Aliquots (20 µL) were taken at designated time points (0.5, 5, 10, 15, 30, 45, 60 and 90 min) and quenched in 6 M urea (20 µL). The quenched samples were then treated with the luminescence buffer (500 µM Tb$^{3+}$ and 500 mM NaCl, 10 µL) for a total volume of 50 µL (final concentrations of sample components: 2.4 M urea, 40 µM ATP, 4 mM MgCl$_2$, 0.4 µM Na$_3$VO$_4$, leupeptin, aprotinin, 50 ng/µL BSA and 10 mM HEPES pH 7.5). Luminescence emission spectra were collected as described above and the area under each spectrum was integrated using GraphPad Prism. An additional aliquot (1 µL) of the kinase reaction mixture was taken at each time point for validation of phosphorylation using an ELISA-based chemifluorescent assay as previously
described. Briefly, each aliquot was quenched with 0.5 M EDTA and incubated in a 96-well Neutravidin coated plate (15 pmol biotin binding capacity per well, Thermo Scientific) in Tris-buffer saline (TBS) containing 0.1% BSA and 0.05% Tween 20 for 1h. Following incubation, each well was washed with the TBS buffer and the incubated with mouse anti-phosphotyrosine monoclonal antibody 4G10 (Millipore, 1:10,000 dilution in TBS buffer) for 1h. Following incubation, each well was washed with TBS buffer and incubated with horseradish peroxidase-conjugated goat anti-mouse immunoglobulin G (IgG) secondary antibody (Abcam) (1:1000 dilution) for 1h. Wells were then washed and treated with Amplex Red reaction buffer (Amplex Red reagent, Invitrogen, 20 mM H$_2$O$_2$ and sodium phosphate buffer) for 30 min. Fluorescence was measured using a Synergy4 multiwell plate reader (Biotek) with an excitation wavelength of 532 nm and emission wavelength of 590 nm.

**Dose-Response Inhibition Assay.** Syk-EGFP (0.4 µg/reaction) was incubated with the kinase reaction buffer described above before adding SAS tide in the presence of DMSO (vehicle) or varying concentrations of piceatannol at 30°C for 10 min prior to the start of the reaction. The reaction was started with the addition of SAS tide (37.5 µM, total reaction volume 20 µL). Each reaction was quenched after 30 min in 6 M urea (20 µL). The samples were then treated with the luminescence buffer (500 µM Tb$^{3+}$ and 500 mM NaCl, 10 µL) for a total volume of 50 µL.

**High-Throughput Screening Calculations.** The Z' factor was calculated according to Eq. 2.

$$Z' = \frac{(\mu_{pos} - 3\sigma_{pos}/\sqrt{n}) - (\mu_{neg} + 3\sigma_{neg}/\sqrt{n})}{(\mu_{pos} - \mu_{neg})} \quad (2)$$

And the signal window was calculated according to Eq. 3.
\[ SW = \left( \mu_{\text{pos}} - 3\sigma_{\text{pos}}/\sqrt{n} \right) - \left( \mu_{\text{neg}} + 3\sigma_{\text{neg}}/\sqrt{n} \right) / (\sigma_{\text{pos}}/\sqrt{n}) \quad (3) \]

Where \( n \) is the number of replicates, \( \mu_{\text{pos}} \) and \( \mu_{\text{neg}} \) are the average luminescence of the positive (pSAStide or uninhibited) and negative (SAStide or inhibitor treated EGFP-Syk) controls respectively; \( \sigma_{\text{pos}} \) and \( \sigma_{\text{neg}} \) are the standard deviation of the positive and negative controls.

3.4 Results

3.4.1 Time-resolved luminescence measurements increase signal to noise

Phosphorylated (pSAStide) and unphosphorylated (SAStide) forms of the peptide were synthesized and steady-state luminescence of each 1:1 Tb\(^{3+}\) complex with excitation at 266 nm through a monochromator (for highly resolved excitation energy control) was measured (Fig. 3-2B). The optimal excitation energy was determined to be 266 nm for pSAStide and 275 nm for SAStide and were used for characterizing each species in complex with terbium. At 266 nm excitation energy, pSAStide exhibited strong Tb\(^{3+}\) sensitization and SAStide displayed weaker luminescence. Based on this we used 266 nm for all further analyses. While SAStide signal was somewhat mitigated by using 266 nm excitation energy, there was still weak but significant luminescence from SAStide so the signal to noise (comparing pSAStide and SAStide luminescence) was low (2:1).

However, time-resolved measurements significantly improved the signal to noise to 32:1 (Fig. 3-2C), demonstrating that taking advantage of the increase in Tb\(^{3+}\) luminescence lifetime in the presence of phosphotyrosine vs. unphosphorylated tyrosine improved the ability to detect phosphorylation of SAStide. The improvement in the signal to noise ratio was accomplished by allowing the short-lived signal of the unphosphorylated peptide to decay prior to collection.
Figure 3–2 Detection of phosphorylation of SAStide through terbium luminescence. A) Sequence alignment of a peptide from α-synuclein, an atypical terbium sensitizing peptide, and SAStide. B) Steady-state luminescence and C) time-resolved luminescence emission plots for SAStide and pSAStide. pSAStide (—) SAStide (---) and Tb (···). Spectra were collected from 15 μM peptide in the presence of 100 μM Tb³⁺ in 10mM HEPES, 100mM NaCl, pH 7.0, λ_ex=266nm, 1000ms collection time, 50 μsec delay time and sensitivity 180. Data represent the average of experiments performed in triplicate.

3.4.2 Physical characterization of SAStide-lanthanide binding and luminescence

Binding studies were performed to determine the stoichiometry and affinity of pSAStide-terbium complexation. The binding stoichiometry of the highest affinity complex was established using the Jobs method of continuous variations. The area under the emission spectrum was used as the metric to quantify the pSAStide-Tb³⁺ binding ratios. pSAStide and terbium individually displayed no detectable luminescence; therefore, any changes in luminescence could then be attributed the formation of the pSAStide-Tb³⁺ complex. The Jobs plot displayed an increase in total area as the mole fraction of terbium increased to 0.5 followed by a linear decrease with further increases in the mole fraction (Fig. 3-3A). These data indicated that the preferred binding stoichiometry of pSAStide-Tb³⁺ binding is 1:1.
Figure 3-3 Characterization of the SAStide/Tb$^{3+}$ physical interaction. A) Jobs Plot of continuous variation. The mole fraction of Tb$^{3+}$ and pSAStide were inversely varied relative to each to maintain a total molar concentration of the two species at 16 µM. Luminescence spectra were collected in 10mM HEPES, $\lambda_{ex} = 266$ nm, 1000ms collection time, 50 µsec delay time and sensitivity 180. Data represent averages ± SEM of experiment performed in triplicate. B) Binding affinity of SAStide-terbium complexation. Spectra were collected using 1 µM pSAStide (●) or SAStide (○) in the presence of varying concentrations of Tb$^{3+}$ in 10mM HEPES, 100mM NaCl, pH 7.5, $\lambda_{ex} = 266$ nm for pSAStide and $\lambda_{ex} = 275$ nm for SAStide, 1000ms collection time, 50 µsec delay time and sensitivity 180. Data represent averages ± SEM of experiment performed in triplicate.

Binding affinities were also determined using Tb$^{3+}$ luminescence as a measure of complexation. Terbium was titrated in the presence of 100 nM pSAStide or SAStide and luminescence emission spectra were collected and integrated. The binding curves displayed a hyperbolic increase in luminescence with increasing terbium concentrations from 0-20 µM (representing a large excess of terbium), with saturation between at 20 µM characteristic of one site binding (Fig. 3-3B). Additional increases in luminescence, ranging up to three fold, were observed with increasing terbium concentrations, which mass spectrometry analysis suggested were likely due to complexes containing multiple terbium ions. However, for the remainder of this work detection of pSAStide was carried out with 6.67 equivalents of terbium relative to peptide, thus the 1:1 binding mode characterized by the initial hyperbolic increase was the most relevant to detection under
conditions used subsequently for assays. The calculated $K_d$ for the 1:1 pSAStide-terbium complex represented by the hyperbolic curve was $1.51 \pm 0.087 \mu M$, which is comparable to the affinities reported for other terbium binding peptides. (15, 20) The unphosphorylated SAStide-terbium complexation displayed significantly weaker binding; the 1:1 complex exhibited a $K_d$ of $7.64 \pm 0.32 \mu M$ (5-fold weaker than for the phosphorylated peptide). These results demonstrate that phosphorylation increased the affinity of SAStide for Tb$^{3+}$. Also, since the greatest fold change in signal for pSAStide-terbium vs. SAStide-terbium was observed for the 1:1 complex, this represented the best ratio to maintain in subsequent kinase assays.

Measurements of the terbium luminescence lifetime were performed to characterize the photophysical properties of pSAStide-terbium and SAStide-terbium complexation. The hydration number (i.e. the number of water ligands ($q$) in the terbium coordination sphere) can be determined via the luminescence lifetime of the complex in H$_2$O vs. D$_2$O, since the terbium excited state is quenched by the $\text{–OH}$ vibrational overtones of H$_2$O but not D$_2$O. Luminescence spectra for the pSAStide:Tb$^{3+}$ and SAStide:Tb$^{3+}$ complexes were collected in various ratios of H$_2$O/D$_2$O. The luminescence lifetimes were fitted to a single exponential decay and were determined to be 2.02 ms and 2.48 ms in H$_2$O and D$_2$O respectively for pSAStide:Tb$^{3+}$. These lifetimes lead to a $q$ value of 0.12 for the phosphopeptide complex, indicating nearly an absence of H$_2$O in the inner coordination sphere of terbium at equilibrium. The unphosphorylated SAStide:Tb$^{3+}$ had a comparably shorter lifetime in H$_2$O at 1.88 ms and a longer lifetime D$_2$O of 2.92 ms resulting in a $q$ value of 0.66. These data suggest the SAStide:Tb$^{3+}$ contain closer to one H$_2$O in the coordination sphere at equilibrium, resulting in more quenching of the terbium
excited state, which manifested as a shorter lifetime. These differences are likely related to the greater luminescence intensity observed for the phosphorylated vs. unphosphorylated peptide, and also to the longer lifetime of the phosphopeptide that enabled gating of the signal for better dynamic range for discriminating the phosphorylated from unphosphorylated species using time-resolved signal collection. (Table 3-1)

Table 3-1 Photophysical properties of the pSAStide-terbium and SAStide-terbium complexes.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>λ&lt;sub&gt;Ex&lt;/sub&gt;</th>
<th>τ&lt;sub&gt;H2O&lt;/sub&gt;</th>
<th>τ&lt;sub&gt;D2O&lt;/sub&gt;</th>
<th>q</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSAStide</td>
<td>266</td>
<td>2.02</td>
<td>2.44</td>
<td>0.12</td>
<td>0.34</td>
</tr>
<tr>
<td>SAStide</td>
<td>275</td>
<td>1.88</td>
<td>2.92</td>
<td>0.66</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

λ<sub>Ex</sub> = maximum excitation wavelength; τ<sub>H2O</sub> = luminescence decay in the presence of H<sub>2</sub>O; τ<sub>D2O</sub> = luminescence decay in the presence of D<sub>2</sub>O; q = number of H<sub>2</sub>O ligands in inner coordination sphere at equilibrium; φ = quantum yield.

The quantum yield for the pSAStide-terbium complex was determined using diffusion-enhanced energy transfer from the complex to fluorescein isothiocyanate (FITC). (30) Luminescence emission spectra were collected in the presence of varying concentrations of FITC with increasing delay times. The time-resolved emission spectra and corresponding lifetime plots display an increase in emission intensity and a simultaneous decrease in the lifetime with increasing concentrations of FITC, and the quantum yield calculated from these data was 0.34.

3.4.3 Quantitative detection of Syk kinase activity using Tb<sup>3+</sup> sensitization

To demonstrate the use of SAStide:Tb<sup>3+</sup> as a biosensor for quantitative detection of Syk activity, a calibration curve was established using mixed ratios of SAStide and pSAStide.
in the presence of the kinase assay components and quenching buffer conditions (MgCl$_2$, BSA, ATP, NaVO$_4$, protease inhibitors, piceatannol, DMSO, urea) at concentrations sufficient to mimic an appropriate background matrix for a kinase assay measurement (Fig. 3-4A). Luminescence emission spectra were collected for the various SAStide/pSAStide ratios and the area under the curves were integrated. Controls showed limited interference from the components of the kinase assay and quenching conditions. The calibration curve demonstrated that the emission spectral area increased linearly and was well correlated with increasing percent phosphorylation (Fig. 3-4B). However, there was an increase in the basal signal in the absence of pSAStide (relative to mixtures of just peptide and simple buffers) that was likely due to complexation of terbium with ATP. The three phosphate groups of ATP can provide an appropriate coordination environment and adenosine provides the appropriate chromophore for excitation ($\lambda_{ex}=259$ nm), giving rise to some background even with time-resolved measurements. Compared to detection of pSAStide vs. SAStide in HEPES buffer alone, this increase in background signal reduced the signal to noise ratio (S/N) by half (Figure 3-2C compared to Figure 3-4B). However, despite this decrease excellent S/N (15.3:1) was still achieved.
Figure 3-4 Calibration and detection of phosphorylation of SASTide. A) pSASTide-Tb³⁺ luminescence emission spectra in the presence of the quenched Syk in vitro kinase assay buffer (2.4 M urea, 40 µM ATP, 4 mM MgCl₂, 0.4 µM Na₃VO₄, leupeptin, aprotinin, 50 ng/µL BSA and 10 mM HEPES pH 7.5). Spectra were collected from 15 µM peptide in the presence of 100 µM Tb³⁺ (representing a ratio of 6.7 Tb³⁺:peptide, which corresponds to the saturated range for 1:1 binding) in 10mM HEPES, 100mM NaCl, pH 7.0, λₑₓ=266nm, 1000ms collection time, 50 µsec delay time and sensitivity 180. B) Emission spectral area calibration curve based on percent phosphorylation in vitro. C) in vitro Syk kinase assay luminescence emission spectra. D) Interpolated percent phosphorylation from Syk in vitro kinase assay. Data indicate averages ± SEM of experiments performed in triplicate.

The limit of detection (LOD) for phosphorylation was 3.8 ± 0.51%, defined as the percentage of pSASTide that gave a signal area corresponding to 3X the standard deviation greater than the baseline for unphosphorylated SASTide in the quenched kinase reaction buffer (the negative control). The limit of quantification (LOQ) for phosphorylation was 7.4 ± 0.52 %, defined as the percentage of pSASTide that gave a signal area 10X the standard deviation greater than the signal in the negative control. The Z’ factor and the signal window (SW) to determine if this sensor
would be appropriate for use in a high throughput screening (HTS) assay. The Z’ factor should be between 0.5 and 1 for an assay to be considered appropriated for HTS, as assays with a Z’ factor in this range exhibit a large dynamic ranges and wide separation of positive and negative results. Assays with a SW greater than 2 are also considered appropriate for HTS assays. Both parameters were calculated from the mean emission and standard deviation of the spectral area from triplicate measurements of the negative control SAStide in the in vitro assay buffer and the positive control pSAStide in the same conditions. The Z’ factor and SW were determined to be 0.82 and 14.63, respectively, indicating that time-resolved terbium luminescence detection of SAStide phosphorylation is an appropriate method for HTS assays.

Detection of Syk activity in vitro was accomplished using Syk-EGFP immunoprecipitated from engineered DT40 chicken B-cells with the kinase reaction buffer and quenching conditions described above. After pre-incubation of the enzyme with the kinase reaction mixture, SAStide substrate was added and aliquots of the reaction were quenched at various time points in urea (to denature the enzyme). Tb$^{3+}$ was added and time-resolved luminescence was measured (Fig. 3-4C). The areas under the emission spectra were calculated, the percent phosphorylation was interpolated from the calibration curve and plotted against time (Fig. 3-4D). These data show that enzymatic phosphorylation of SAStide can be detected using time-resolved Tb$^{3+}$ luminescence. As a control to verify phosphorylation, an additional aliquot was used for detection of phosphorylation using an ELISA-based chemifluorescent assay.

We also examined the potential of this strategy to be used for inhibitor screening. The effect of the Syk inhibitor piceatannol was assayed in a dilution series from 10 nM
to 10 mM. Luminescence emission spectra were collected and integrated (Fig. 3-5A). The areas were normalized to the DMSO control and reported as percent activity. The observed IC$_{50}$ for piceatannol was 178±1.4 µM (Fig. 3-5B), consistent with that found in the literature. (35) The Z’ factor and SW were determined in the context of the dose-response inhibition assay, calculated from the standard deviation and mean from the normalized percent activity from triplicate measurements of the negative control (10 mM piceatannol) and the positive controls (10 nM – 500 µM piceatannol). Over all the positive controls the Z’ factor was greater than 0.5 and the SW was greater than 2 demonstrating that the application of pSAStide:Tb$^{3+}$ maintains its appropriateness as a HTS tool in practice.

Figure 3-5 Detection of dose-dependent Syk inhibition. A) in vitro Syk kinase assay luminescence emission spectra. Spectra were collected from 15 µM peptide in the presence of 100 µM Tb$^{3+}$ in the presence of the quenched kinase reaction buffer (2.4 M urea, 40 µM ATP, 4 mM MgCl$_2$, 0.4 µM Na$_3$VO$_4$, leupeptin, aprotinin, 50 ng/µL BSA and 10 mM HEPES pH 7.5), $\lambda_{ex}=266$nm, 1000ms collection time, 50 µsec delay time and sensitivity 180. B) Dose-response inhibition of Syk by the inhibitor piceatannol. The extent of biosensor phosphorylation was interpolated from the calibration curve and normalized to the uninhibited DMSO control. Data indicate averages ± SEM of experiments performed in triplicate.
3.5 Discussion

In summary, we have demonstrated that time-resolved Tb\(^{3+}\) luminescence measurements substantially increase signal to noise and thus dynamic range for quantitative analysis of peptide phosphorylation. Time-resolved luminescence detection has been employed in FRET-based assays for kinase activity (e.g. the LanthaScreen\(^{®}\) assay from Life Technologies),\(^{(36)}\) however these assays rely on Tb\(^{3+}\) chelation by a macromolecular carrier conjugated to an anti-phosphosite antibody coupled with a fluorescently labeled substrate, and Tb\(^{3+}\) itself is not involved in binding to the phosphorylated product of the kinase reaction—therefore the LanthaScreen\(^{®}\) technique depends on antibody availability and is an indirect, “off-line” measure of substrate phosphorylation. Exploiting the binding and sensitization of Tb\(^{3+}\) directly with a phosphorylated substrate is essentially “label-free,” since neither antibodies nor fluorophore labels are required. In our studies, time-resolved measurements significantly improved the signal to noise of detection compared to steady-state measurements by minimizing background signal. The ability to discriminate between the coordination environments for Tb\(^{3+}\) binding to the unphosphorylated vs. phosphorylated peptides based on luminescence lifetime facilitated the improvement we observed in distinguishing between species. Accordingly, phosphorylation of the Syk substrate peptide SAStide could be applied for rapid, quantitative and sensitive detection of Syk kinase activity and inhibition by small molecule inhibitors with little to no interference from the components of the kinase reaction.

Besides the specific application to Syk kinase described here, this strategy has broad significance for detecting phosphorylation using lanthanide sensitization. Time-
resolved detection should expand the possibilities for other peptide- and protein-based lanthanide sensitization approaches to achieve better dynamic range and sensitivity. This will enable us and others to leverage existing Tb³⁺-sensitizing substrates for other kinases, as well as open a new avenue for development of novel substrates to achieve high-throughput compatibility for other kinase targets important in therapeutic development, that otherwise may not have provided sufficient signal to noise with steady-state measurements.
3.6 References


CHAPTER 4. PIPELINE TO DEVELOP PHOSPHORYLATION-DEPENDENT TERBIUM SENSITIZING TYROSINE KINASE BIOSENSORS

4.1 Abstract

Non-receptor protein tyrosine kinases represent a major class of enzymes essential for cellular homeostasis, and thus are a major focus of current drug discovery efforts. Many of these enzymes display differences in substrate specificity but are structurally similar even between families. This makes the design of substrates that can distinguish between tyrosine kinase families a challenge. Peptide substrates that can enhance lanthanide ion luminescence upon tyrosine phosphorylation enable rapid, sensitive screening of kinase activity. However, the bottleneck caused by the laborious task of identifying and optimizing sequences has slowed the development of efficient and selective substrates for lanthanide-based kinase assays. Streamlined strategies to identify sequences that strike an ideal balance between phosphorylation efficiency, selectivity and lanthanide binding are needed. Here we present a method to circumvent this bottleneck and rapidly identify kinase specific substrates that are capable of detecting tyrosine kinase activity through phosphorylation-dependent terbium (Tb$^{3+}$) sensitization. We exploited positional scanning peptide libraries and brought in existing proteomic data from natural substrates to identify substrate preference features that can be applied to achieve substrate specificity while enabling Tb$^{3+}$ binding via key substitutions at residues that are unimportant for specificity. We applied this strategy to develop biosensors for
several non-receptor tyrosine kinases, including Abl, Jak2 and Src family tyrosine kinases. We demonstrated that these substrates are selective within kinase families and are amenable to high throughput drug screening applications. Overall, this strategy represents a pipeline for developing efficient and specific assays that use lanthanide-based detection for virtually any tyrosine kinase/kinase family.

4.2 Introduction

Protein kinases catalyze the reversible phosphorylation of proteins and play a ubiquitous role in the regulation of signal transduction pathways directing cellular processes including proliferation, survival and adhesion. Phosphorylation of a protein can result in changes in activity, conformation, and stability as well as facilitate protein-protein interactions through phospho-recognition domains. The human genome encodes for more than 500 protein kinases, 32 of which are non-receptor tyrosine kinases (NRTKs). (1) This class of kinases represents the primary mechanism for integrating signaling events initiated at the membrane to affect gene regulation in the nucleus. Deregulation of these kinases can occur through overexpression or constitutive activation, which are facilitated by two mechanisms: 1. Loss of function of tumor suppressors (e.g. tyrosine phosphatases PTEN and SHP-1) or loss of autoinhibitory mechanisms (e.g. intramolecular binding of protein interaction domains), and 2. Gain of function mutations (as seen in Philadelphia chromosome-positive chronic myelogenous leukemia, CML, and the Jak2 V617F mutation in myeloproliferative disorders). (2-4) The association of many NRTKs with a multitude of disease states has led to a large drug discovery effort, resulting in the development of 24 FDA-approved small molecule non-receptor tyrosine kinase inhibitors since 2001. (5) However, despite the preclinical and
clinical success of these particular drugs, small molecule inhibitors have only been approved for 30 kinases from 12 families (5.8% of the kinome and 9.8% of the families). Although most of these are marketed as “specific,” many actually lack selectivity for a single target (which is arguably a factor in their success). Therefore, the development of tools that allow for sensitive and specific detection of the activity of many kinases, including the dissection of multiple targets, is critically important for understanding the full scope of on- and off-target effects in kinase inhibitor biology. This integrated information will contribute fundamentally to the development of new therapeutics and also enable relevant treatment monitoring strategies in the clinic.

Typical strategies for monitoring kinase activity use radioactive ATP, antibodies, or proteomics to detect phosphorylation of native substrates. (6-8) While these methods have successfully generated a wealth of information about kinase activity they each suffer from several disadvantages. The dependency of many kinases on upstream activities makes interpreting phosphorylation of endogenous substrates difficult. The ability of kinases to compensate for each others’ loss can also confound the assignment of endogenous phosphorylation sites as targets for a specific enzyme. Artificial peptide substrates (as opposed to endogenous substrates) offer an attractive strategy for examining kinase activity either in vitro or intact cells due to their diverse chemistries, compatibility with a wide variety of detection platforms, and their ability to directly report the function of a particular enzyme. These have been used with analytical approaches including capillary electrophoresis, voltammetry, mass spectrometry, antibody-based detection (e.g. ELISA), light scattering based methods using SERS and RLS, and fluorescence-based methods e.g. chelation enhanced fluorescence (CHEF),
FRET and fluorescence quenching. (9-19) In particular, CHEF methods that sensitize lanthanide ions such as terbium (Tb$^{3+}$) in a phosphorylation-dependent manner can enable high sensitivity and analytical reproducibility. Previously, we described the application of a kinase specific peptide substrate (SAStide) for the sensitive detection of spleen tyrosine kinase (Syk) activity \textit{in vitro} through phosphorylation-dependent enhanced sensitization of Tb$^{3+}$ luminescence. The peptide biosensor was comprised of kinase recognition motif as well as an acidic motif that is appropriately arranged for Tb$^{3+}$ chelation. (20) The luminescence signal is generated when phosphorylation of the tyrosine residue results in exclusion of water and completion of the Tb$^{3+}$ coordination sphere. Phosphorylation also alters the excitation wavelength of the aromatic side chain and increases the binding affinity for the peptide, resulting in a large increase in luminescence (16-fold in the case of SAStide). However, other than this example of a serendipitous case, most CHEF substrates are designed primarily to achieve optimal metal binding, which often comes at the expense of kinase selectivity and enzyme kinetics. Currently there is no general, streamlined method to identify and develop novel substrates that are simultaneously specific for an individual kinase and strong metal chelators.

To develop such an approach, both elements (specificity and binding) must be taken into account. Kinase specificity is determined by two factors: protein-protein interaction domains and the kinase domain interaction with the substrate sequence. (21, 22) Typically, once a kinase has been recruited to a substrate through a protein-protein interaction (e.g. a local scaffolding protein or a domain in the kinase itself), the kinase domain’s recognition of the substrate sequence determines the efficiency of phosphotransfer. Therefore, while scaffolding and protein-protein interactions are very
important, the kinase substrate recognition factor is also critical for the specificity of signal transduction. Phage display, degenerate peptide libraries, and related techniques have successfully identified substrate specificity profiles for many kinases. (22-25) However, these techniques have not universally yielded motifs that can be directly translated into efficient, selective kinase substrates. Within the sequence space defined by a given motif (which can comprise tens of thousands of permutations), computational prediction of preference is not 100% accurate, so not all sequences that score strongly via prediction algorithms are efficient substrates when tested empirically. Additionally, “positive” sequences of amino acids are selected based on the best residues for catalytic turnover in a given assay, which usually only tests one kinase. Therefore, many motifs look similar across non-receptor tyrosine kinases, and data from these methods seem to imply that there are few selectivity determinants. This suggests that there would be limited ability to discriminate between kinase specificities with an individual peptide sequence.

Amino acid residues that prevent the interaction of a protein domain with a particular sequence (due to e.g. steric clash or charge-based repulsion), termed “non-permissive” residues, have recently been identified as important determinants in SH2 domain specificity. (26) Anecdotal evidence exists in the literature supporting the idea that such a hypothesis could apply to kinase substrate interactions as well. For example, spleen tyrosine kinase (Syk) has been shown to prefer acidic residues both up- and downstream of the phosphorylation site. Substitution analysis of the acidic endogenous Syk substrate BLNK Y178 via point mutations of the GST-tagged substrate peptides demonstrated that single substitutions of basic residues at the -1 or +1 positions were
sufficient to attenuate phosphorylation of the sequence by Syk. (27) Similar results were observed when a peptide microarray containing 1,433 known phosphorylated peptides was screened to determine Abl substrate preference. (28) Peptides containing lysine, arginine, proline, and aspartic acid at the -2, -1, +1 and +3 positions, respectively, reduced Abl’s ability to phosphorylate these sequences. The determinants for non-permissive residues have even been demonstrated through structural homology models for the catalytic domains of Abl and Src tyrosine kinase based on the crystal structure of insulin receptor tyrosine kinase. Site directed mutations of the kinase domains guided by these homology models displayed dramatic changes in kinetic parameters. (29, 30) These observations suggest that incorporating the concept of non-permissive residues into a substrate development model could enable the fine-tuning of motifs for better selectivity.

In this report, we present a pipeline to develop peptide substrates for tyrosine kinases (using the NRTKs as a model system) that are compatible with phosphorylation dependent sensitization of Tb\(^{3+}\). Using a novel high-density microarray screening assay, we analyzed positional scanning peptide libraries against representative members of most of the non-receptor tyrosine kinase families and tested peptides based on the consensus motifs that were generated out of these analyses against a panel of NRTKs. We then refined these motifs by bringing in curated collections of known endogenous substrate sequences and combined the concepts of motif profiling and non-permissive residue identification in an in silico model that enabled the rapid identification of selectivity determinants and the relative “importance” of maintaining certain residues at each position. We used this information and Tb\(^{3+}\) motif alignment as sequence-space-filtering criteria to narrow down the potential substrate library generated from the motif for a
given kinase to a manageable handful of sequences that could be empirically tested and thoroughly characterized. We then applied this pipeline to generate biosensors for Abl, Jak2, and Src-Family tyrosine kinases and demonstrated HTS assays using the Abl substrate against a small molecule library to identify novel Abl inhibitors.

4.3 Methods

**Peptide synthesis and purification.** Peptides were synthesized using a Protein Technologies Prelude Parallel peptide synthesizer on Rink-amide resin. Coupling of standard Fmoc-protected amino acids were achieved with HCTU in the presence of NMM in DMF for two 10 min couplings. Fmoc deprotection was preformed in 20% piperidine in DMF for two 2.5 min cycles. Split-pool libraries were synthesized as a single reaction until the +1 position where the resin was split into equal portions and each amino acid was coupled with 4 equiv of the Fmoc-amino acid/HCTU/NMM for the same length of time as above. Following coupling of the -1 position the resin of all reactions with substitutions at the -1 position were combined. The reaction was continued until all substitutions were synthesized.

**Feature extraction.** Positional Scanning Peptide Library – The positional scanning peptide library was generated and assayed using recombinant enzyme and $\gamma$-P$^{32}$-ATP by Dr. Yang Deng in Prof. Benjamin Turk’s laboratory at Yale University. Data (unpublished) were provided to the Parker laboratory in the form of tables representing the radioactive phosphate signal intensity for the spots on the PSPL arrays. To combine the PSPL data with the endogenous substrate information in the filtering algorithm, peptide phosphorylation signals for each array were quantified based on the median
intensity for each spot. The median intensity values were then background corrected and
signal intensity were then normalized by the following equation:

\[ Z_{i,j} = m \frac{S_{i,j}}{\# S_{j,i}} \]

where \( Z_{i,j} \) stands for the normalized score of amino acid \( j \) at position \( i \) having a signal
score \( S_{i,j} \), and \( m \) stands for the total number of amino acids. \( S_{i,j} \) is the signal score of
amino acid \( j \) at position \( i \) where \( i \) is defined in the summation of all the \( m \) amino acids.

**Positional Probability (from endogenous substrates)** – To characterize a sequence of nine
amino acids, four on either side of the phosphorylation site (-4 to +4) were considered in
the score. These positions were chosen based on the disorder in known substrates, which
increases substantially beyond that distance from the phosphorylation site. We computed
the biological probability matrix, BPM, as follows. It is experimentally known that kinase
\( k \) phosphorylates \( n \) substrates (\( n_1, n_2, \ldots, n_n \)) consisting of nine amino acids, four on each
side of the phosphorylation site. The frequency of each amino acid at each position in the
collection of substrates was computed, \( f_{j,i} \), where \( j \) is amino acid (A, C, …, W, Y) at
position \( i \) (-4, -3, …, 1). Due to the limitation of identified substrates for some kinases,
when \( j = 0 \) for those amino acids the value of \( j = 1/n \), where \( n \) is the number of substrate
sequences for kinase \( k \). The matrix values were computed by comparing the observed
frequency, \( f_{i,j} \), within the substrates to the expected frequency (background frequency),
\( b_{i,j} \), derived from the frequency of each amino acid in each protein containing a substrate
sequence as well as non-phosphorylated interacting proteins (defined using protein-
protein interaction databases (Human Protein Reference Database, String 9.0 and
BioGRID 3.2) and curated by literature analysis. We constructed the probability matrix 20 x 9 for each amino acid and position defined as $s_{ij} = f_{ij} / b_{ij}$.

**Positional Scoring Matrix** – The two individual matrices, BPM and PSPLM, were then multiplied together to form the positional scoring matrix, PSM. The value for each amino acid was then used to identify favorable and unfavorable residues at each position. Values greater than 0.9 were considered favorable or permissive for the kinase, while values less than 0.9 were considered unfavorable or impermissive.

For an nonapeptide of a given amino acid sequence the product of all $s_{ij}$ values yields the raw probability score, $S_R$.

$$S_R = \prod_{i=1}^{8} s_{i,j}$$

The raw score was normalized by probability of any nonapeptide being a substrate for kinase $k$, $P_s$. $P_s$ was determined by the number of kinase substrates collected, $n$, plus the number of significantly peptides from the PSPL compared to the total number tyrosine center nonapeptides seen in substrate and interacting proteins and the 200 peptides from the PSPL for kinase $k$.

$$P_s = \frac{n + x}{n_T + 200}$$

$$S = \frac{S_R}{S_R + \frac{1}{P_s}}$$

Positional selectivity was determined by the ratio of the number of significantly abundant amino acids found at the subsite, $n_{ij}^{\text{sig}}$, to that the expected abundance from a random
distribution, \( n_{i,j}^{\text{sig}} \), multiplied by the ratio of the number of significantly abundant amino acids at the subsite to the total number of residues, \( n_{i,j}^{\text{aa}} \).

\[
S_i = \frac{\sum_j n_{i,j}^{\text{sig}}}{\sum_j n_{i,j}^{\text{aa}}} \times \frac{\sum_j n_{i,j}^{\text{sig}}}{\sum_j n_{i,j}^{\text{aa}}}
\]

An amino acid was defined as being significantly abundant if its frequency was found to be greater than two standard deviations above the mean frequency observed in the total sequence library.

**Performance Evaluation.** To evaluate the predictive performance of the algorithm, receiver operator characteristic (ROC) curves were calculated and plotted based on the specificities (Sp) and sensitivities (Sn) by taking different thresholds.

\[
\text{Specificity (Sp)} = \frac{TN}{TN + FP}
\]

\[
\text{Sensitivity (Sn)} = \frac{TP}{TP + FN}
\]

where TP is the number of true positive predictions (phosphorylation predictions that were also present in the positive data set), TN is the number of true negative predictions (non-phosphorylated predictions that were also present in the negative data set), FN is the number of false negative (non-phosphorylated predictions that were also in the positive data set), and FP is the number of false positives (phosphorylation predictions that were also in the negative data set). Areas under ROC curves (AROC) were also calculated based on the integration and used as the Mann-Whitney U statistic.

Additional characterization was performed by the following:

\[
\text{Accuracy (Ac)} = \frac{(TP + TN)}{(TP + FN + TN + FP)}
\]

\[
\text{Precision (Pr)} = \frac{TP}{TP + FP}
\]
Equal Error Rate (EER) = \((1 – Sn) \times (TP + FN) + ((1 – Sp) \times (1 – (TP + FN))) \times 0.01\)

\[
MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}
\]

**Generation of Kinase Focused Peptide Libraries.** Kinase focused peptide libraries were generated based on the values of the PSM. All \(s_{ij} > 0.9\) were chosen as potential residues at each position. Combinatorial peptide sequences were generated from these residues and scored against each kinase. Those peptides that scored positive for the kinase (or kinase family) of interest and negative for all other kinases (or kinase families) were then selected and added to virtual “focused libraries” for further screening.

**Terbium Binding in silico Screening.** Following the generation of focused putative kinase substrate libraries, sequences were screened for the potential to bind terbium in a phosphorylation-dependent manner. A BLOSUM 62 matrix was used to generate a sequence similarity score between the focused library of potential kinase substrates and the known terbium sensitizing sequence \(\alpha\)-syn Y125 (DPDNEAYEMPSEE). (31, 32)

**In vitro kinase assays.** Recombinant kinases were incubated with the kinase reaction buffer (15 nM kinase, 100 \(\mu\)M ATP, 10mM MgCl\(_2\), 125 ng/\(\mu\)L BSA and 25 mM HEPES pH 7.5, total volume 180 \(\mu\)L) containing 12.5 \(\mu\)M biosensor at 30\(^\circ\)C. Aliquots (20 \(\mu\)L) were taken at designated time points (0.5, 5, 10, 15, 30, 45 and 60 min) and quenched in 6 M urea (20 \(\mu\)L). The quenched samples were then treated with the luminescence buffer (500 \(\mu\)M Tb\(^{3+}\) and 500 mM NaCl, 10 \(\mu\)L) for a total volume of 50 \(\mu\)L (final concentrations of sample components: 2.4 M urea, 40 \(\mu\)M ATP, 4 mM MgCl\(_2\), 50 ng/\(\mu\)L BSA and 10 mM HEPES pH 7.5). Luminescence emission spectra were collected as described above and the area under each spectrum was integrated using GraphPad Prism.
An additional aliquot (2 µL) of the kinase reaction mixture was taken at each time point for validation of phosphorylation using an ELISA-based chemifluorescent assay as previously described. (33) Briefly, each aliquot was quenched with 0.5 M EDTA and incubated in a 96-well Neutravidin coated plate (15 pmol biotin binding capacity per well, Thermo Scientific) in Tris-buffered saline (TBS) containing 0.1% BSA and 0.05% Tween 20 for 1h. Following incubation, each well was washed with the TBS buffer and then incubated with mouse anti-phosphotyrosine monoclonal antibody 4G10 (Millipore, 1:10,000 dilution in TBS buffer) for 1h. Following incubation, each well was washed with TBS buffer and incubated with horseradish peroxidase-conjugated goat anti-mouse immunoglobulin G (IgG) secondary antibody (Abcam) (1:1000 dilution) for 1h. Wells were then washed and treated with Amplex Red reaction buffer (Amplex Red reagent, Invitrogen, 20 mM H₂O₂ and sodium phosphate buffer) for 30 min. Fluorescence was measured using a Synergy4 multiwell plate reader (Biotek) with an excitation wavelength of 532 nm and emission wavelength of 590 nm.

**Dose-response Inhibition Assay.** Kinase (15 nM) was incubated with the kinase reaction buffer described above in the presence of DMSO (vehicle) or varying concentrations of kinase inhibitors (nilotinib, bosutinib, ruxolitinib) at 30°C for 10 min prior to the start of the reaction by adding the peptide substrate. The reaction was started with the addition of biosensor (12.5 µM, total reaction volume 20 µL). Each reaction was quenched after 30 min in 6 M urea (20 µL). The samples were then treated with the luminescence buffer (500 µM Tb³⁺ and 500 mM NaCl, 10 µL) for a total volume of 50 µL.

**Growth inhibition curves.** K562 cells were seeded into 96-well plates at 10,000 cells per well in Iscove’s Modified Dulbecco’s Medium supplemented with 10% fetal bovine
serum and pen/strep. The cells were dosed with the indicated inhibitor at the indicated concentrations (n = 4), and allowed to incubate for 3 days at 37° C. Following incubation, XTT reagent (ATCC) was added according to manufacturers protocol, and allowed to incubate at 37° C for 3 hours. Absorbance at 475 nm was measured on a Biotek Synergy4 plate reader. Values were calculated as percent of vehicle (0.1% DMSO), plotted in Graphpad Prism 6, and IC_{50} values generated by fitting a variable slope (four parameter) curve.

**High-Throughput Screening Calculations.** The Z’ factor was calculated according to Eq. 2.

\[
Z' = \frac{\left( \mu_{pos} - \frac{3\sigma_{pos}}{\sqrt{n}} \right) - \left( \mu_{neg} + \frac{3\sigma_{neg}}{\sqrt{n}} \right)}{\mu_{pos} - \mu_{neg}} \tag{2}
\]

The signal window was calculated according to Eq. 3

\[
SW = \frac{\left( \mu_{pos} - \frac{3\sigma_{pos}}{\sqrt{n}} \right) - \left( \mu_{neg} + \frac{3\sigma_{neg}}{\sqrt{n}} \right)}{\frac{\sigma_{pos}}{\sqrt{n}}} \tag{3}
\]

where n is the number of replicates, \(\mu_{pos}\) and \(\mu_{neg}\) are the average luminescence of the positive (phosphorylated peptide or uninhibited) and negative (unphosphorylated peptide or control inhibitor-treated) controls respectively; \(\sigma_{pos}\) and \(\sigma_{neg}\) are the standard deviation of the positive and negative controls.
4.4 Results

4.4.1 Design and Development of Kinase Specific Substrates For Terbium-Based Detection of Phosphorylation

Identifying optimal peptide substrate sequences for a kinase out of the enormous sequence space of a completely random sequence library (e.g. 25 billion possible combinations for a 9 amino acid peptide with a fixed phosphorylation site residue) has traditionally been a major technical challenge, and has created a bottleneck in the development of new specific substrates. The use of degenerative peptide and one-bead-one-peptide libraries has been successfully applied for the identification of catalytically favorable sequences for individual kinases; however, most of these sequences, even when phosphorylated, are insufficient to sensitize Tb$^{3+}$ luminescence. The birth of global proteomics and advances in proteomic technologies including protein arrays and positional scanning peptide libraries has resulted in an explosion of kinase substrate data over the past two decades, and catalyzed the identification of novel phosphorylation sites and linear motifs for individual kinases. We aimed at taking advantages of these technologies and data to develop a pipeline to identify peptide substrates that would be not only efficiently phosphorylated by a given kinase, but also selective to that kinase/kinase family as well as able to enhance Tb$^{3+}$ luminescence in a phosphorylation-dependent manner.

Inspired by the general design rules of previous CHEF-based sensors for detection of kinase activity (i.e. Sox-Mg$^{2+}$ and EF-hand-Tb$^{3+}$ sensors), we aimed at developing a general approach to design biosensors for tyrosine kinase activity analysis using phosphorylation-dependent enhanced Tb$^{3+}$ luminescence. Sensors were designed to
combine nonreceptor tyrosine kinase substrate specificity with the excitation and chelation elements governing Tb³⁺ luminescence. To achieve this, each kinase biosensor was developed to contain an optimized substrate sequence with an embedded Tb³⁺ coordination motif, similar to that previously identified from the 14-residue fragment of α-synuclein surrounding Y125. (Figure 4-1A) Based on our previous work, we reasoned that this design strategy was plausible due to Tb³⁺ being a hard acid, which gives it preferred coordination with hard bases such as the carboxylic acids in the side chains of glutamic and aspartic acid. (20) These amino acids are also generally preferred amino acids in tyrosine kinase substrate motifs. (34) To further utilize the hard acid character of Tb³⁺ for detection of kinase activity, the incorporation of phosphate into the sequence at tyrosine, catalyzed by the kinase, provides an additional bidentate hard base moiety. The expectation of phosphate addition by the kinase allows for designed control of the number of coordinating water molecules to Tb⁵⁺, otherwise known as the hydration number. The luminescence of Tb³⁺ is susceptible to quenching through nonradiative energy transfer to O-H oscillators; therefore, water molecules coordinated to Tb³⁺ can quench luminescence emission. The biosensors were designed to use phosphorylation of the peptide to decrease the hydration number of Tb³⁺. This led to an increase in the luminescence lifetime as well as increased the binding affinity of the peptide for Tb³⁺, ultimately resulting in enhanced Tb³⁺ luminescence compared to the unphosphorylated form. The final design consideration was excitation of Tb³⁺, which is Laporte forbidden, which results in Tb³⁺ having an extremely low molar extinction coefficient. This requires in a chromophore to absorb the UV-visible radiation and transfer it to Tb³⁺ to facilitate excitation. Sequences containing tyrosine not only provide an appropriate chromophore
for sensitization of Tb$^{3+}$, but also maintain the appropriate distance between the chromophore and the chelated Tb$^{3+}$ ion for Förster resonance energy transfer. An additional advantage of tyrosine for Tb$^{3+}$ excitation and detection of kinase activity is the resulting shift in the excitation wavelength upon phosphorylation from 275 nm to 266 nm. This allows for the selective excitation of the phosphorylated biosensor through the use of a monochromator. Together these phosphorylation-dependent physical changes in the biosensor enable enhanced Tb$^{3+}$ luminescence of the phosphorylated biosensor compared to the unphosphorylated form. (Figure 4-1B)

For each kinase, a focused virtual library of peptide biosensors was designed, optimized and selected 

*in silico* through our bioinformatic pipeline, KINATEST-ID (Kinase Terbium Emission Sensor Identification) (Fig 4-1C). This method starts with the generation of a positional scoring matrix (PSM) (as described below) for a given kinase, which uses biologically validated phosphorylation sites for individual kinases as well as empirically observed effects of amino acids from positional scanning peptide library data (Figure 4-1C Step 1). This matrix represents the relative preference the kinase has for each amino acid at each position within the sequence. Additionally, a positional selectivity matrix was also generated using these data. These matrices were used to guide the generation of a focused library of possible kinase specific peptide substrates, where motif-based substitutions were assessed based on site selectivity. In other words, particular sites that lack selectivity (thus having more flexibility for a given amino acid at that position) but are required for Tb$^{3+}$ binding could be substituted with the appropriate Tb$^{3+}$ binding residue as opposed to a residue suggested by the catalytic preference motif (Figure 4-1C Step 2). Each sequence in the focused library was given a score based on
the PSM for the given kinase, as well as a score for all other kinases included in the analysis using their respective PSMs. (Figure 4-1C Step 2) The focused library was then filtered based on classifying the sequences as predicted “substrates” or “nonsubstrates” for each kinase as well as “specific” or “nonspecific” for the given kinase. All nonsubstrate and nonspecific sequences were then filtered from the library. (Figure 4-1C Step 3) Cutoff scores for classifying the sequences as substrates or nonsubstrates for each kinase were selected based on the algorithm training parameters to give the lowest false discovery rate for the kinase of interest and the highest sensitivity for all off target kinases. This ensured that all remaining sequences in the library would have a maximal likelihood to be substrates for the desired kinase and not for the other kinases. The remaining sequences were compared to the atypical Tb\(^{3+}\) sensitizing peptide derived from the \(\alpha\)-synuclein Y125 center peptide using BLOSUM sequence alignment scoring. (Figure 4-1C Step 4) Sequences with a BLOSUM score below the threshold of 25% similarity were removed from the library. This ultimately yielded a compressed library of potential kinase-specific peptide substrate sequences that were also likely to sensitize Tb\(^{3+}\) luminescence, which were then used in further studies examining kinase specificity and Tb\(^{3+}\) sensitization. (Figure 4-1C Step 5).
Figure 4-1 Design and development of phosphorylation-dependent enhanced Tb\(^{3+}\) luminescence tyrosine kinase peptide biosensors. A) General biosensor design strategy for kinase biosensors capable of phosphorylation induced enhanced Tb\(^{3+}\) luminescence. B) The detection strategy using the phosphorylation-dependent physical changes in the biosensors that result in enhanced Tb\(^{3+}\) luminescence. C) To develop a kinase specific peptide based biosensor, we first obtain all known phosphorylated substrates for a given kinase as the foreground as well as all unphosphorylated tyrosine centered sequences that interact with the kinase as the background. In addition position scanning peptide libraries screens were perform to provide empirical values for each amino acid. (1) A library of sequences were generated in silico based on substrate preferences at each positions along with the Tb\(^{3+}\) binding motif using the positional selectivity score. The library is scored against the kinase of interest as well as all other tyrosine kinases. (2) The library is filtered to remove any nonspecific or nonsubstrate sequences for the kinases based on the PSM scores. (3) The remaining sequences are scored using BLOSUM matrix to access the similarity to the phosphorylation-dependent Tb\(^{3+}\) sequence α-syn Y125. (4) The remaining sequence are validated for kinase specificity and photophysical properties associated with Tb\(^{3+}\) luminescence. (5)
4.4.2 Identification and validation of kinase specific residues

**Endogenous substrate analysis.** For the endogenous substrate information, an individual prediction model was trained from our collection of non-redundant phosphorylated sequences. “Positive” substrates for each kinase were gathered from phosphorylation site repositories including Phosphosite Plus, Phospho.ELM and the Human Protein Reference Database or the literature. “Negative” sequences were derived from the proteins containing substrate sequences and any validated interacting proteins identified from HPRD, BioGRID 3.2, String 9.0 and the literature that contained no currently known substrate sequences for the kinase of interest—in other words, proteins that are known to exist in the proximity of the kinase but which it has not been observed to phosphorylate. We reasoned that this would best represent the tyrosine-containing sequences that had the highest chance of being phosphorylated by the kinase due to the potential for high local concentration, therefore if they are not reported to be phosphorylated, they are more likely to represent true negatives. To characterize a sequence of nine amino acids, four residues on either side of the phosphorylation site were considered. This range of positions was chosen based on the previously characterized disorder in positions beyond -4 and +4 in substrates. (35) Both the identity and the physiochemical properties of the residues at each position were taken into account. Substrate sequences were aligned based on the tyrosine phosphorylation site, and the frequency of each amino acid at a given position was calculated and compared to the frequency of that amino acids at the equivalent position in the entire dataset (negative and positive). To identify property-based features in a substrate motif, the sequence alignment was analyzed to identify redundant properties of the various amino acid side chains (i.e. acidic or basic residues,
hydrophobic or hydrophilic residues, etc.). We computed a biological probability matrix, BPM, based on the frequency of each amino acid in the substrate sequences compared to a background corrected frequency. These BPMs were used to generate a score for a given kinase and a given substrate. The higher the score, the more closely the sequence matched the pattern represented in the substrate preference motif.

**Combined model.** One limitation of using biological data to derive a prediction matrix is that the biological sequence space does not cover all possible sequence combinations, as sequences have evolved to perform multiple functions in a signaling cascade. To enhance the model, score outputs from positional scanning peptide libraries (PSPLs) and the biological probability matrices (BPMs) were multiplied together and taken as features to generate a positional scoring matrix (PSM) for each kinase, using the product of the scores as the classifier for each sequence. Combining the positional scanning peptide library analysis with these endogenous matrices compensated for the information missing from the endogenous substrate dataset. The prediction model was then trained using the collection of endogenous positive “substrate” and negative “non-substrate” sequences through 4-fold cross validation. Comparing to literature motifs as well as several machine learning quality control parameters confirmed the relative quality of the model. An additional feature determined for each kinase was the selectivity of each position (defined as positional selectivity) within the substrate sequence. The positional selectivity was used to describe and determine which positions within the substrate sequence for a given kinase were the most influential in determining kinase substrate selectivity and specificity.
Using the values obtained from the PSM and considering the position selectivity values, the potential effect of each amino acid at each position in the substrate could be predicted. Positional selectivity scores demonstrated that all kinases analyzed displayed strongest substrate selectivity based on the residues present at the -1, +1 and +3 positions relative to the phosphorylation site. Analysis of amino acid variation at these positions revealed that the +3 position displayed the highest selectivity scores, and there was limited variation in the preferred residues at this position as all kinases preferred large hydrophobic residues (I/L/M/F/V) and/or proline. This lack of variation of residues would not allow for manipulation of kinase substrate specificity through amino acid selection at the +3 position. However, analysis of the favorable residues at the -1 and +1 positions demonstrated more diverse preferences at both positions across the kinases studied, providing a promising avenue for selectivity tuning.

Amino acid preference at the -1 position classified nonreceptor tyrosine kinase specificity into two main clusters based on preference for acidic residues (D/E), which was observed as strong for Jak2 and Syk, or large hydrophobic residues (I/L/V), preferred by all other kinases analyzed including Src-family kinases (SFK), Abl, Fer family kinases (FFK), Pyk2, Csk and Btk. The “large hydrophobic” cluster could be further partitioned based on additional weaker preferences. SFK, Abl and FFK demonstrated weak preference for acidic residues, Csk displayed strong preference for glutamine while Pyk2 and Btk exclusively hydrophobic. (Figure 4.2A)

The +1 position demonstrated increased variation in amino acid preference compared to the -1 position, enabling kinases to be classified into four groups. Src-family, Abl-family and Fes-family kinases displayed strong preference for acidic residues
as well as moderate preference for small aliphatic and hydroxyl residues (A/G/S/T). Syk, Btk and Pyk2 displayed strong preference for acidic residues with weaker favorability for large hydrophobic residues (I/L/V). Jak2 exclusively preferred large hydrophobic residues and Csk displayed strong preference for glutamine and weaker preference for large hydrophobic residues. (Figure 4.2B)

**Experimental evaluation of the model using mutations of known substrates.** These differences in amino acid preference at the -1 and +1 positions offer an opportunity to control amino acid substrate specificity through selection of the appropriate residues for that are favorable for one kinase and unfavorable for the others. To examine the ability for our approach to control kinase substrate specificity using these preference definitions, a focused library of peptides based on the canonical endogenous substrates often used to monitor the cellular activity of Abl, Csk, Jak2 and Syk, with single substitutions at the -1 or +1 position, were synthesized using split and pool peptide synthesis. (36-39) The library was then used in an *in vitro* kinase assay with the kinase of interest and analyzed using liquid chromatography-mass spectrometry (LC-MS).

Large hydrophobic residues (I/L/V) were anticipated to be strongly favored at the -1 position for Abl. A pool of peptides based on the CrkL Y207 centered sequence, with substitutions of the alanine at -1 to arginine and glutamic acid, was analyzed to examine the effect of changing the charge of the position. Bioinformatic analysis predicted that the substitutions to either residue would decrease substrate phosphorylation. Both substitutions resulted in a significant decrease in phosphorylation compared to the wild-type substrate. Glutamic acid allowed for lower levels of phosphorylation compared to the wild-type substrate, while substitution with arginine resulted in complete abrogation
of phosphorylation. This suggested that arginine may be a non-permissive residue in that position. To further examine this, the optimal Abl specific substrate Abltide (EAIYAAPF), which is robustly and rapidly phosphorylated under normal conditions, was synthesized with arginine substituted at the -1 position. Substitution with arginine at -1 was able to completely abrogate phosphorylation of Abltide compared to the wild-type substrate, providing strong evidence for non-permissivity of this substitution.

At the +1 position, acidic (D/E) and small hydrophobic residues (A/G) were predicted to be strongly preferred by Abl. For the peptide pool, substitutions at the +1 position examined the effect of changing the size as well as the charge of the side chain through substitutions of alanine to Glu, Leu, Phe, Val and Gly. It was predicted that alanine would be the most preferred followed by glutamic acid and glycine, while leucine, phenylalanine and valine would be unfavorable and potentially non-premissive. In kinase assays on the pool, alanine and glycine substitutions resulted in similar levels of phosphorylation; however, glutamic acid drastically reduced the level of phosphorylation. In agreement with the prediction, Leu, Phe and Val all significantly reduced phosphorylation of the peptide. (Figure 4.2C)

Bioinformatic analysis of Csk displayed strong preference for large hydrophobic residues and glutamine both the -1 and +1 positions. The substitution of the -1 position to isoleucine did not result in the expected increase in the level of phosphorylation, while arginine completely abolished phosphorylation, which was expected. At the +1 position the substitutions of methionine and alanine resulted in a significant improvement in phosphorylation of Src Y530 sequence, which was not expected. Another unexpected result was the low level of phosphorylation associated with the substitution to
phenylalanine, which was expected to be favorable and was previously shown to be the most abundant residue at that position in oriented peptide library screening. (Figure 4-2D) The lack of accuracy in the prediction of Csk substrate determinant is most likely attributed to the bias of substrates generated in the literature, which is fairly sparse for Csk compared to several of the other kinases analyzed in the pipeline. It should also be noted that synergy between positions was observed, as the dual substitution of isoleucine and phenylalanine at the -1 and +1. This result suggests that a relationship exists between the two positions, which is lost in the PSM scoring method.

Previous studies as well as our bioinformatic analysis of Syk substrate specificity demonstrated a strong preference for acidic residues at both the -1 and +1 positions, which was unique amongst all kinases. Substitutions of the BLNK Y178 center sequence examined the effects of charge and hydrophobicity at both positions. It was predicted that substitutions of acidic residues to large hydrophobic amino acids would result in a significant decrease in substrate phosphorylation; this prediction was borne out in the experimental results. The smaller decrease associated with the substitution to phenylalanine (as opposed to leucine or isoleucine) can most likely be attributed to its similarity to tyrosine, which is a favorable residue in those positions. (Figure 4-2E)

Jak2 substrate preference also showed a strong preference for acidic residues at the -1 position, and large hydrophobic residues at the +1 position. The substrate peptide used was derived from the sequence surrounding STAT5 Y694. Substitution from glycine at the -1 position to an acidic aspartic acid residue significantly improved phosphorylation of the substrate (as predicted), while arginine performed comparably to glycine, and glutamine and isoleucine substitutions significantly reduced
phosphorylation. Substitution from valine at the +1 position to phenyalanine resulted in comparable levels of phosphorylation, which was also as expected. Glycine, with its less bulky side chain, resulted in a significant decrease in phosphorylation. Glutamic acid at the +1 position was also tolerated, which was not predicted. (Figure 4.2F)

While the majority of the predictions for individual effects of the substitutions were correct, approximately 25% were incorrect. These results suggest two possibilities: 1) the data obtained and used are insufficient to identify the appropriate value for every amino acid at each position in the substrate for each kinase—the limitations of the data can arise due to limited data set sizes, bias in the data due to the context and focus of studies for the kinase and its substrates, or other mechanisms that regulate specificity aside from substrate sequence recognition; or 2) The other possible explanation is the computational and prediction approach is incorrect, resulting in the wrong interpretation of the data and the incorrect calculation of the PSMs. However, based on the correct prediction of the majority of the substitutions, it is reasonable to believe that the incorrect predictions are a result of the former rather than the latter.
Analysis of amino acid selectivity at the -1 (A) and +1 (B) positions relative to the phosphorylation site. The heat map were generated from the positional scoring matrices (PSMs) using bidirectional heiracrchial clustering. Relative quantification of split pool peptide libraries using a singly-substituted canonical endogenous substrate for each kinase. Abl (C), Csk (D), Syk (E), and Jak2 (F) pools were synthesized using split-pool peptide synthesis and assayed by in vitro kinase assays with recombinant active kinase (3 nM for Abl, Syk and Jak2 and 250 nM for Csk) for 1 hr.
4.4.3 Identification and Characterization of Abl, Jak2 and Src-family kinase biosensors

To demonstrate the application of KINATEST-ID for the identification of kinase-specific peptide biosensors, the algorithm was applied to generate Tb$^{3+}$-sensitizing biosensors predicted to be specific for Abl, Jak2, and Src-family kinases. Initial potential substrate sequence libraries were generated by listing all possible permutations of each kinase’s preference motif (as determined by the PSM). These virtual libraries started with ~43,000, 92,000 and 5,500 sequences for Abl, Jak2 and Src-family kinases, respectively. These libraries were then filtered by PSM scores for each kinase in the analysis to remove predicted nonspecific and nonsubstrate sequences, which drastically reduced the library size by ~99% for each kinase. The Tb$^{3+}$ binding alignment score filter was then applied, which reduced the size of the libraries by a further ~50%, leaving libraries ranging in size from 11-250 sequences. The remaining sequences for each of the kinases contained the identified kinase substrate motifs as well as the !-syn Y125 Tb$^{3+}$ binding motif or slight shifts in that motif.

From these libraries, several sequences were selected to evaluate empirically for each kinase as kinase artificial substrate peptides (KASTides) for Abl (AbASTide), Jak2 (JAStide), and Src-family (SFASStide) kinases. The specificity of these sensors was assessed by screening the peptides against a panel of kinases representing at least one member of each family of nonreceptor tyrosine kinases. The ability of the kinases to phosphorylate a given peptide was determined using an endpoint in vitro kinase assay. Phosphorylation of each peptide was determined quantitatively using a chemifluorescent ELISA-based assay previously developed in our lab. Relative fluorescence units, which corresponded to the amount of phosphorylated peptide present, were measured and
percent phosphorylation was interpolated from a calibration curve. (Figure 4-3B) As predicted by KINATEST-ID all sequences were specific substrates of the intended kinases. AbAStide did display some nonspecific phosphorylation by Csk; however this was only observed at a concentration of Csk 83-fold greater (250 nM) than the concentration of Abl. This result suggested that, while not explicitly measured, the $k_{cat}$ and catalytic efficiency of AbAStide for Csk is significantly lower than those for Abl.

Analysis of Jak2 preference amongst the pool of substrates demonstrated that JAStide-E was the most preferred substrate displaying significantly more phosphorylation compared to the other substrates ($P<0.0001$ for JAStide-A and D and $P<0.001$ for JAStide-B and C). This result is consistent with the substrate substitution analysis preference in Jak2 substrates, for which the -1 position demonstrated the greatest preference for acidic residues (JAStide-E) and reduced favorability for arginine (JAStide-A, B, C) and phenylalanine (JAStide-D). The Src family substrates displayed comparable levels of phosphorylation across all Src family kinases, while maintaining selectivity against all other families. The variation in residue properties between the sequences at the -3 and +2 positions demonstrated that SFKs tolerate substitutions at these positions with little effect on phosphorylation (which is in accordance with the positional selectivity matrix results).

These sequences are the first demonstrated to our knowledge of Jak2 specific substrates (JAStide-A-E), and specific substrates for all Src-family kinases (SFAS tide A and B). The sequence we term SFAS tide-A has been previously identified as an optimal Lyn kinase motif, however it has not previously been analyzed in the context of the entire Src family.
AbASTide, JAStide-E and SFASTide-A were selected for further characterization based on their specificities in the non-receptor tyrosine kinase screening panel as well as their robust phosphorylation by their designated kinases. These biosensors were characterized using steady-state kinetics to model the kinase-substrate interaction and subsequent phosphorylation of the substrates. In vitro kinase reactions were carried out at various concentrations of the substrates. The initial reaction velocities for each kinase-substrate pair were determined and fitted to the Michaelis-Menten equation. The kinetic parameters for each substrate were determined based on the fit, including $K_m$, $V_{max}$, turnover number ($k_{cat}$) and catalytic efficiency ($k_{cat}/K_m$). (Figure 4-3C-E, Table 2) The $K_m$ values for all the peptides were within ~50-200 µM, which is lower than the range typically observed for kinase-substrate pairs for endogenous substrates. (40) Compared to the previously published optimal peptide substrates for these kinases, the $K_m$ values measured for the selected KASTides are ~2-5-fold greater than those reported in the literature. Relatedly, the catalytic efficiencies demonstrated that the AbASTide, JAStide-E, and SFASTide-A sequences are effective substrates for their respective kinases, and these efficiencies are comparable to those reported for the “optimal” kinase substrates. These results demonstrate that KINATEST-ID is capable of identifying sequences with a high likelihood of being selective substrates that have comparable kinetic parameters to the optimal substrates previously identified using traditional, fully empirical methods.
Table 1: Kinase Substrate Sequences and Predicted Scores for Panel of Kinases

<table>
<thead>
<tr>
<th>Kinase</th>
<th>Substrate</th>
<th>Sequence</th>
<th>Abl</th>
<th>Btk</th>
<th>Csk</th>
<th>Jak2</th>
<th>Pyk2</th>
<th>Src</th>
<th>Syk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab1</td>
<td>AbAStide</td>
<td>GGEDEEVAQPEDGKGK(biotin)GG</td>
<td>52.3</td>
<td>0.49</td>
<td>5.59</td>
<td>0.51</td>
<td>27.5</td>
<td>19.4</td>
<td>25</td>
</tr>
<tr>
<td>Ab2</td>
<td>AbAStide-A</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>18.7</td>
<td>0.00</td>
<td>5.36</td>
<td>50.6</td>
<td>0.01</td>
<td>2.03</td>
<td>25</td>
</tr>
<tr>
<td>Ab2</td>
<td>AbAStide-B</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>0.89</td>
<td>1.34</td>
<td>3.50</td>
<td>99.9</td>
<td>0.01</td>
<td>2.36</td>
<td>10.5</td>
</tr>
<tr>
<td>Jak2</td>
<td>JAStide-A</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>2.97</td>
<td>0.12</td>
<td>0.07</td>
<td>39.5</td>
<td>5.04</td>
<td>7.88</td>
<td>13.3</td>
</tr>
<tr>
<td>Jak2</td>
<td>JAStide-B</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>26.1</td>
<td>0.04</td>
<td>0.12</td>
<td>17.1</td>
<td>99.8</td>
<td>8.25</td>
<td>10.8</td>
</tr>
<tr>
<td>Jak2</td>
<td>JAStide-C</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>34.4</td>
<td>0.03</td>
<td>0.03</td>
<td>15.7</td>
<td>99.7</td>
<td>7.88</td>
<td>28</td>
</tr>
<tr>
<td>SFK</td>
<td>SFAStide-A</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>37.4</td>
<td>19.7</td>
<td>65.1</td>
<td>51.3</td>
<td>0.00</td>
<td>98.4</td>
<td>88.2</td>
</tr>
<tr>
<td>SFK</td>
<td>SFAStide-B</td>
<td>GGEDEEVPDQEDGKGK(biotin)GG</td>
<td>34.4</td>
<td>19.7</td>
<td>65.1</td>
<td>51.3</td>
<td>0.00</td>
<td>98.4</td>
<td>88.2</td>
</tr>
</tbody>
</table>

Figure 4-3 Identification, validation, and characterization of kinase specific biosensors using KINATEST-ID. A) The kinase substrate sequences selected for further evaluation and their prediction scores for the panel of kinases used in the assay. B) Screening of kinase substrate against a panel of purified recombinant kinases. Color-coded values represent the mean of experiments performed in triplicate. C-E) Recombinant, active Abl, Jak2 and Lyn (3 nM) were used to carry out the kinase reactions with 100 μM ATP and increasing concentrations of AbAStide, JAStide-E and SFAStide-A. Initial rates of phosphorylation of the kinase specific biosensors (picomoles of phosphorylated product per minute) for AbAStide (C) JAStide-E (D) and SFAStide-A using Lyn (E) were calculated and fitted to the Michaelis-Menten equation. Values represent the mean ± s.e.m of experiments performed in triplicate.
Table 4-1 Kinetic constants for kinase specific peptide biosensors

<table>
<thead>
<tr>
<th>Sensor</th>
<th>$K_m$ (µM)</th>
<th>$V_{max}$ (pmol/min)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$k_{cat}/K_m$</th>
<th>R$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbASStide</td>
<td>99.2 ± 22.0</td>
<td>107.3 ± 7.2</td>
<td>1430 ± 95.5</td>
<td>14.4</td>
<td>0.92</td>
</tr>
<tr>
<td>JASStide-E</td>
<td>186.4 ± 26.9</td>
<td>672.2 ± 37.1</td>
<td>2689 ± 148.4</td>
<td>14.4</td>
<td>0.96</td>
</tr>
<tr>
<td>SFASStide-A</td>
<td>62.4 ± 10.3</td>
<td>269.7 ± 16.1</td>
<td>3596 ± 148.4</td>
<td>57.6</td>
<td>0.96</td>
</tr>
</tbody>
</table>

4.4.4 Tb$^{3+}$ luminescence characterization of KINATEST-ID identified biosensors

The biosensors that displayed appropriate specificity in the screening panel were further evaluated for phosphorylation-dependent enhanced Tb$^{3+}$ luminescence. The ability of the biosensors to have not only appropriate kinase specificity but also to enable large increases in Tb$^{3+}$ luminescence upon phosphorylation is important for robust detection of kinase activity. Phosphorylated and unphosphorylated forms of the peptides were synthesized and Tb$^{3+}$ luminescence emission spectra were collected. Steady-state measurements of the biosensors revealed a modest range of enhancement (~1-2 fold) in Tb$^{3+}$ luminescence upon phosphorylation. (Table 3.2) However, as we have previously observed for a Syk-specific peptide substrate, time-resolved measurements significantly improved the enhancement of Tb$^{3+}$ luminescence to the range of ~5-11 fold (approximately 3-5 fold improvement compared to steady-state measurements). The observed enhancement of Tb$^{3+}$ luminescence could be attributed to the differences in properties of the sensors including binding affinity, hydration number, luminescence lifetime, and excitation wavelength (characterized as described in Table 4.2), which validated the phosphorylation-dependent design of the sensors.

Biosensors were designed such that the incorporation of phosphate into the peptide sequence would result in several physical changes leading to enhanced Tb$^{3+}$...
luminescence. Starting with the excitation wavelength, phosphorylation shifted the excitation wavelength of tyrosine from 275 nm to 266 nm. Relative to excitation at 275 nm, selective excitation of the phosphorylated biosensor at 266 nm through the use of a monochromator significantly improved detection of the phosphorylated biosensors while having little effect on the signal from unphosphorylated forms. The addition of the phosphate group to the biosensors also significantly altered the Tb$^{3+}$ binding affinities for the peptides. The unphosphorylated sequences exhibited binding constants ($K_d$) in the range of 9-80 µM, which were significantly weaker than the range of 1-12 µM observed for the phosphorylated biosensors (Table 4.2). The differences in binding affinities between the two forms of each biosensor resulted in the Tb$^{3+}$ binding equilibrium favoring the phosphorylated peptides over the unphosphorylated peptides. Finally, time-resolved measurements significantly improved the signal to noise for all the phosphorylated biosensors by leveraging the change in luminescence lifetime between the two forms of each peptide for favorable detection of the phosphorylated form. The luminescence lifetimes of the all the biosensors were increased by an amount in the range of 100-200 µsec upon phosphorylation. The time-resolved measurements allowed the shorter-lived signal of the unphosphorylated peptide to decay prior to collection. As we previously described, these changes in lifetime can be attributed to the change in hydration number of the biosensor-Tb$^{3+}$ complexes upon phosphorylation. As incorporated into the design strategy, the phosphorylated tyrosine group most likely provides an additional bidentate chelating moiety, which in turn displaces two water molecules from the coordination sphere of Tb$^{3+}$. As in our previous work, this interpretation is supported by experiments comparing the luminescence lifetimes in water
vs. deuterium hydroxide (which does not quench the Tb$^{3+}$ luminescence). Together these results demonstrate that this general design strategy can be applied to diverse tyrosine kinase substrates, and that these predicted substrates exhibit robust Tb$^{3+}$ luminescence sensitization with photophysical properties consistent with the anticipated detection mechanism.

Table 4-2 Biosensor characterization of physical and photophysical properties

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Steady-State S:N</th>
<th>Time-Resolved S:N</th>
<th>$K_d$ (μM)</th>
<th>$\tau^{-1}_{H2O}$ (ms)</th>
<th>$\tau^{-1}_{D2O}$ (ms)</th>
<th>q</th>
<th>Δq</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFAStide-A</td>
<td>2.1:1</td>
<td>10.6:1</td>
<td>1.35 ± 0.17</td>
<td>1.35</td>
<td>0.51</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.76 ± 0.54</td>
<td>1.83</td>
<td>0.53</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>SFAStide-B</td>
<td>2.1:1</td>
<td>6.2:1</td>
<td>8.16 ± 1.01</td>
<td>1.73</td>
<td>0.71</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>54.12 ± 10.31</td>
<td>2.13</td>
<td>0.67</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>JAStide-B</td>
<td>1.1:1</td>
<td>4.6:1</td>
<td>1.59 ± 0.21</td>
<td>1.26</td>
<td>0.81</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15.31 ± 0.46</td>
<td>1.63</td>
<td>0.71</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>JAStide-E</td>
<td>1.9:1</td>
<td>6.8:1</td>
<td>7.89 ± 6.21</td>
<td>1.61</td>
<td>0.62</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>67.6 ± 8.73</td>
<td>2.10</td>
<td>0.72</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>AbAStide-A</td>
<td>1.3:1</td>
<td>7.3:1</td>
<td>12.24 ± 2.56</td>
<td>1.65</td>
<td>0.61</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.55 ± 3.72</td>
<td>2.21</td>
<td>0.74</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

4.4.5 *In vitro* detection of nonreceptor tyrosine kinase activity

AbAStide, JAStide-E, and SFAStide-A all demonstrated excellent photophysical properties and were further characterize for their performance for *in vitro* detection of kinase activity. Conditions for optimal detection of the biosensors were established in the kinase reaction conditions as well as calibration curves for each biosensor using various ratios of phosphorylated and unphosphorylated forms of the sensors. All sensors displayed linear increases in Tb$^{3+}$ luminescence with increasing percent phosphorylation allowing for quantitative determination of phosphorylation. High-throughput screening parameters were also derived from the calibration curves, including the Z-factor and
signal widow (SW), which reflect the robustness of detection. All sensors displayed appropriate HTS parameters (Z-factor < 0.5 and SW < 2) that support their potential application in HTS screening assays (Table 4.3).

Table 4-3 Characterization of biosensor high-throughput screening parameters.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Limit of Detection</th>
<th>Limit of Quantification</th>
<th>Z’ factor</th>
<th>Signal Window</th>
<th>Signal:Noise</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFAStide-A</td>
<td>7.9</td>
<td>21.8</td>
<td>0.79</td>
<td>13.7</td>
<td>9.1:1</td>
<td>0.95</td>
</tr>
<tr>
<td>JAStide-E</td>
<td>3.1</td>
<td>9.6</td>
<td>0.90</td>
<td>30.8</td>
<td>6.0:1</td>
<td>0.98</td>
</tr>
<tr>
<td>AbAStide-A</td>
<td>4.8</td>
<td>12.1</td>
<td>0.65</td>
<td>5.86</td>
<td>4.2:1</td>
<td>0.95</td>
</tr>
</tbody>
</table>

The sensors were applied for quantitative detection of kinase activity in vitro using recombinant active kinases over a 60-minute time course. To obtain kinase activity measurements aliquots of the reaction were quenched and Tb³⁺ time-resolved luminescence emission spectra were obtained and percent phosphorylation was extrapolated from calibration curves. These data show that enzymatic phosphorylation of the biosensors by the kinases results in increasing Tb³⁺ luminescence over the time course. The percent phosphorylation determined by the Tb³⁺ based read out were in agreement with those obtain using the quantitative ELISA-based read out. (Figure 4.4 A-C)

Next, the biosensors were applied to monitor dose-response inhibition of Abl, Jak2, and Hck kinase activity by their respective inhibitors imatinib, ruxolitinib, and dasatinib. The effects of these inhibitors were assayed in a dilution series from 10 pM to 500 µM. Luminescence emission spectra were collected and normalized to the vehicle (DMSO) control and reported as percent control. The observed IC₅₀ values were 3.88 ± 1.27 nM, 2.89 ± 1.44 nM, and 2.33 ± 1.57 nM for imatinib, ruxolitinib, and dasatinib,
respectively. These values are in agreement with those reported in the literature for each kinase. (41-43) The Z-factor and SW for these assays were assessed to determine if HTS parameters were maintained under the assay conditions. The Z-factors were found to be greater than 0.5 and the SW maintained values greater than 2 for all the sensors in the assays. These results demonstrate that these sensors can be applied for both the detection of kinase activity and inhibition while maintaining robust detection parameters.

![Graphs showing quantitative time-resolved phosphorylation-enhanced Tb³⁺ luminescence detection of nonreceptor tyrosine kinase activity and inhibition. Kinase reaction progress curves for Abl (A) Hck (B) and Jak2 (C). Dose-response inhibition of Abl with imatinib, Hck with dasatinib, and Jak2 ruxalotinb.(D-F)]
4.4.6 Application of AbAStide for High Throughput Screening of Small Molecule Inhibitions for Chronic Myeloid Leukemia (CML)

Chronic myeloid leukemia (CML) is caused by the reciprocal translocation of chromosomes 9 and 22, which results in the formation of the Philadelphia chromosome or the BCR-ABL fusion gene. This gene product results in the Bcr-Abl fusion protein, which results in constitutive Abl kinase activity and drives uncontrolled myeloid proliferation. Treatment of CML has achieved great success in the clinic by inhibition of Bcr-Abl tyrosine kinase activity using the Abl specific inhibitor imatinib. However, resistance to imatinib often arises through the development of a broad range of point mutations within the kinase domain, which impairs inhibitor binding, as well as other off-target mechanisms. The development of the second-generation Abl and dual Abl–Src family kinase inhibitors nilotinib, ponatinib, dasatinib, and bosutinib have allowed many clinically important mutations to be matched to an appropriate and effective inhibitor. However, several mutations still remain insensitive to the currently available inhibitors, including the E255V mutation, and others are continually emerging, indicating that further development of Abl inhibitors that can address the rapidly evolving clinical situation is still warranted. Here we demonstrate the application of AbAStide for the identification of novel Abl inhibitors in a high-throughput screening format.

To scale the assay to a 384-well plate format for reproducible detection of Abl tyrosine kinase activity and inhibition using AbAStide, replicate in vitro kinase assays were performed in the presence (N=96) or absence (N=96) of imatinib. The peptide substrate was incubated with Abl kinase in kinase reaction buffer containing ATP and MgCl₂ for 60 min. Reactions were quenched in Tb³⁺ luminescence buffer (containing urea
and Tb$^{3+}$) and time-resolved luminescence measurements were taken on a Synergy4 plate reader (Biotek). To evaluate assay quality, luminescence of positive and negative control wells (containing the phosphorylated form and unphosphorylated form of the biosensor, respectively, N=96 for each) was also measured. The in vitro kinase assay replicates gave highly reproducible detection of AbAStide phosphorylation and Abl inhibition by imatinib. (Fig. 4.5A) The Z-factor and SW for the kinase reaction replicates were 0.56 and 84 respectively, comparable to those for control well readings. These results demonstrate that in vitro detection of Abl kinase activity and inhibition using AbAStide-enhanced Tb$^{3+}$ luminescence as a read-out is scalable and robust.

The robustness of AbAStide for detecting Abl kinase activity was leveraged in a high-throughput screen using the GSK PKIS library to identify compounds that inhibit Abl kinase activity. Imatinib was used as the positive control and DMSO (vehicle) alone was used as the negative control for in these assays. In addition to these experimental controls, the phosphorylated and unphosphorylated biosensors in the absence of Abl were also used as quality controls. The GSK PKIS library screened consisted of 364 compounds arrayed in 96-well plates as single compounds at 10 mM in DMSO. The library was screened at a constant 1:1000 dilution, with 10 µM final concentration of compound in each well (1% DMSO). Compounds were incubated with the kinase for 30 minutes prior to start of the kinase reaction, which was initiated by the introduction of the biosensor substrate. The kinase reaction was allowed to proceed for one hour before being quenched with the Tb$^{3+}$ luminescence buffer (containing urea and Tb$^{3+}$). The time-resolved Tb$^{3+}$ emission intensity was measured and the “percent inhibition” was determined compared to the biological positive and negative controls (known inhibitor
imatinib and no inhibitor, respectively) (Fig. 4.5B). Primary hits were identified as compounds reducing Tb$^{3+}$ luminescence by greater than 3-fold (the top 5% most potent inhibitors, which were the top 18 compounds). The primary hits were tested in a secondary screen using the same kinase reaction conditions, but employing a chemifluorescent ELISA-based detection instead of Tb$^{3+}$-based detection. The secondary screen confirmed that all of the hits inhibited Abl kinase activity by at least 50% compared to vehicle.

The three most potent inhibitors from the HTS and validation screens were GW693917A, GW711782X and GW513184X, developed to target TIE2/VEGFR2, ALK5 and GSK3β respectively. These were selected for further evaluation of their potency in vitro. In addition to the validated hits, two negative compounds (SB-358518 and GW607049C) were selected as controls for the assay. Potency (IC₅₀) of the compounds was assessed in a dose-response kinase assay with AbAStide. All three hits potently inhibited Abl kinase activity, with IC₅₀ values of 0.52 nM, 1.91 nM, and 0.35 nM for GW693917A, GW711782X and GW513184X respectively. The selected negative compounds did not display inhibition of Abl. (Fig. 4.5C) To determine whether the results of the in vitro inhibition studies translate to a CML model, the compounds were tested in cellular viability assays against the human CML cell line K562. Cellular IC₅₀ for GW693917A was comparable to imatinib, at 81 nM compared to 147 nM. GW711782X and GW513184X were less potent in the cell viability assay, at 20 µM and 3.24 µM respectively. (Fig. 4.5D) Together these results indicate that AbAStide is a biosensor that provides a robust method for the detection of Abl kinase activity, enabling the high-throughput identification and characterization of novel Abl kinase inhibitors.
Figure 4-5 A high-throughput chemical screen using AbAS tide biosensor identifies new inhibitors of Abl tyrosine kinase. A) The AbAS tide in vitro kinase assay shows highly reproducible signal upon imatinib treatment. Green: synthetically phosphorylated peptide (positive control); Black: kinase reaction; Red: kinase reaction + imatinib; Blue: unphosphorylated peptide (negative control). B) Distribution of compound activities identified in a high-throughput screen performed with the GSK PKIS library using AbAS tide-sensitized Tb luminescence. C) Dose-response inhibition of Abl kinase activity by selected compounds from the GSK PKIS. The extent of biosensor phosphorylation was interpolated from an externally generated calibration curve (not shown) and normalized to vehicle (DMSO) control. D) XTT cytotoxicity assay for selected compounds (as in 5C), showing potencies in K562 cells.

4.5 Discussion

Although other methods using synthetic peptide libraries such as oriented peptide libraries and encoded peptide libraries including phage and mRNA displays have been applied to identify determinants of kinase substrate specificity, these methods require
costly, laborious and time consuming experimentation, processing and analysis resulting in a bottleneck in the development of novel kinase specific peptide substrates. Computational approaches offer an alternative to circumvent an empirical bottleneck in the generation and evaluation of kinase specific peptide substrates. They are currently used to facilitate the discovery of novel phosphorylation sites and to identify the kinases responsible for catalyzing the phosphorylation and have yet to be explored for the development of artificial kinase specific sequences. Furthermore, these techniques fail to simultaneously identify sequences that are not only compatible with the various peptide-based detection methods, but will also facilitate robust, rapid, sensitive, specific detection of kinase activity as is seen with phosphorylation-dependent enhanced Tb$^{3+}$ luminescence. To address these challenges we developed an effective strategy, termed KINATEST-ID, which combines the identification of kinase specificity determinants with the prediction of kinase-substrate phosphorylation and peptide:Tb$^{3+}$ complex formation to generate kinase specific biosensors for phosphorylation-dependent enhanced Tb$^{3+}$ luminescence detection. A synergy of approaches are used in the development of the algorithm, which combines positional scanning peptide library data with validated biological phosphorylation sites to predict the individual effects of each amino acid on kinase specificity and BLOSUM substitution matrix sequence alignment to evaluate phosphorylation dependent Tb$^{3+}$ luminescence. Together these approaches identify novel peptide sequences that are kinase specific and produce robust phosphorylation-dependent enhanced Tb$^{3+}$ luminescence for detection of kinase activity.

Several lines of evidence suggest that KINATEST-ID is an appropriate and effective approach for the development of kinase specific peptide biosensors. The quality
of the prediction is demonstrated when compared with other existing phosphorylation prediction methods. Comparing the favorable substrate features extracted from the data using KINATEST-ID to the recently developed M3 (Motif discovery based on Microarry and MS/MS) and NetPhorest there is significant overlap in the identified kinase substrate motifs. This overlap in motifs is an encouraging result to confirm the reliability of the algorithm, since all three methods use different computational approaches and data sets to arrive at similar results. The quality of these features is further demonstrated by the empirical validation using substrate substitution analysis of canonical endogenous kinase substrates. Among the 35 substitutions across the four kinase-substrate pairs examined, the effect of 71.4% (25/35) are correctly predicted based on the PSM values. The weakest prediction is for Csk substrates (35.6%, 3/8), which is attributed to the skewed biological substrate set due to the context of Csk function and the limited study of its kinase activity. This result highlights the importance of obtaining a comprehensive and well curated dataset for the development of these prediction algorithms as the model is only as good as the data. The ability to correctly determine the individual effect of each amino acid on substrate specificity offers the opportunity to generate artificial sequences de novo to serve as kinase specific substrates by selecting appropriate residues for individual kinases. In the past the exhaustive sequence space covering 2.5 billion sequences was screened to identify the most abundant amino acids at each position to give an “optimal” substrate. Using the same initial sequence space and KINATEST-ID for prediction of specificity the number of sequences to be assayed was significantly reduced by seven orders of magnitude. This is a major advantage of using computational approaches as the
sequence space can be reduced to a size that is feasible to screen and monitor each sequence individual rather than identifying the effect as a population.

The sequences presented here are predicted to be specific for individual kinases and displayed specificity in agreement with this prediction. The characterization of these substrates demonstrated that they are all effective reporters of kinase activity. Comparing the performance of the substrates to those reported in the literature reveal that the kinetic parameters were between those found in optimal substrate and those for endogenous substrates. Several Src specific peptide substrates have been developed using one-bead-one-peptide and oriented peptide libraries with $K_m$ values between 20-55 $\mu$M, similar to SFAStide ($K_m = 62 \mu$M). AbAStide however displayed a significant increase in $K_m$ compared to it’s the optimal substrate ($K_m = 4 \mu$M) and comparable to the endogenous substrate CrkL Y207 ($K_m = 134 \mu$M). However, AbAStide was shown to be a better substrate compared to CrkL Y207 due to a higher $V_{max}$ and catalytic efficiency. As for JAStide ($K_m = 186 \mu$M) this is the first report of a specific substrate compared to the commonly used endogenous phosphorylation site STAT5 Y694 ($K_m = 306 \mu$M) the $K_m$ was 2-fold lower. This drop in kinetic parameters is to be expected as catalytic parameters are sacrificed to accommodate Tb$^{3+}$ luminescence detection of phosphorylation. Although the sequences presented here did not perform as well kinetically as previously identified “optimal” kinase specific substrates the design of the sensors can be modified to improvement of kinetic constants. The incorporating of ligands targeting protein-protein interaction domains that helps govern kinase substrate recognition and specificity can increase substrate affinity and the efficiency of
phosphorylation. This has previously been done using the D-domain of Erk as well as the
SH2 and SH3 domains of Abl and Hck.

Aside from the development of the specific peptide substrate another challenge in
the development of peptide biosensors using sulfonamide-oxine (Sox) fluorophores,
environmentally sensitive fluorophores and genetically encoded FRET pairs is the time
required to optimize the phosphorylation responsive signal based on the placement of the
fluorophore or the molecular switch. The ability to predict the response of the sensor and
the detection parameters would increase the throughput of development are alleviate the
optimization time. The design rules applied here for phosphorylation responsive detection
attempt to facilitate predictable response of the biosensors, but do not provide a metric to
measure the potential response. The use of BLOSUM substitution matrix sequence
alignment to identify the similarities between previously identified peptides capable of
phosphorylation-dependent Tb$^{3+}$ sensitization and the sequences generated using
positional scoring matrix. As predicted by the BLOSUM scores and the design rules the
sequences presented here produced robust Tb luminescence and high throughput
screening parameters demonstrating the compatibility of the sensors with HTS screening
assays. Lanthanide-based detection has several desirable properties for high throughput
screening assays including affordability, non-radioactive, provides label-free and
antibody-free detection, it is generic, and it is compatible with time-resolved
measurements. The application the sensors here for HTS presented here offers high
sensitivity and reproducibility as well as quantitative monitoring of nonreceptor tyrosine
kinase activity. The ability to generate novel sensors de novo with similar outstanding
detection parameters is another advantage of applying computational methods for the development of biosensors.

The rapid development of new kinase specific probes presented here will help alleviate the bottleneck in biosensors and facilitate new substrate sequences for novel kinases. Aside from generating kinase specific phosphorylation-dependent Tb\(^{3+}\) luminescence peptide biosensors the application of the algorithm can be a resource for the research community for the development of new artificial peptide substrates for not only tyrosine kinases, but also serine/theronine kinases. These sequences can be applied for detection of kinase activity using previously developed methods including Sox-based CHEF, mass spectrometry, genetically encode FRET-based biosensors and electrochemical assays.
4.6 References


41. Quintas-Cardama, A., Vaddi, K., Liu, P., Manshouri, T., Li, J., Scherle, P. A.,
Caulder, E., Wen, X., Li, Y., Waeltz, P., Rupar, M., Burn, T., Lo, Y., Kelley, J.,
Covington, M., Shepard, S., Rodgers, J. D., Haley, P., Kantarjian, H., Fridman, J.
inhibitor INCB018424: therapeutic implications for the treatment of

42. O'Hare, T., Walters, D. K., Stoffregen, E. P., Jia, T., Manley, P. W., Mestan, J.,
Cowan-Jacob, S. W., Lee, F. Y., Heinrich, M. C., Deininger, M. W., and Druker,
against clinically relevant imatinib-resistant Abl kinase domain mutants, *Cancer
research* 65, 4500-4505.

43. Manley, P. W., Drueckes, P., Fendrich, G., Furet, P., Liebetanz, J., Martiny-
kinase profile and properties of the protein kinase inhibitor nilotinib, *Biochimica
et biophysica acta* 1804, 445-453.
CHAPTER 5. SIMULTANEOUS MUTILPLEXED DETECTION OF LYN AND SYK TYROSINE KINASE ACTIVITIES USING TIME-RESOLVED LANTHANIDE-BASED RESONANCE ENERGY TRANSFER (TR-LRET)

5.1 Abstract

SAStide, a Syk specific peptide biosensor, and SFAStide, a Src-family kinase specific peptide biosensor, combine the biochemical principles of kinase specificity and inorganic properties of terbium coordination for simultaneous rapid, sensitive and quantitative detection of Syk and Lyn activity through phosphorylation dependent terbium sensitization and time-resolved lanthanide-based resonance energy transfer (TR-LRET).

5.2 Introduction

Numerous leukemias and lymphomas have been characterized by the clonal expansion of B-lymphocytes due to the deregulation of the B-cell receptor signaling pathway. (1, 2) Tyrosine kinases Lyn, Syk and Btk are the main signal transducers in this pathway making them ideal therapeutic targets for small molecule inhibitors. (3) Despite the identification of this pathway as the cause of disease there are few therapeutic options targeting the B-cell receptor pathway or these kinases. Currently there is a need for the development of detection platforms that offer sensitive and specific detection of kinase activity that can enhance the depth of information obtained in a single assay. Often, kinase activities are dependent on each other and affect the efficacy of inhibitor drugs.
targeting common pathways. In order to improve upon existing techniques for
discovering and evaluating new inhibitors, multiple events need to be monitored
simultaneously.

Förster resonance energy transfer (FRET) based assays have been developed to
monitor multiple dynamic cellular processes simultaneously in a single assay. (4-8)
However, while useful in some applications, FRET based methods using organic
fluorophores and fluorescent proteins suffer from a number of disadvantages including
small dynamic ranges, the requirement for genetic expression, small Stokes shifts, and
wide emission peaks resulting in spectral bleed through. This ultimately places a limit on
the number of events that can be detected in a single assay. Lanthanides (Ln$^{3+}$) have been
explored as probes in biological assays for the detection of ligand binding, enzyme
activity, and protein-protein interactions due to their unique optical properties. (9-16)
Compared to organic fluorophores and fluorescent proteins, Ln$^{3+}$ have narrow emission
bands, large Stokes shifts and long photoluminescence lifetimes, which offer high
sensitivity and specificity of detection due to the reduced background interference. These
can also allow for multiplexed detection due to the multiple distinct emission bands
associated with each Ln$^{3+}$. (17-19)

Previously, we have demonstrated the development of kinase specific peptide
biosensors capable of detecting kinase activity through phosphorylation-enhanced
terbium (Tb$^{3+}$) luminescence. Here we present the development of a multiplexed
detection platform for the simultaneous monitoring of Lyn and Syk kinase activities
through intramolecular energy transfer from the luminescence donor phosphopeptide-
Tb$^{3+}$ complex to orthogonal fluorophores conjugated to the individual substrates.
5.3 Methods

**Peptide synthesis, fluorophore conjugation and purification.** Peptides were synthesized at 50 μmol scale using a Prelude Parallel peptide synthesizer (Protein Technologies, Inc.) on MBHA-amide resin (Peptides International). Coupling of standard Fmoc (9-fluorenlymethoxy-carbonyl)-protected amino acids (4 equiv) (Peptides International) were achieved with HCTU (3.8 equiv) in the presence of NMM (8 equiv) in dimethylformamide (DMF) for two 10 min couplings. Fmoc deprotection was achieved in 20% piperidine in DMF for two 2.5 min cycles. Side-chain deprotection and peptide cleavage from the resin was performed in 5 ml of a cocktail of trifluoroacetic acid (TFA)/water/ethanedithiol(EDT)/triisopropylsilane (TIS) (94:2.5:2.5:1). Peptides were precipitated and washed three times with cold diethyl ether. The peptides were dissolved in acetonitrile/water/TFA (50:50:0.1), flash frozen in liquid nitrogen, and lyophilized. The peptides were purified by preparative reverse-phase HPLC (Agilent Technologies 1200 Series) using a C18 reverse-phase column. Peptides were characterized by LC-MS and MALDI-TOF analysis.

SASTide was labeled with AlexaFluor-488-maleimide (Invitrogen) and Cy5-maleimide (Lumiprobe) in TCEP (10 mM) and phosphate buffer (100 mM) at pH 6.5. Reaction progress was monitored by MALDI-TOF MS and was found to be complete after 2 h. The labeled peptide was purified using a C18 cartridge (50 mg, Waters) and lyophilized. The labeled peptides were then characterized by LC/MS analysis (see supporting information for characterization data). Peptides were dissolved in distilled water and diluted using Tris buffer (20 mM), pH 9.0. UV-Vis spectroscopy of 5-FAM, AF488, or Cy5 absorbance was determined and the concentration of the peptide solution...
was calculated according to Beer’s Law using the respective fluorophore’s $\lambda_{\text{max}}$ extinction coefficients.

**Luminescence Emission measurements.** Emission spectra were collected on a Biotek Synergy4 plate reader in black 384-well plates (Greiner FluorTrac 200). Spectra were collected from 450-800 nm after excitation at 266 nm (using a monochromator) with a delay time of 50 µsec and collection time of 1 msec. Sensitivity was adjusted to 180.

**Luminescence lifetime measurements.** Luminescence emission intensities were collected on a Biotek Synergy4 plate reader at 23°C in black 384-well plates (Greiner FluorTrac 200). Spectra were collected from the donor peaks (495 nm for AF488 and 5-FAM and 670 nm for Cy5) after excitation at 266 nm with a delay time starting at 100 µsec, increasing in steps of 100 µsec to 5000 µsec, and a collection time of 1 ms.

Samples were prepared with 1 µM pSASTide or SAStide with 10 µM Tb3+ in 100 mM NaCl, 10 mM HEPES pH 7.5 in 100 µL total volume. Emission intensities were normalized to the intensity obtained at 100 µsec delay and fitted to a single exponential decay curve using GraphPad Prism and lifetimes.

**Determination of LRET parameters.** LRET follows the same principles as FRET and can have the same theory applied to calculate the distance between the fluorophore acceptor and the terbium-peptide complex donor pair. The fundamental concept of Förster theory is that resonance energy transfer is proportional to

$$ R = R_0[(1/E)(1)]^{1/6} \quad (1) $$

Where $E$ is the energy transferred, $R_0$, is the Förster distance, and $R$ is the distance between the donor and acceptor. The percentage of energy transfer, $E$, can be determined
from the lifetime measurements of the donor in the absence of the acceptor, $\tau_D$, (peptide-terbium complex (donor) without the conjugated fluorophore (acceptor)) and the donor in the presence of the acceptor, $\tau_{DA}$.

$$E = \frac{1 - \tau_{DA}}{\tau_D} \quad (2)$$

$R_0$, the Förster distance is determined for each acceptor/donor pair as the distance between the acceptor and donor that results in 50% energy transfer by the following equation.

$$R_0 = 0.211(\kappa^2 \eta^{-4} Q_D J) \quad (3)$$

Where $\kappa^2$ is the orientation factor, $Q_D$ is the quantum yield of the donor (which was determined in our previous studies for both peptide-Tb$^{3+}$ complexes), $\eta$ is the the refractive index, which is 1.4 for biological samples in water, and $J$ represents the spectral overlap of the emission spectrum of the donor and the excitation spectrum of the acceptor. $J$ is determined by the following equation

$$J = \frac{\sum [F_D(\lambda)\varepsilon(\lambda)\lambda^2\Delta\lambda]}{\sum [F_D(\lambda)\Delta\lambda]} \quad (4)$$

where $F_D(\lambda)$ is the corrected fluorescence intensity of the donor, $\varepsilon$ is the extinction coefficient of the acceptor.

**In vitro kinase assay.** Purified Lyn was obtained from a commercial source (Millipore). His-tagged Syk was stably transfected into and purified from HEK293 cells. Cells were lysed and Syk was purified using Ni$^{2+}$ magnetic beads (Promega). Syk protein concentration was determined using quantitative BCA assay. Each individual kinase
enzyme (15 nM) or a combination of both was pre-incubated with the kinase reaction buffer (100 μM ATP, 10 mM MgCl₂, leupeptide, aprotinin, and HEPES pH 7.5) for 10 time, followed by initiating the kinase reaction by adding either SAStide-Cy5 (12.5 μM ), SFAS tide-5-FAM (2.5 μM ) or both at 37 °C. Aliquots were removed at designated time points and quenched in 6 M Urea pH 7.5. The quenched samples were then treated with 10 μM Tb³⁺ bring the sample to a volume of 100 μL.

5.4 Results

Time-resolved lanthanide-based energy transfer (TR-LRET) detection of Lyn and Syk kinase activities was achieved by employing the phosphopeptide-Tb³⁺ complexes as the energy donors and 5-carboxyfluorescein (5-FAM) and cyanine 5 (Cy5), respectively, as the energy acceptors (Figure 5.1A.) Phosphorylation of the peptide substrates results in several physical changes in the peptide that ultimately lead to detection of kinase activity, including enhancing the Tb³⁺ binding affinity, reducing the Tb³⁺ chelate hydration number, increasing the Tb³⁺ luminescence lifetime, and shifting the excitation wavelength of tyrosine. (20) Sensitization through tyrosine results in excitation of the pSFAStide-Tb³⁺ complex and triggers energy transfer to 5-FAM and results in emission from 5-FAM at its characteristic wavelength (~520nm) (Figure 5.1B). 5-FAM was selected as the energy acceptor to couple with the pSFAStide-Tb³⁺ complex because its broad excitation peak at 495 nm matches well with the ⁵D₄ → ⁷F₆ emission band of Tb³⁺ centered at 495 nm. Similarly, detection of Syk kinase activity was achieved with SAStide-Cy5 based on the overlap of the excitation band with the ⁵D₄ → ⁷F₄ and ⁵D₄ → ⁷F₃ emission bands of Tb³⁺ centered at 595 nm and 620 nm (Figure 5.1C).
Figure 5-1 Multiplexed detection using time-resolved lanthanide-based resonance energy transfer (TR-LRET) and fluorophore conjugated peptide biosensors. (A) Emission spectrum of phosphopeptide-Tb\(^{3+}\) complex (black), excitation (dashed lines) and emission (solid lines) spectra of the two acceptor fluorophores 5-FAM (green) and Cy5 (red). Schematic illustrating TR-LRET detection of Lyn (B) and Syk (C) tyrosine kinase activities.

Phosphorylated and unphosphorylated forms of SAStide-Cy5 and SFAStide-5-FAM were synthesized as controls. Excitation of each peptide biosensor in the presence of Tb\(^{3+}\) resulted in the four characteristic luminescent emission peaks from Tb\(^{3+}\) as well as the fluorescence emission peak from the conjugated fluorophore label (5-FAM at 520 nm and Cy5 at 670 nm). (Figure 5.2A,B) Both sensors displayed increased Tb\(^{3+}\) luminescence and fluorescence of their respective organic dyes for the phosphorylated peptides relative to the unphosphorylated peptides. Control experiments in the presence and absence of Tb\(^{3+}\) showed that the sensitization and resulting emission spectra were
Tb$^{3+}$- and LRET-dependent rather than arising from direct excitation of the fluorophores for both biosensors. Comparing the differences in Cy5 signals between the phosphorylated and unphosphorylated SAStide displayed a 25-fold increase in Cy5 signal for the phosphorylated peptide. (Figure 5.2A) Comparing luminescence emission spectra of SFAS tide-5-FAM to pSFAS tide-5-FAM, a 3.9-fold increase in the 5-FAM emission intensity was observed. (Figure 5.2B) These changes in the intensity of the fluorophores upon phosphorylation of their respective peptides provide sensor-specific spectral features that can be monitored to determine phosphorylation of the sensors and consequently kinase activity.

In order to achieve multiplex detection in the same sample, the reaction and detection conditions must be optimized to have limited cross-interference between sensors. The cross-interference could be evaluated by detecting the presence of the fluorophore signal from an unphosphorylated sensor in the presence of the other phosphorylated biosensor. To accomplish this, the concentrations of the biosensors and Tb$^{3+}$, as well as the delay time, were varied and and TR-LRET spectra collected. Quantification of fluorescence signal was accomplished by Gaussian fitting of the fluorophore emission peaks and integrating the resulting curves for each peak. The conditions were initially optimized using the phosphorylated SAStide sensor (pSAStide-Cy5) with the unphosphorylated SFAS tide-5-FAM peptide. Reducing the concentration of SFAS tide, increasing the delay time, and varying the concentration of the Tb$^{3+}$ successfully mitigated any interference from the 5-FAM signal caused by intermolecular LRET. (Figure 5.2C) Using the same conditions, the cross-interference from SAStide-Cy5 was examined in the presence of pSFAS tide-5-FAM. The TR-LRET spectra
displayed minimal signal from Cy5, while displaying significantly stronger 5-FAM signal. (Figure 5.2D)

Figure 5-2 Time-Resolved Lanthanide-based Resonance Energy Transfer (TR-LRET) detection of phosphorylation-dependent signals and cross-interference. (A) Time-resolved luminescence emission spectra for SAStide-Cy5 and pSAStide-Cy5 and (B) SFASStide-5-FAM and pSFASStide-5-FAM. Spectra were collected from 15 μM peptide in the presence of 100 μM Tb³⁺ in 10 mM HEPES, 100 mM NaCl, pH 7.5, λₑₓ = 266 nm, in 50 μL total volume 1 ms collection time, 50 μs delay time, and sensitivity 180. Data represent the average of experiments performed in triplicate. (C) pSAStide cross-interference with SFASStide signal and (D) pSFASStide cross-interference with SAStide signal. Spectra were collected from 0.5 μM SFASStide-5-FAM and 2.5 μM SAStide-Cy5 in the presence of 10 μM Tb³⁺ in 10 mM HEPES, 100 mM NaCl, pH 7.5, Urea, ATP, BSA, MgCl₂, λₑₓ = 266 nm, in 100 μL total volume 1 ms collection time, 50 μs delay time, and sensitivity 180. Data represent the average of experiments performed in triplicate.
The distance between the Tb$^{3+}$ ion and the fluorophore is a critical parameter for energy transfer, and is directly proportional to the fluorescence signal displayed in the emission spectra. By obtaining the luminescence lifetimes of the biosensors in their fluorophore conjugated and unconjugated forms, the energy transfer and LRET parameters for each sensor were characterized. SAStide was also conjugated with AlexaFluor 488 (AF488) as a control for the measurements to demonstrate the agreement in LRET parameters between using different fluorophores. TR-LRET measurements demonstrated efficient energy transfer (in the range of 89-93%) from Tb$^{3+}$ to the various fluorophores. The radius representing the distance between Tb$^{3+}$ and the fluorophores, R, and the Förster radius, R$_0$, were determined from these measurements as well. The R and R$_0$ values ranged from 50-55 Å and 35-40 Å respectively. As indicated by the efficient energy transfer, these parameters are within the optimal range for TR-LRET measurements. (21)

Figure 5-3 Luminescence lifetime measurement of phosphopeptide-Tb$^{3+}$ complexes in the conjugate and unconjugated forms for (A) SAStide-Alexafluor488 (B) SAStide-Cy5 (C) SFAStide-5-FAM.

To demonstrate the utility of these biosensors for quantitative detection of Lyn and Syk tyrosine kinase activities, a calibration curve was established for each sensor in
the presence of the unphosphorylated form of the other biosensor and the kinase reaction buffer (to best mimic the conditions of a multiplexed kinase reaction). Quantification of phosphorylation was calibrated by integrating the signal centered at 520 nm for 5-FAM and 670 nm for Cy5. The high signal to noise ratio observed in the initial control experiments was maintained in the presence of the reaction buffer with 7.6:1 for Cy5 and 5.8:1 for 5-FAM. Z'-factor and signal window (SW) values were shown to be appropriate for HTS, with Z-factor values of 0.72 and 0.78, and SW of 13.27 and 12.65, for SAStide-Cy5 and SFASStide-5-FAM, respectively.

Figure 5-4 TR-LRET quantitative calibration of biosensor phosphorylation. (A) pSAStide-Tb³⁺ emission spectra with increasing proportions of phosphorylated biosensor compared to unphosphorylated. (B) Cy5 emission spectral area calibration curve based on spectra from (A). (C) pSFASStide-Tb³⁺ emission spectra at increasing proportions of phosphorylated biosensor compared to unphosphorylated. (D) 5-FAM emission spectral area calibration curve based on (C).
Detection of Syk and Lyn activities \textit{in vitro} was accomplished and quantified using the purified kinases with kinase reaction buffer and detection conditions described above. After pre-incubation of the kinases with the reaction buffer (for 10 min) the kinase reaction was initiated by the addition of the biosensor(s). Aliquots were removed from the reaction and quenched with urea, treated with Tb$^{3+}$ and brought to a volume of 100 µL. TR-LRET emission spectra for each respective biosensor displayed an increase in the conjugated dye’s fluorescence signal (with minimal bleed through or background interference from the other fluorophore) over the time course of the reaction in the presence of only one of the kinases (Figure 5.5 A,B,C,D). These results confirmed the specificity of each biosensor for its individual kinase. To demonstrate multiplex detection, both biosensors were applied with both kinases in a single reaction. A simultaneous increase in fluorescence of both fluorophores was seen over the time course, indicating an increase in phosphorylation of both peptides. The negative levels of phosphorylation observed from the nonsubstrate biosensor indicate that the cross interference signal was reduced compared to the calibration curve. This reduction in signal could be the result of photobleaching of the fluorophores over the time course of the assay and the sample handling.
Figure 5-5 Simultaneous multiplexed \textit{in vitro} detection of Syk and Lyn kinase activities. (A) \textit{In vitro} Lyn assay luminescence emission spectra in the presence of both SFAStide-5-FAM and SAStide-Cy5. (C) \textit{In vitro} Syk assay luminescence emission spectra in the presence of both SFAStide-5-FAM and SAStide-Cy5. (E) \textit{In vitro} Lyn and Syk assay luminescence emission spectra the presence of both SFAStide-5-FAM and SAStide-Cy5. (B, D, F) Quantification of SFAStide-5-FAM phosphorylation and SAStide-Cy5 phosphorylation interpolated for each assay from the calibration curves in Fig. 5.4.
5.5 Discussion

In summary, we have presented the development of a multiplexed platform for detection of kinase activity that leverages the overlap of the multiple distinct emission bands of Tb$^{3+}$ with orthogonal fluorescently labeled peptide substrates that are capable of phosphorylation-enhanced Tb$^{3+}$ luminescence. Previously, multiplexed kinase activity detection has been accomplished using the environmentally sensitive fluorophores oxazine and cascade yellow conjugated to Lyn and Abl specific substrates, where phosphorylation of the substrates results in enhanced fluorescence. LRET-based phosphospecific antibodies conjugated with lanthanide chelates (LanthaScreen) and fluorescently labeled peptide substrates have also been applied to simultaneously monitor two kinases in a single sample. However, in the case of environmentally sensitive fluorophores, they suffer from the disadvantages of the having a small dynamic range and background fluorescence, while LanthaScreen is limited by the requirement of phosphospecific antibodies to facilitate detection of phosphorylation. The approach presented here allows for direct measurements of kinase activity through the incorporation of phosphate to the biosensors—circumventing the need for antibody labels.
5.6 References


VITA
VITA

Andrew M. Lipchik
Department of Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University

Education
B.S., Chemistry, 2008, Xavier University, Cincinnati, Ohio
Ph.D., Medicinal Chemistry and Molecular Pharmacology, 2013, Purdue University, West Lafayette, Indiana

Research Experience
Graduate Research Assistant, Purdue University January 2009-Present
Department of Medicinal Chemistry and Molecular Pharmacology
Developed bioinformatic tools for the development of kinase specific peptide biosensors for detection of kinase activity using lanthanide sensitization. Conducted research on the detection of intercellular Abl tyrosine kinase activity using surface enhanced Raman spectroscopy (SERS) and silver nanoparticle biosensors.

Undergraduate Research Assistant, Xavier University January 2006 – May 2008
Department of Chemistry
Conducted research on the development of nickel(II) catalysts to activate oxygen as an oxidizing agents. Designed and synthesized electron-donating ligands to incorporate into nickel(II) complexes. Characterized complexes using spectroscopic techniques.

Teaching Experience
Teaching Assistant, Purdue University August 2008-May 2009
Department of Medicinal Chemistry and Molecular Pharmacology January 2011-May 2011
Organic Chemistry Laboratory

Teaching Assistant, Purdue University August 2010-December 2010
Department of Medicinal Chemistry and Molecular Pharmacology
Integrated Laboratory
Conference Presentations

Oral


**Lipchik, A.M.** “Development of a Syk Peptide Biosensor.” Cell Identity and Signaling (CIS) Seminar Series, Purdue University, West Lafayette, IN, USA. (2012).

Posters


Patents


Awards

Cancer Prevention Interdisciplinary Program Fellowship August 2011-August 2012
Dollens Life Sciences Scholarship August 2009-May 2010

Professional Organizations

American Chemical Society (ACS) 2008-Present
American Association for the Advancement of Science (AAAS) 2010-Present
US Human Proteomics Organization (US HUPO) 2012-Present
American Association for Cancer Research (AACR) 2012-Present


