NORTHWESTERN UNIVERSITY

# Using SAMDI-MS and Peptide Arrays to Study Protein Tyrosine Phosphatases and Discover New Cellular Regulatory Mechanisms

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemistry

By

Che-Fan Huang

EVANSTON, ILLINOIS

June 2021

# Abstract

The opposing activities of phosphatases and kinases determine the phosphorylation status of proteins, yet kinases have received disproportionate attention in studies of cellular processes, with the roles of phosphatases remaining less understood. This dissertation describes the use of self-assembled monolayer laser desorption/ionization mass spectrometry (SAMDI-MS) together with peptide arrays to directly assay phosphatase activity, profile their substrate specificity and reveal novel cellar regulatory mechanisms protein tyrosine phosphatases (PTPs) are critically involved in.

SHP2 is the first identified oncogenic phosphatase. In the first part of this work, we applied high-throughput SAMDI-MS assays for SHP2 drug discovery and reported an FDA-approved compound, adapalene, as a potent SHP2 inhibitor. We identified that the adamantyl functional group is crucial for its inhibition. SHP2 disease mutants were profiled with phosphotyrosine-containing peptide arrays to study the alteration of substrate specificity. We found that some mutants particularly favored aromatic residues at the -1 position adjacent to a phosphotyrosine.

Extending the study to the whole PTP family in human proteome, twenty-two tyrosine phosphatases were characterized with the arrays to give a profile of their specificities. An analysis of the data revealed that certain residues in the substrates had a conserved effect on activity for all enzymes tested, including the general rule that inclusion of a basic lysine or arginine residue on either side of the phosphotyrosine decreased activity. This insight also provides a new perspective on the role of an R1152Q mutant in the insulin receptor, which is

known to exhibit a lower phosphorylation level, and which our work suggests may be due to an increased activity towards phosphatase enzymes.

We then identified more than 6,000 cancer mutations involving basic residues adjacent to known phosphotyrosine sites through a database search. Using two β-catenin mutants associated with cancer (T653R/K) and a mouse model for intellectual disability (T653K), we showed that T653-basic mutant β-catenins are less efficiently dephosphorylated by SHP1 phosphatase, leading to sustained Y654 phosphorylation and elevated downstream Wnt signal. This example rationalized how basic mutations proximal to phosphotyrosines can restrict counter-regulation by phosphatases, providing new mechanismistic and treatment insights for 6,000+ potentially relevant cancer mutations.

Lastly, we showed that the same principle can be applied for a phosphorylation/chare-altering modification crosstalk via PTP. Using citrullination for example, we found the modification neutralized the basic arginine residue and increased the peptide dephosphorylation $k_{cat}/K_M$ by 2.3-fold. This novel regulatory mechanism is generalizable because all PTPs lack activity towards substrates that have a basic residue proximal to the phosphotyrosine.

This dissertation demonstrates the use of SAMDI-MS to provide a rapid and quantitative assay of phosphatase enzymes will be important to gaining a more complete understanding of the biochemistry and biology of this important enzyme class.

# Acknowledgements

First and foremost, I would like to express my deepest gratitude and appreciation to my advisor, Professor Milan Mrksich, who has always been an inspirational scientist, a dedicated educator, and a supportive mentor throughout my PhD career. He gave me the freedom to explore at the intersection of chemistry and biology and encouraged me to become a critical thinker and a comprehensive scientist. I am extremely grateful for his guidance.

Next, I would like to thank Professor Cara Gottardi, who started as my research collaborator but went above and beyond to accommodate me working in her lab for over a year and advised me closely as a secondary mentor. I am grateful for the opportunity to learn advanced cell biology under her training and for all her support and effort on my projects.

I would also like to thank Professor Rick Silverman, Professor Neil Kelleher, and Professor Karl Scheidt for serving on my committees providing valuable discussion and suggestions.

Much of my work would not be made possible without collaborations. I am extremely grateful to work with Dr. Elizabeth Eklund, Dr. Marco Tartaglia, Dr. Simone Martinelli and Dr. Lorenzo Stella on the SHP2 project; Dr. Jennifer Grant on the citrullination crosstalk; Dr. Rama Mishra on molecular modeling; Dr. Lihua Zou on cancer database search. I am very appreciative of the mentorship I received from Dr. Alex Yemelyanov and Annette Flozak during my time working in the Gottardi lab. They both helped me with all the cell biology experiments including creating a CRISPR knockout cell line. I feel tremendously lucky to learn from all these experts in their respective fields.

I am very thankful to the whole Mrksich group for being extremely supportive throughout the past few years. I thank Dr. Hsin-Yu Kuo, Dr. Liang Lin, Dr. Lindsey Szymczak and Dr. Sarah Wood for their mentorship when I first joined the group; Dr. Justin Modica, Dr. Eric Berns, Dr. Kevin Metcalf, Dr. Patrick O'Kane, Dr. Alexei Ten, Dr. Jose-Marc Techner, Dr. Maria Cabezas and Dr. Pradeep Bugga for their advice throughout my time in the group. I want to thank Dr. Jennifer Grant and Blaise Kimmel for taking great care of the MALDI instrument which made all the challenging phosphopeptide works possible. I thank Elamar Hakim Moully and Sarah Anderson for taking up a lot of the lab responsibilities including safety and making the lab runs smoothly. I would also like to thank the administrators in the group: Yael Mayer, Rebecca Willingham, Brianna Bullock and Erin Wallace. They assisted me with everything from ordering, scheduling to submitting papers and shared every sweet and bitter moment together.

I want to thank my classmates who started their PhD with me either in the Chemistry program or International Summer Institute. Thank you for being supportive friends and colleagues, Dr. Vivian Tang, Dr. Natalia Powers-Riggs, Yuan Liu, Kevin Kang, Dr. Jennifer Park, Yintai Ma. I would also like to extend my appreciation to all the International Summer Institute faculty and staff for their input into the international student community.

I am grateful to have the strongest and most supportive friendships with the alums of National Taiwan University Chemistry Department. A lot of us shared the experience of pursuing postgraduate degrees in foreign countries and therefore built strong connections despite being in different places. I thank Dr. Timothy Wang and Yin-Peng Chen for their constant help from my day one in Chicago. I would not have been able to establish myself in a new city so smoothly without your kind hospitality. I thank Dr. Anna Chen for always providing incredible

assistance from my PhD application to job interviews and thesis proofreading. And together with her husband Jimmy Liu, they are the most amazing friends one could ever have. I also want to thank Dr. Hua-Chia Tai, Min-Hung Tsai, Chien-I Yang, Tsung-Hung Yao, Dr. Louis Li, Ming-Uei Hung, Hao-Li Huang, Dr. Yi-Chung Dzeng, Jonathan Lee and Dr. Jason Tsai for their incredible remote support and visits throughout my time at Northwestern.

Life without music is unimaginable for me. I am grateful to be a part of the Lakeside Pride Music Ensemble and Northwestern Medical Orchestra and I thank them for giving me the opportunity to nurture my soul through music. I would like to thank my brass quintet members, Brandon Strawn, Will Biby, Travis Cuckler and Jes Klass for sharing very rewarding moments with me on and off the stage.

I want to thank Dr. Jimmy Chang for a lot of valuable discussions and Chicago living perks. I thank Sean Campbell for his support during and beyond my time working at the Pulmonary Department. I thank Danny Pedroza for being the best wingman. I thank my high school friends Dr. Wei-An Lin and Dr. Alex Jen for their timeless friendship. I thank Andy Chester, Troy Steiner and Dr. Thomas Koberda for inspiring me to become the best person. I thank Dr. Jod Taywaditep and Dr. MaryBeth Napier for their companies during difficult moments. And I thank Joey Weiss for his amazing friendship.

Finally, I would like to thank my family, my parents and my siblings. Although I have missed numerous family events for being thousands of miles away, I appreciate the constant messages and pictures, which reminded me that I am still a part of the family and will always have your strongest support no matter where my career takes me to.

# List of Abbreviations

| | |
|---|---|
| SAMDI | Self-assembled Monolayer Laser Desorption/Ionization |
| MS | Mass Spectrometry |
| PTM | Post-translational Modification |
| Pi | Phosphate Ion |
| PK | Protein Kinase |
| PP | Protein Phosphatase |
| PTK | Protein Tyrosine Kinase |
| PTP | Protein Tyrosine Phosphatase |
| ATP | Adenosine Triphosphate |
| RPTP | Receptor-like Protein Tyrosine Phosphatase |
| DSP | Dual Specificity Phosphatase |
| pS | Phosphoserine |
| pT | Phosphothreonine |
| pY | Phosphotyrosine |
| RTK | Receptor Tyrosine Kinase |
| MW | Molecular Weight |
| MALDI | Matrix-assisted Laser Desorption/Ionization |
| TOF | Time-of-flight |
| SAM | Self-assembled Monolayer |
| m/z | Mass-to-charge Ratio |
| SPPS | Solid Phase Peptide Synthesis |

| | |
|---|---|
| HDAC | Histone Deacetylase |
| KDAC | Lysine Deacetylase |
| ELISA | Enzyme-linked Immunosorbent Assay |
| JMML | Juvenile Myelomonocytic Leukemia |
| PNPP | *para*-Nitrophenyl Phosphate |
| DIFMUP | 6,8-Difluoro-4-methylumbelliferyl Phosphate |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| WT | Wildtype |
| Mut | Mutant |
| AUC | Area Under the Curve |
| THAP | 2,4,6-Trihydroxyacetophenone |
| $IC_{50}$ | Half Maximal Inhibitory Concentration |
| a.a. | Amino Acid |
| IFD | Induced Fit Docking |
| Kd | Dissociation Constant |
| PCR | Polymerase Chain Reaction |
| DMF | Dimethylformamide |
| NMM | N-methylmorpholine |
| TFA | Trifluoroacetic Acid |
| TCEP | Tris(2-carboxyethyl)phosphine Hydrochloride |
| DMSO | Dimethyl Sulfoxide |
| AML | Acute Myeloid Leukemia |

| | |
|---|---|
| ALL | Acute Lymphocytic Leukemia |
| NS | Noonan Syndrome |
| LS | LEOPARD Syndrome |
| ELISA | Enzyme-linked Immunosorbent Assay |
| GST | Glutathione S-transferase |
| β-cat | β-catenin |
| *Bfc* | Batface |
| MBP | Myelin Basic Protein |
| CNS | Central Nervous System |

*All amino acids are abbreviated using the standard 3-letter or 1-letter codes.

# Table of Contents

**Abstract** ..... **3**

**Acknowledgements** ..... **5**

**List of Abbreviations** ..... **8**

**Table of Contents** ..... **11**

**List of Figures** ..... **16**

**List of Tables** ..... **18**

**Chapter 1. Introduction to Protein Tyrosine Phosphatases, SAMDI-MS, and Peptide Arrays** ..... **19**

1.1. Protein Tyrosine Phosphatases ..... 20

1.1.1. Protein Phosphorylation, Kinases, and Phosphatases ..... 20

1.1.2. Structural Diversity of Protein Tyrosine Phosphatases ..... 22

1.1.3. Regulatory Roles of Protein Tyrosine Phosphatases ..... 25

1.1.4. Drug Development Targeting Phosphatases ..... 27

1.2. SAMDI-MS and Peptide arrays ..... 30

1.2.1. Self-assembled Monolayers on Gold ..... 30

1.2.2. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry ..... 31

1.2.3. SAMDI-MS and Its Applications ..... 33

1.2.4. Combining Peptide Arrays and SAMDI-MS ..... 34

1.2.5. Profiling Enzyme Activities with SAMDI-MS and Peptide Arrays ..... 35

1.3. Dissertation Overview ..... 38

**Chapter 2. Using SAMDI-MS to Assay an Oncogenic Phosphatase SHP2 And Discover Its Inhibitors ..... 40**

2.1. Introduction ..... 41

2.2. Results and Discussion ..... 43

2.2.1. SHP2 Protein Expression ..... 43

2.2.2. SAMDI-MS Phosphatase Assay ..... 44

2.2.3. SHP2 Inhibitor Screening ..... 46

2.2.4. Determination of $IC_{50}$ Using SAMDI-MS ..... 48

2.2.5. Molecular Modeling and Modification of Adapalene ..... 50

2.3. Conclusion ..... 52

2.4. Methods ..... 53

**Chapter 3. Combining SAMDI-MS and Peptide Arrays to Profile SHP2 Disease Mutants ..... 57**

3.1. Introduction ..... 58

3.2. Results and Discussion ..... 60

3.2.1. Profiling SHP2 Activities With SAMDI-MS and Peptide Arrays ..... 60

3.2.2. Comparing SHP2 WT and Constitutively Active Mutant E76K ..... 62

3.2.3. Exploring Activities and Substrate Selectivities of SHP2 Mutants ..... 63

3.2.4. Using SHP2Δ104 Constructs to Exam Changes in Substrate Selectivity..... 65

3.3. Conclusion ..... 70

3.4. Methods..... 71

**Chapter 4. Profiling Protein Tyrosine Phosphatase Specificity with SAMDI-MS and Peptide Arrays..... 75**

4.1. Introduction..... 76

4.2. Results and Discussion ..... 79

4.2.1. Preparation of Phosphopeptide Array..... 79

4.2.2. Profiling Activities of DEP1 (PTPRJ)..... 81

4.2.3. Profiling Activities of 22 Phosphatases. ..... 83

4.2.4. Quantitative Analysis of Heatmaps. ..... 83

4.2.5. Basic Residues Generally Reduce Substrate Activity ..... 88

4.2.6. Acidic Residues Play a Complex Role ..... 89

4.2.7. Glycine Has Different Effects at X and Z Positions..... 90

4.2.8. Proline Is a Strong Inhibitor for PTPs..... 90

4.2.9. Aromatic and Hydrophobic Residues Are General Activators..... 91

4.2.10. Polar Neutral Residues Are Less Involved in PTP Specificities ..... 91

4.2.11. Comparisons to Prior Studies..... 92

4.2.12. Categorizing PTPs According to Their Substrate Selectivities. ..... 93

4.2.13. A Potential Mutation/Modification Crosstalk Involving PTP1B..... 96

4.3. Conclusion ..... 98

4.4. Methods..... 99

**Chapter 5. Tyrosine Phosphatase Activity Is Restricted by Basic Charge Substituting Mutation of Substrates: Implications for Missense Mutations Associated with Disease ... 102**

5.1. Introduction..... 103

5.2. Results..... 106

5.2.1. Profiling Activities of SHP1 ..... 106

5.2.2. Basic Residues Inhibit PTP Activity..... 107

5.2.3. A Charge-Mutation/PTP-Restriction Mechanism..... 108

5.2.4. Database Search Identifies 6,000+ Relevant Cancer Mutants. ..... 109

5.2.5. β-cat T653 Basic Mutants Show Greater Phosphorylation at Y654. ..... 111

5.2.6. β-cat Basic Charge Mutants Manifest Enhanced Wnt-Signaling. ..... 114

5.3. Discussion and Conclusion..... 116

5.4. Methods..... 119

5.5. Creating β-cat$^{KO}$ HEK 293T cell lines using CRISPR-Cas9 ..... 124

**Chapter 6. A Novel Citrullination/Phosphorylation Crosstalk via Protein Tyrosine Phosphatases..... 127**

6.1. Introduction ..... 128

6.2. Results and Discussion ..... 131

6.2.1. Basic Residues Antagonize PTP Activity ..... 131

6.2.2. Citrullination Restores PTP Activity. ..... 132

6.2.3. MBP Modification Crosstalk Between Y261 And R264 ..... 134

6.2.4. Citrullination Increases PTP $k_{cat}/K_M$ ..... 136

6.3. Conclusion ..... 138

6.4. Methods ..... 139

**Chapter 7. Summary and Perspectives ..... 143**

**References ..... 148**

# List of Figures

Figure 1-1 Protein phosphorylation. ..... 21

Figure 1-2 The general mechanism of the PTP-catalyzed reaction. ..... 22

Figure 1-3 The structures of PTPs. ..... 24

Figure 1-4 Mechanism of SHP2-mediated signal transduction and the effect of SHP2 mutations in leukemia. ..... 26

Figure 1-5 Formation of alkanethiolate self-assembled monolayers on gold. ..... 31

Figure 1-6 SAMDI-MS. ..... 34

Figure 1-7 WT and mutant OmpT activity profiles. ..... 37

Figure 2-1 Expression of active SHP2 and its mutants ..... 43

Figure 2-2 SAMDI-MS phosphatase assay ..... 45

Figure 2-3 Assaying SHP2 activities with SAMDI-MS. ..... 46

Figure 2-4 Using SAMDI-MS for high-throughput SHP2 inhibitor screening. ..... 47

Figure 2-5 $IC_{50}$ measurements of 4 SHP2 inhibitors. ..... 49

Figure 2-6 Molecular docking of adapalene and SHP2 PTP-domain. ..... 51

Figure 3-1 Profiling PTP activities using SAMDI-MS and peptide array. ..... 60

Figure 3-2 Profiling SHP2 WT and E76K activities. ..... 61

Figure 3-3 Profiling full-length SHP2 mutant activities. ..... 64

Figure 3-4 Profiling SHP2Δ104 mutant activities. ..... 68

Figure 3-5 Analysis in amino acid specificity of SHP2Δ104 and mutants ..... 69

Figure 4-1 Profiling phosphatase activities using peptide arrays and SAMDI-MS ..... 80

Figure 4-2 Activity heatmaps for each of 22 phosphatases profiled on the peptide array having 361 sequences of form Ac-TRDXpYZTC-NH2. ..... 82

Figure 4-3 The ratio of the difference (Δ) heatmap of 22 PTPs. ..... 85

Figure 4-4 Histograms showing the role of residues found in PTP specificities. ..... 87

Figure 4-5 Specificity tree of the 22 phosphatases profiled. ..... 95

Figure 4-6 Dephosphoraylation of peptides corresponding to WT and mutant INSR by the PTP1B phosphatase. ..... 97

Figure 5-1 Phosphatase activity profiling using peptide arrays and SAMDI-MS. ..... 107

Figure 5-2 Basic residues antagonize PTP activity..... 108

Figure 5-3 β-cat phosphorylation at Y654 is enhanced by basic amino acid replacement at T653. ..... 113

Figure 5-4 β-cat T653 basic charge mutants display enhanced Wnt-signaling. ..... 115

Figure 5-5 β-cat charge-mutation/PTP-restriction mechanism that enhances Wnt signaling. ..... 118

Figure 5-6 Creating β-cat$^{KO}$ HEK 293T cell lines and their characterization. ..... 126

Figure 6-1 Profiling phosphatase activity using peptide arrays and SAMDI-MS. ..... 132

Figure 6-2 Citrullination increases PTP activity..... 133

Figure 6-3 Dephosphorylation of MBP peptides. ..... 135

Figure 6-4 Kinetic comparison of citrullinated/uncitrullinated MBP peptide dephosphorylation. ..... 137

# List of Tables

Table 1-1 Selected list of phosphatase inhibitors in clinical trials..... 29

Table 3-1 SHP2Δ104 mutants and their predicted effect on SHP2 function..... 65

Table 3-2 Primer list for full length SHP2 mutants. ..... 71

Table 4-1 Summary of PTP selectivity..... 86

Table 4-2 Reaction conditions for profiling phosphatase activities..... 101

# Chapter 1. Introduction to Protein Tyrosine Phosphatases, SAMDI-MS, and Peptide Arrays

## *1.1. Protein Tyrosine Phosphatases*

### *1.1.1. Protein Phosphorylation, Kinases, and Phosphatases*

Protein phosphorylation is a post-translational modification (PTM) that adds a phosphate ion (Pi) to amino acid sidechains. It often occurs on serine (86.4%), threonine (11.8%), and tyrosine (1.8%) through phosphoester bond formation.[1] Recent studies have demonstrated a widespread phosphorylation of human proteins on other residues such as histidine, aspartate, glutamate, cysteine, arginine and lysine.[2] Phosphorylation provides a mechanism for the reversible and rapid control of protein function and the regulation of diverse cellular processes, development and signal transduction. It is estimated that 30% of proteins are phosphorylated and 3.5% of human genes are devoted to the expression of proteins that directly regulate protein phosphorylation, including kinases and phosphatases.[3]

Kinases catalyze the addition of phosphate ions and phosphatases catalyze the removal of phosphate ions (Fig. 1-1). For serine and threonine, there are 428 protein kinases (PKs) and 30 protein phosphatases (PPs) responsible for regulating their phosphorylation; for tyrosine phosphorylation, there are 90 protein tyrosine kinases (PTKs) and 107 protein tyrosine phosphatases (PTPs) in charge.[4-5] While the phosphorylated states of proteins are determined by the balance of opposing kinase and phosphatase activities, the overwhelming majority of research work has been carry out to address the roles of kinases and their substrates in regulating phosphorylation, and has generally assumed that phosphatases serve a non-regulatory housekeeper role.[6] However, this assumption lacks justification and appears inconsistent with the roughly equal numbers of PTPs and PTKs encoded in the human genome. Furthermore, recent

work has illustrated a regulatory role for PTPs and reported their sophisticated modes of regulation.[7-10] Dysregulated PTP activities have also been directly linked to disease and cancer. For example, SHP2 (PTPN11) has been identified as the first oncogenic phosphatase.[11-13] Therefore, the understanding of phosphorylation and signaling would greatly benefit from advancing our understanding of the roles that PTPs play using novel bioanalytical technologies.

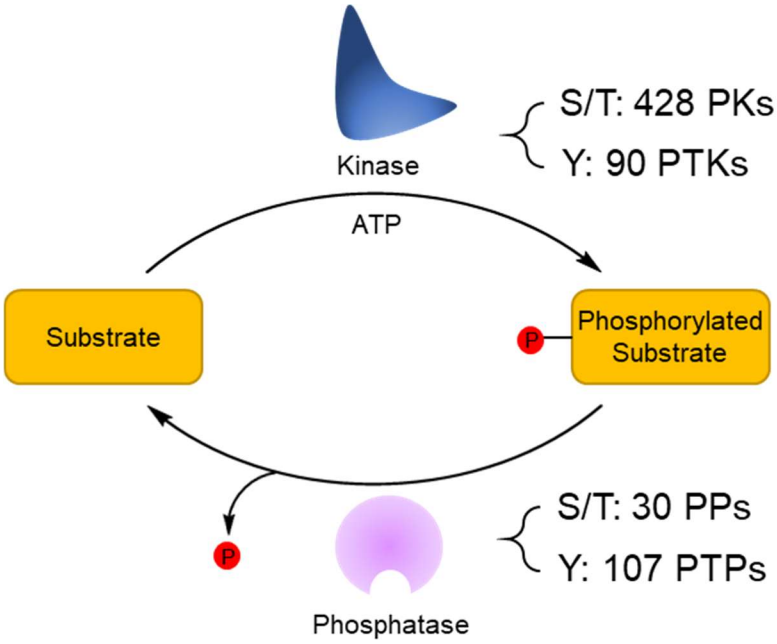



**Figure 1-1 Protein phosphorylation.**
The phosphorylated states of proteins are determined by the balance of opposing kinase and phosphatase activities. Their expression is regulated by 3.5% of human genes.

We note that serine and threonine phosphatases rely on very diverse subunits to exhibit its regulatory roles.[14] Therefore, the significant smaller number of PP compared to PK does not imply that PP is non-regulatory. However, the study of PP is beyond the scope of this dissertation.

### *1.1.2. Structural Diversity of Protein Tyrosine Phosphatases*

Classical PTPs are defined by a signature active site sequence motif HCXXGXXRS(T) or $HC(X)^5R$ where a cysteine residue acts as the nucleophile and the conserved arginine residue folds back toward the phosphate-binding pocket to assist in substrate binding and catalysis (Fig. 1-2).[15] The dephosphorylation reaction is completed by another conserved motif, the WPD-loop, protonating the tyrosine residue and catalyzing the hydrolysis of the phosphate leaving group.[16] Identification of these motifs and the complete sequencing of the human genome made it possible to catalogue the 107 genes that comprise the PTP family.

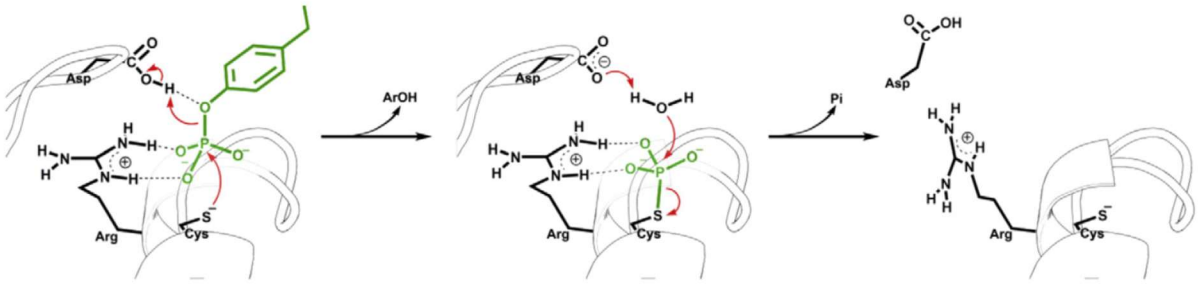



**Figure 1-2 The general mechanism of the PTP-catalyzed reaction.**
PTP-conserved $HC(X)^5R$ motif and WPD-loop catalyze the tyrosine dephosphorylation reaction. Reprinted with permission from Brandao, T. A.; Johnson, S. J.; Hengge, A. C., The molecular details of WPD-loop movement differ in the protein-tyrosine phosphatases YopH and PTP1B. *Arch. Biochem. Biophys.* **2012,** *525* (1), 53-9. Copyright © 2012 Elsevier Inc.

The approximately equal number of PTP and PTK genes suggest similar levels of structural and functional complexity between the two families. Although all PTPs share the same catalytic mechanism, their structures are extremely diverse. Among 37 classical PTPs (Fig. 1-3A), there are 21 transmembrane, receptor-like PTPs (RPTPs). The diversity in the extracellular segments of the RPTPs presumably reflects a similar diversity of the ligands they respond to. However, the

identity and function of these ligands remains largely unknown.[17] The other 16 classical PTPs are non-transmembrane/cytoplasmic. The structural diversity of cytoplasmic PTPs can be defined by a wide selection and combinations of regulatory sequences or adaptor domains that flank the catalytic domain. Some domains modulate the PTP activity to specific substrates directly. For example, SH2 domains found on SHP1 and SHP2 recognize phosphotyrosine. Others control the subcellular localization of the protein. For example, FERM domains that are widely expressed in 5 different PTPs are involved with localizing proteins to the plasma membrane.[9, 18]

Non-classical PTPs are broadly defined as dual specificity phosphatases (DSPs). Although they are also characterized by the presence of the $HC(X)^5R$ active site motif and share the same catalytic mechanism as the classical PTPs, the structure of the DSP active site allows them to accommodate phosphoserine (pS) and phosphothreonine (pT) residues in addition to phosphotyrosine (pY) residues in substrates. DSPs are more structurally diverse than the classical PTPs and possess a much smaller conserved catalytic domain in combination with a wide variety of regulatory sequences or adaptor domains (Fig. 1-3B).[9]

The number of PTP genes only illustrates the minimal level of diversity in the family. Additional complexity is introduced using alternative gene promoters, alternative mRNA splicing and post-translational modifications.[9] The structural diversity of PTPs indicates the importance of PTP function in the regulation of cell signaling.

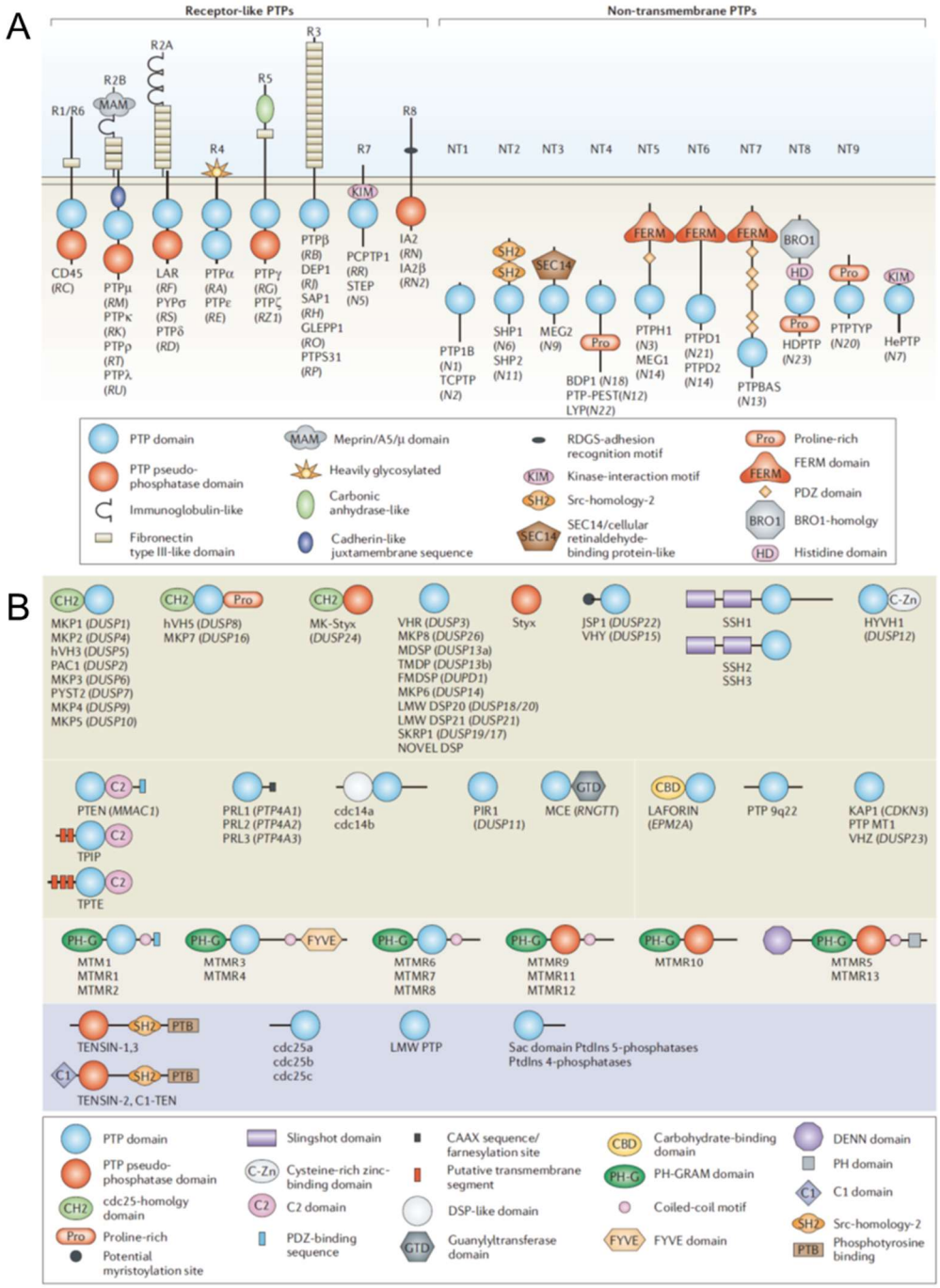



**Figure 1-3 The structures of PTPs.**

A) The classical PTPs. B) The dual specificity phosphatases. Reprinted with permission from Tonks, N. K., Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* **2006,** *7* (11), 833-46. Copyright © 2006 Nature Publishing Group.

### *1.1.3. Regulatory Roles of Protein Tyrosine Phosphatases*

The greatest misconception about the roles of phosphatases in regulation of cell signaling is that phosphatases are merely constitutively active housekeeping enzymes. In contradiction, phosphatases are specific, essential regulators of signal transduction and serve complimentary roles in coordination with kinases, like yin and yang.[19] One PTP in particular, SHP2 has drawn great attention from the pharmaceutical and drug development industry. Below, we will discuss its regulatory roles highlight the importance of the PTP family.

SHP2 (PTPN11) has been identified as the first oncogenic phosphatase.[11-13] The structure of SHP2 contains one PTP domain and two adaptor domains (N-SH2 and C-SH2) at the N-terminus, both of which are SH2 domains that recognize and bind to phosphotyrosine residues. In its native conformation, the E76 residue in the N-SH2 domain forms an important hydrogen bonding interaction with the S502 sidechain in the PTP domain, bringing both domains together into a closed-form structure hindering the PTP active site from its substrate.[20] Activation of surface receptor tyrosine kinases (RTKs), and the subsequent tyrosine phosphorylation of binding sites for the SHP2 SH2 domains lead to recruitment of SHP2. Occupation of the SH2 domains by phosphotyrosines results in a conformational change that resolves the inhibitory interaction of the N-SH2 domain and the PTP domain (Fig 1-4a).[21] Activated SHP2 dephosphorylates substrates that are inhibitory for signaling toward ERK1/ERK2, AKT or STAT5. These include binding sites for cSRC kinase, a negative regulator of the Ras-pathway, leading to its activation.[22]

Leukemia-associated mutations in SHP2 such as E76K, E76Q and S502P impair the inhibitory interaction of the N-SH2 domain and PTP domain, leading to increased basal PTP

activity and increased affinity of the SH2 domains for pY ligands (Fig 1-4b).[21] Alterations of substrate selectivity of the enzyme by the aforementioned mutations was also suggested.[23] The constitutively active SHP2 structure caused by mutations facilitates upstream ligand activation and increases downstream signaling related to cell proliferation and migration and drives disease development.[22]

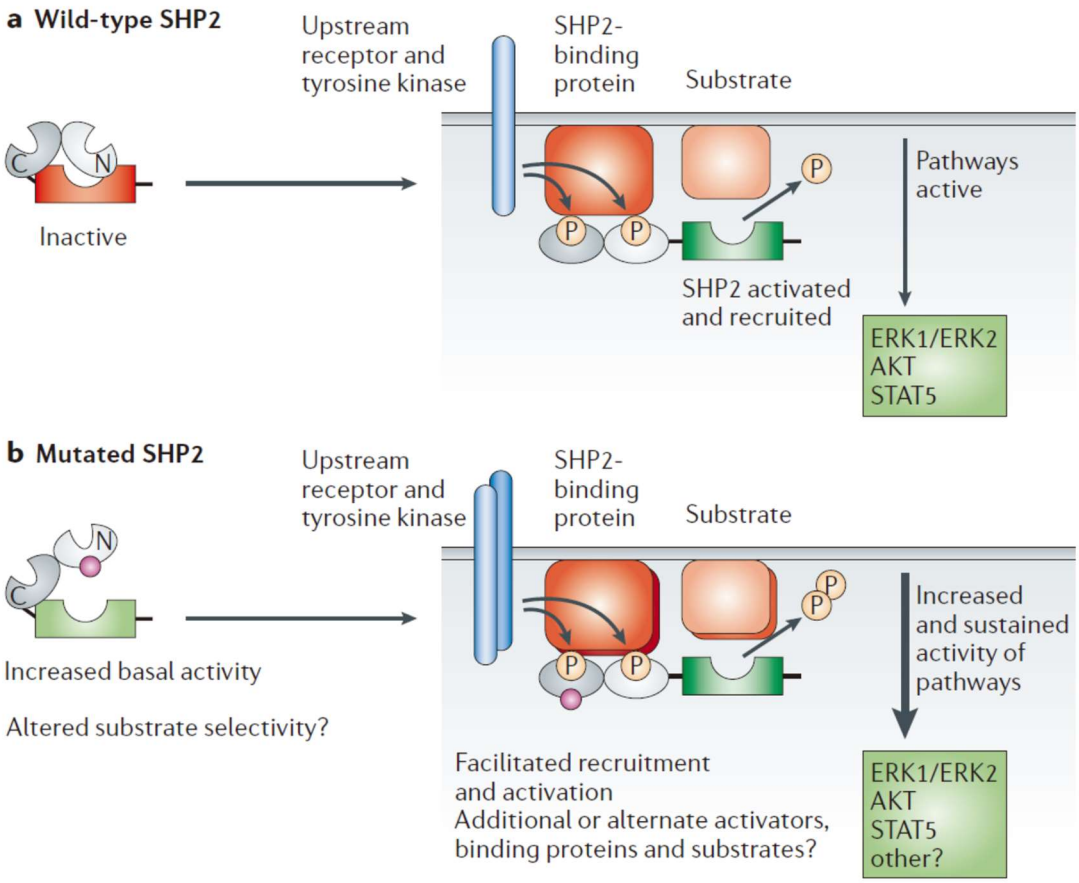



**Figure 1-4 Mechanism of SHP2-mediated signal transduction and the effect of SHP2 mutations in leukemia.**

a. WT SHP2 is inactive in the absence of upstream activation. b. Leukemia-related SHP2 mutant is constitutively active. Reprinted with permission from Ostman, A.; Hellberg, C.; Bohmer, F. D., Protein-tyrosine phosphatases and cancer. *Nat. Rev. Cancer* **2006,** *6* (4), 307-20. Copyright © 2006 Nature Publishing Group.

In addition to SHP2, other PTPs has emerged in cancer research recently. PRL1 and PRL3 involved in AKT and ERK signaling pathways are also identified as oncoproteins leading to prostate, colorectal, ovarian, and pancreatic cancers. PTPRT, PTEN and PTPRD are reported tumor suppressors for their unique kinase counteracting roles in common cancer-related pathways such as EGFR and STAT3 signaling.[24] Beyond cancers, PTP1B, a redox sensitive PTP where its activity can be tuned reversibly with reactive oxygen species, is found closely related to insulin insensibility in diabetic and obese patients.[25] All these examples suggest that PTPs are not merely constitutively active housekeeping enzymes. PTP activities are carefully regulated by delicate mechanisms. PTPs are critically involved in key signaling pathways and their dysregulation directly leads to disease.

### *1.1.4. Drug Development Targeting Phosphatases*

The imbalanced effort put into kinase and phosphatase research is best reflected in their drug development. Till the end of 2020, FDA has approved 62 kinase inhibitors, a number that has doubled in the past 5 years and continues to grow rapidly.[26] Among those approved, 48 inhibitors target receptor or non-receptor tyrosine kinases, 10 for serine/threonine kinases and 4 for dual functional kinase. This number reflects the key roles tyrosine phosphorylation plays despite its 1.8% composition of the phosphoproteome in the cells.[1] On the other hand, the development of phosphatase drugs are much slower the kinase. There are only two FDA-approved phosphatase inhibitors: FK506 (Tacrolimus) and cyclosporin A (Ciclosporin). At the time of the approval, their detailed mechanism of action was not yet clear. It was later shown that they bind to the LxVP-SLiM-binding site on the calcium and calmodulin dependent serine/threonine protein

phosphatase (calcineurin, PP2B).[27-28] To date, there are still no FDA approved drugs targeting PTPs.

PTPs have long been stigmatized as undruggable for their extremely conserved, and hydrophilic active sites across the family. It is very challenging to target one specific PTP without inhibiting others in the same family. In addition, active site inhibitors are often too hydrophilic to cross the cell membranes. Therefore, drug specificity and bioavailability are two major obstacles in the development of PTP active site inhibitors.[29] With the discovery of critical roles phosphatases play in disease development in the last few years, interest in phosphatase biology and in phosphatase drug discovery has been strongly renewed.[30] In 2016, Novartis introduced a new concept to inhibit SHP2 that changed the game. They reported an allosteric inhibitor SHP099 that binds into the cleft between the three domains, trapping SHP2 in its closed conformation to inhibit its activity.[31-32] The strategy cleverly bypassed the active sites to overcome the long-standing specificity and bioavailability issues with PTP inhibitors. The work gained momentum in the field and quickly inspired more design of allosteric phosphatase inhibitors.[33] Selected phosphatase inhibitors in clinical trials are listed in Table 1-1. Phosphatases have regained strong attention and are now actively pursued in drug development.

| drug | company | target | property | status | indication | NCT identifier |
|---|---|---|---|---|---|---|
| **IFB-088** | InFlectis BioScience | PPP1R15A | Allosteric inhibitor | Phase I | Charcot–Marie–Tooth disease | NCT03610334 |
| **LB-100** | Lixte Biotechnology | PP2A | Catalytic inhibitor | Phase I | Solid tumors | NCT01837667 |
| | | | | Phase II | Recurrent glioblastoma | NCT03027388 |
| **AKB-9778** | Aerpio Pharmaceuticals | VE-PTP | Catalytic inhibitor | Phase II | Retinal vein occlusion | NCT02387788 |
| | | | | | Nonproliferative diabetic retinopathy | NCT03197870 |
| | | | | | Diabetic macular edema | NCT01702441 |
| | | | | | | NCT02050828 |
| **PRL3-zumab** | A*STAR | PRL3 | Monoclonal antibody | Phase I | Advanced solid tumors | NCT03191682 |
| **MSI-1436** | Genaera Corporation | PTP1B | Allosteric inhibitor | Phase I | Diabetes mellitus | NCT00806338 |
| **JAB-3068** | Jacobio | SHP2 | Allosteric inhibitor | Phase I | Advanced solid tumors | NCT03518554 |
| **TNO155** | Novartis | SHP2 | Allosteric inhibitor | Phase I | Advanced solid tumor | NCT03114319 |
| **RMC-4630** | Revolution Medicines/Sanofi | SHP2 | Allosteric inhibitor | Phase I | Relapsed/refractory solid tumor | NCT03634982 |
| **RLY-1971** | Relay Therapeutics | SHP2 | Allosteric inhibitor | Phase I | Advanced or metastatic solid tumors | NCT04252339 |

**Table 1-1 Selected list of phosphatase inhibitors in clinical trials.**

Reprinted with permission from Yuan, X.; Bu, H.; Zhou, J.; Yang, C. Y.; Zhang, H., Recent Advances of SHP2 Inhibitors in Cancer Therapy: Current Development and Clinical Application. *J. Med. Chem.* **2020,** *63* (20), 11368-11396. Copyright © 2020 American Chemical Society.

## *1.2. SAMDI-MS and Peptide arrays*

### *1.2.1. Self-assembled Monolayers on Gold*

Self-assembled monolayers (SAMs) are ordered molecular assemblies formed by the adsorption of an active surfactant on a solid surface. The formation of these two-dimensional molecular structure is driven by a spontaneous chemical reaction at the interface, as the system approaches thermodynamic equilibrium. SAMs of functionalized long-chain hydrocarbons, such as fatty acids, organosilicon derivatives and alkanethiolates are the most frequently used building blocks of supramolecular structures.[34]

Alkanethiolates are the most common monomers used to form SAMs on gold surfaces through unique gold-thiolate bonds. The Mrksich group pioneers in using such surfaces for a wide variety of applications including the development of different surface chemistries[35-37], electrochemical biosensors[38], enzyme assays[39], molecular patterning and cellular controls[40-41]. These alkanethiolates are usually terminated with tri(ethylene glycol) group (Fig. 1-5A) to provide an inert background which is resistant to non-specific protein adsorption for biological applications.[42] At a surface density between 5-20%, the surface is presented with various reactive functional groups for substrate immobilization. For examples, azides/alkynes for click chemistry[43], benzoquinones/cyclopentadienes for Diels-Alder reaction[44], biotins for streptavidin capture[45], diazirines for photocrosslinking[46] and most commonly, maleimides for the Michael addition with thiols (Fig. 1-5B)[47]. The maleimide presenting SAMs (Fig. 1-5C) are extremely versatile to immobilize any cysteine-containing peptides and are used for all works in this dissertation.

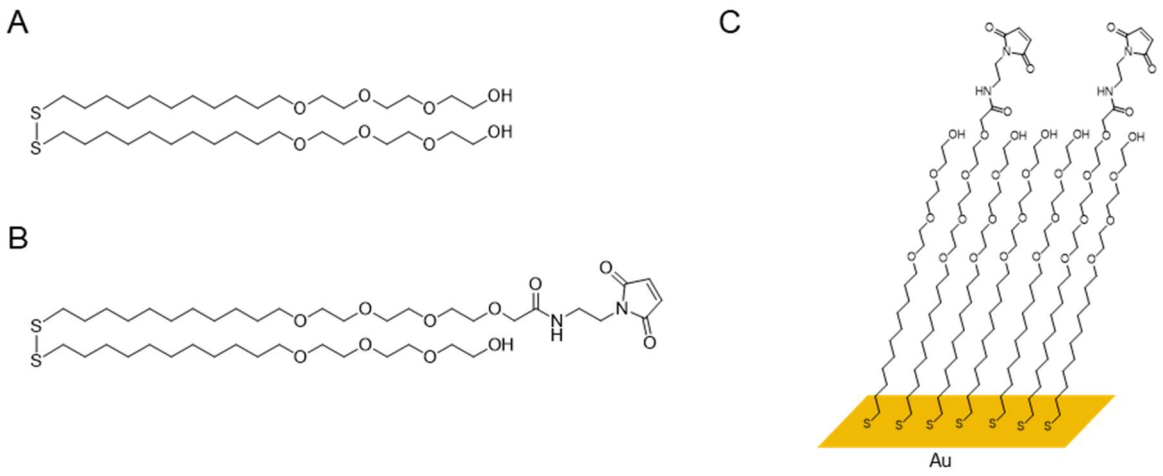



**Figure 1-5 Formation of alkanethiolate self-assembled monolayers on gold.**
A) Tri(ethylene glycol)-terminated $C_{11}$-alkane disulfide. B) $C_{11}$-alkane disulfide with one terminal tri(ethylene glycol) and one terminal maleimide C) Alkanethiolate self-assembled monolayer on gold. The SAMs present functional maleimide groups with a background of tri(ethylene glycol) groups.

### *1.2.2. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry*

Mass spectrometry (MS) is an analytical technique to measure mass-to-charge ratio of ions. The analyte's molecular weight (MW) ranges from several Da to several hundreds of kDa. The versatility of MS makes it extremely useful in biological research: from characterization of small molecule metabolites to large biomolecules such as proteins and DNA.[48] Among all ionization methods, matrix-assisted laser desorption/ionization (MALDI) is a unique soft ionization process that allows direct MS characterization of solid analytes. In MALDI, the sample is first mixed with a matrix material and applied to a metal plate. Then, a pulsed laser irradiates the sample. The energy is absorbed by the matrix and subsequently transferred to the analyte, triggering the

desorption and ionization of the target analyte. The instrument is often coupled with a time-of-flight (TOF) analyzer to resolve ions by their mass-to-charge ratio.[49]

The ionization method MALDI provides at the solid-gas ion interface allows the rapid analysis of surfaces and the preservation of spatial information. Therefore, MALDI-MS is widely used in high-throughput screening and imaging.[50-51] There are also some earlier works that combine MALDI-MS and SAM surfaces. For example, the surface-enhanced laser desorption/ionization (SELDI) method employs SAM surfaces on bead for immunoassay.[52-53] However, it is not concerned with characterizing the alkanethiolates of the monolayer. Another approach, surface-assisted laser desorption/ionization (SALDI) characterizes SAMs on gold nanoparticles. While the method directly characterize the alkanethiolates of the monolayer, it is limited to nanoparticle surfaces and only allows for small (<1,000 Da) analytes.[54-55] Wilkins and Hanley first applied laser desorption mass spectrometry to characterize SAMs – including alkanethiols and disulfides – and their complexes with gold atoms.[56-57] However, their work did not use a standard MALDI-TOF mass spectrometers and did not investigate monolayers derivatized from alkanethiols that were substituted with functional groups and their interfacial reactions.

### *1.2.3. SAMDI-MS and Its Applications*

The Mrksich group advanced the technology characterizing SAMs on gold and their modification by MALDI-MS and introduced self-assembled monolayer laser desorption/ionization mass spectrometry (SAMDI-MS).[58] In SAMDI-MS (Fig. 1-6), SAMs are coated with matrix and irradiated with laser. Similar to MALDI, the matrix absorbs energy and transfers to SAMs, breaking the gold-thiol bonds and causing the desorption and ionization of the alkanethiolate monomers or corresponding alkyldisulfide molecules, which are subsequently resolved by a TOF m/z analyzer. This method integrates well with the standard MALDI-TOF instrumentation and allows the investigation of monolayers derived from alkanethiols that were functionalized. It was first demonstrated to characterize peptides, proteins and carbohydrates immobilized to the SAM surface[59] and quickly developed into a high-throughput, label-free platform to monitor chemical and enzymatic reactions on a wide variety of substrates utilizing different immobilization chemistries.[39] In these assays, treatment of the monolayers with a solution containing the reactants or enzymes lead to modifications of the substrate, which is accompanied by a corresponding change in mass. The reaction products can then be quantitated with SAMDI-MS. The assay has been used to measure a broad range of enzyme activities including acetyltransferases, deacetylases, proteases, kinases, glycosyltransferases and more.[58, 60-63] The technology is also commercialized by SAMDI Tech Inc., which uses SAMDI-MS to assist clients in the pharmaceutical industry in early phase drug discovery.[64]

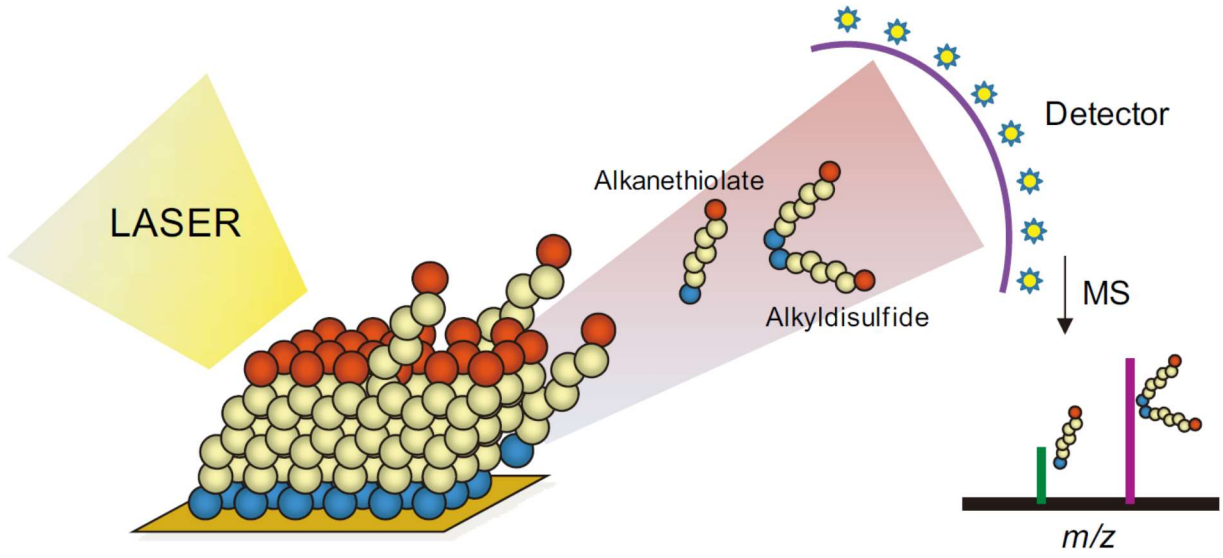



**Figure 1-6 SAMDI-MS.**

SAMDI-MS combines self-assembled monolayers and MALDI mass spectrometry to rapidly characterize the masses of alkanethiolates in a monolayer, useful for monitoring chemical and biochemical reactions at surfaces. Reprinted with permission from Mrksich, M., Mass spectrometry of self-assembled monolayers: a new tool for molecular surface science. *ACS Nano* **2008,** *2* (1), 7-18. Copyright © 2008 American Chemical Society.

### *1.2.4. Combining Peptide Arrays and SAMDI-MS*

Biomolecular arrays are planar substrates that have large numbers of molecules immobilized on a surface in patterns where each region presents one unique molecule. Most arrays are prepared from either small molecules, oligonucleotides, peptides, or proteins. Oligonucleotide arrays are the most developed, commercially available, and widely used in clinical application and genetic research. Motivated by the success of DNA arrays, peptide arrays are drawing attention of researchers because they may provide insights to questions beyond the genetic level,

such as post-translational modifications (PTMs), protein−protein interactions, protein−substrate interactions, and enzyme sequence specificity.[65]

Synthesis of large peptide libraries rely on Fmoc-chemistry-based solid phase peptide synthesis (SPPS).[66] There are two approaches to construct peptide arrays with synthetic peptides. First is the *in-situ* approach where peptides are directly synthesized on the solid support (typically membrane or glass) of the arrays. SPOT[67], particle-based synthesis[68], and photolithographic[69] peptide arrays all employed this strategy. Despite its convenience, this approach often requires specialized instrumentation to generate the arrays, the quality of the peptides is lower and unverifiable, and the solid supports may cause background adsorption issues. Another approach is to synthesize peptides with standard SPPS and then immobilize/pattern them onto the array surface through chemical reactions.[70] The latter approach integrates particularly well with SAMDI-MS assays because peptide immobilization chemistry to the SAM surface is extremely versatile and specific. Combined with the high-throughput measurement capacity of SAMDI-MS, it is an ideal platform to characterize enzymatic reactions on large peptide arrays.[65]

### *1.2.5. Profiling Enzyme Activities with SAMDI-MS and Peptide Arrays*

Many PTMs of proteins – including methylation, acetylation, glycosylation, and phosphorylation – can be measured to investigate both enzyme activity and substrate specificity using the combination of SAMDI-MS and peptide arrays. For example, multiple studies from the Mrksich group have elucidated the biochemistry of acetylation. Histone deacetylases (HDACs)

and lysine deacetylases (KDACs) were profiled to understand their amino acid sequence specificities.[71-72] It was also demonstrated that SAMDI-MS and peptide arrays can be used to profile KDAC activities in cell lysates and showed that KDAC profiles shifted as a result of different cell developmental stages and responses to inhibitors.[73] By treating peptide arrays with acetic anhydride, a sequence-specific acetylation mechanism of peptides containing a histidine-tyrosine (HY) dyad was discovered; varying the amino acids around the dyad revealed the disfavored proximity of acidic residues towards acetylation.[74]

Sequence specificity information acquired by SAMDI-MS and peptide arrays has also helped improve peptide-based assays drastically. For example, the detection of a bacterial outer-membrane protease OmpT was enhanced by 400-fold with the optimized FRRV sequence.[75] It was also shown that OmpT activity profiles can reveal changes in substrate specificities caused by mutations (Fig. 1-7). In another example, tags were designed to incorporate glycosylation into proteins by a systematic platform for glycosylation sequence characterization and optimization by rapid expression and screening (GlycoSCORES), which combines cell-free protein synthesis and SAMDI-MS/peptide arrays.[76] The technologies enable the development of substrate-specific glycosyltransferases and the analysis of glycoprotein structure-function relationship.[77-78]

Combining SAMDI-MS and peptide arrays is a powerful approach to comprehensively characterize a wide range of chemical and enzymatic modifications. Importantly, this can be achieved for enzymes notoriously difficult to assay, such as phosphatases. This dissertation aims to combine SAMDI-MS and peptide arrays to study PTPs and gain insight in their cellular regulatory roles.

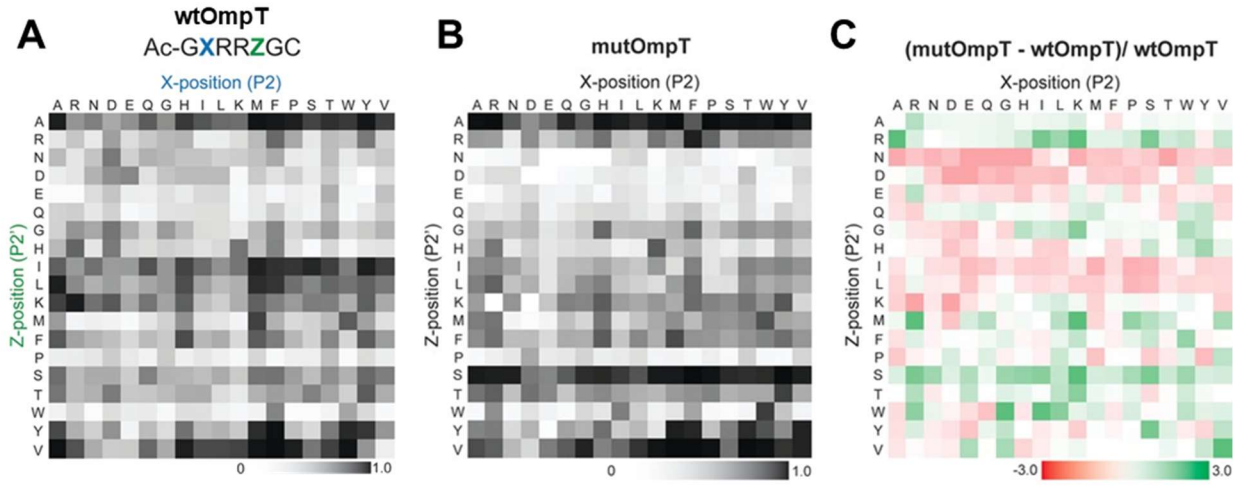



**Figure 1-7 WT and mutant OmpT activity profiles.**

A) WT OmpT. B) Mutant OmpT. C) Differential profile reveals that mutOmpT significantly prefers a serine residue at Z-position, but the WT does not. Reprinted with permission from Wood, S. E.; Sinsinbar, G.; Gudlur, S.; Nallani, M.; Huang, C. F.; Liedberg, B.; Mrksich, M., A Bottom-Up Proteomic Approach to Identify Substrate Specificity of Outer-Membrane Protease OmpT. *Angew. Chem. Int. Ed. Engl.* **2017,** *56* (52), 16531-16535. Copyright © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

## *1.3. Dissertation Overview*

In this dissertation, I present the original work using SAMDI-MS and peptide arrays to study protein tyrosine phosphatases and discover new cellular regulatory mechanisms they are critically involved in.

Chapter 2 reports the SAMDI-MS high-throughput PTP assay and its application in drug discovery for the oncogenic phosphatase SHP2. We use the assay to screen an FDA-approved compound library and identify several potential hits. In Chapter 3, we introduce a pY peptide library to profile SHP2 mutants and understand their roles in disease developments. We report mutation hot spots and classes of SHP2 mutants involving changes in substrate specificities. The data would contribute to better understanding of SHP2 disease pathology.

In Chapter 4, we expand the study using pY peptide library to other PTPs in the proteome. From 22 PTP profiles, we report distinct PTP classes as well as the general rule that inclusion of a basic lysine or arginine residue on either side of the phosphotyrosine decreases PTP activity. We discuss a new perspective on the role of an R1152Q mutant in the insulin receptor, which is known to exhibit a lower phosphorylation level, and which our work suggests may be due to an increased activity towards phosphatases. We hypothesize that PTP activity restriction by basic amino acids adjacent to phosphotyrosines may have a broader impact in disease development.

We next sought to validate this hypothesis in Chapter 5 using two β-catenin mutants associated with cancer (T653R/K) and a mouse model for intellectual disability (T653K). These mutants introduce a basic residue next to Y654, an established phosphorylation site where modification shifts β-catenin from cell-cell adhesions and towards its essential nuclear role as

Wnt-signaling effector. We show that T653-basic mutant β-catenins are less efficiently dephosphorylated by phosphatases, leading to sustained Y654 phosphorylation and elevated Wnt signals, similar to those observed for Y654E phospho-mimic mutant mice. This model rationalizes how basic mutations proximal to phosphotyrosines can restrict counter-regulation by phosphatases, providing new mechanismistic and treatment insights for 6,000+ potentially relevant cancer mutations.

In Chapter 6, we report a novel PTM crosstalk mechanism that the normal cells the utilize this PTP activity restriction by basic amino acids for regulation. We demonstrate that citrullinating an arginine adjacent to a phosphotyrosine restores the substrate's susceptibility to phosphatases, resulting in further dephosphorylation. We validate this crosstalk model using peptides adapted from myelin basic protein with known citrullination and phosphorylation sites proximal to each other. We find that citrullination increases the peptide dephosphorylation $k_{cat}/K_M$ by 2.3-fold. The results suggest that one modification can lead to a cascade of signals and provide new insights for phospho-regulation by phosphatases that could be generalized to other charge-altering PTMs.

The dissertation concludes in Chapter 7 with a summary of described research and a discussion about further research for advancing the understanding of phosphatases.

# Chapter 2. Using SAMDI-MS to Assay an Oncogenic Phosphatase SHP2 And Discover Its Inhibitors

## *2.1. Introduction*

Phosphorylation plays an important role in regulating diverse cellular process and signal transduction, including the MAPK pathway.[79] In this pathway, the signal is initiated by EGF ligands associating with transmembrane receptor EGFR, which then dimerizes and activates its receptor-tyrosine kinase (RTK) activities, inducing EGFR autophosphorylation. Subsequently, a complex of SHP2, GAB1 and GRB2 assembles, resulting in the release of the activated Raf. Then, the signal is passed down by a series of phosphorylation modification catalyzed by MAPK and its upstream kinases. Finally, ERK phosphorylates transcription factors that are able to enter the nucleus to initiate the expression of proteins regulating proliferation, cell migration, and apoptosis.[22] Abnormity in this delicate machinery is associated with disease and Leukemia is one where MAPK pathway upregulation is involved.[80]

Among all the kinases, SHP2 is the only phosphatase that are critically involved in the regulation of MAPK pathway and its dysregulation is directly linked to cancer. It is found that 35% of juvenile myelomonocytic leukemia (JMML) patients have mutations in SHP2. E76 mutations are commonly observed (E76K, E76Q, E76G, E76A and E76V).[81-82] These mutants are constitutively active and exhibit elevated phosphatase activities. The MAPK pathway is therefore upregulated, resulting in abnormal cell proliferation.[20-21, 83-85] SHP2 has been identified as the first oncogenic phosphatase and drawn a lot of attention from the pharmaceutical and drug development industry.[11-13]

Unlike kinases, for which many effective assays are available, it remains difficult to measure phosphatase activity, particularly in complex samples such as cell lysates.[86-88] Current assays for

monitoring phosphatase activity rely on measuring decreases in the amount of phospho-substrate – typically with $^{32}$P-labeling or ELISA – or by measuring the Pi by-product.[89-91] Other assays such as PNPP and DIFMUP are non-selective where different phosphatases interfere with each other.[87-88] Therefore, the field will benefit from novel assays to selectively measure PTP activity and use for high-throughput screening.

Our laboratory has developed SAMDI-MS, a label-free, high-throughput analytical method for measuring enzyme activities, including phosphatase activity.[58, 60, 92-93] SAMDI-MS uses SAMs of alkanethiolates on gold that present a substrate (mostly peptides) for the enzyme of interest. In this chapter, we describe a SAMDI-MS assay for PTP. This assay is suitable for measuring SHP2 activity in cell lysates and demonstrates excellent validity and reliability with a screening Z'-factor of 0.86. We report several inhibitors identified using the assay to screen an FDA-approved compound library. Finally, we use molecular modeling to assist in understanding the inhibitory mechanism of one molecule – adapalene – and designing new structures to improve binding affinity.

## *2.2. Results and Discussion*

### *2.2.1. SHP2 Protein Expression*

SHP2 plasmids, WT/pGEX1, E76K/pcDNA1 (constitutively active) and C463S/pcDNA1 (global deactivated) were obtained from Dr. Elizabeth A Eklund's lab. SHP2 WT, E76 and C463S were subcloned to pET21d plasmids with N-terminal his-8-tags. The plasmids were confirmed by DNA sequencing, transformed into BL21(DE3) *E. coli.* and induced with IPTG for expression. The expressed proteins were purified with cobalt columns (Fig. 2-1A) and the SHP2 activities were confirmed by PNPP assay in which WT and E76K exhibited PTP activities but C463S did not in lack of active site cysteine (Fig. 2-1B). Constitutively active E76K mutant was more active than WT, consistent with the structural prediction.

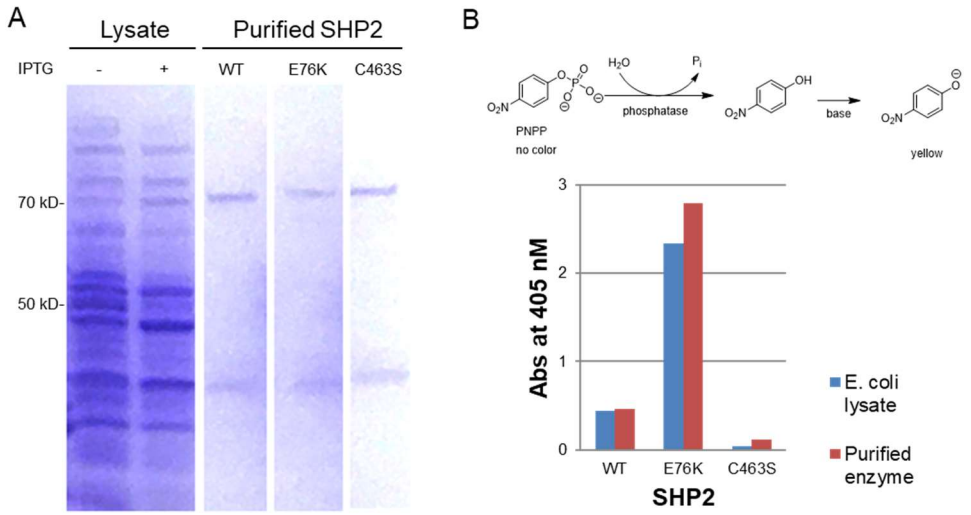



**Figure 2-1 Expression of active SHP2 and its mutants.**

A) Plasmids containing SHP2 and its mutants were transformed to *E. coli.* and induced with IPTG for expression. The proteins were purified with cobalt columns using the his-tags. B) The activities of expressed proteins were confirmed by PNPP assay.

### *2.2.2. SAMDI-MS Phosphatase Assay*

We next used SAMDI-MS to assay the activity of SHP2. The workflow is illustrated in Fig. 2-2. First, self-assemble monolayers were prepared on evaporated gold surface with tri(ethylene glycol)-terminated $C_{11}$-alkane disulfide presenting 10% maleimide. Second, a synthetic peptide Ac-TRDIpYETC, a known substrate of SHP2[94], was immobilized to the surface at pH 7.5 through maleimide-cysteine thiol conjugation. Enzymes (SHP2 WT, E76K and C463) were then applied to the surface and incubated in a humidified chamber at 37 ºC for 30 min, rinsed with water and ethanol, and finally treated with THAP (2,4,6-trihydroxyacetophenone) matrix. The plate was analyzed using an AbSciex 5800 MALDI-TOF mass spectrometer to acquire mass spectra for each spot, which revealed separate peaks corresponding to the substrate and product of the reaction. The dephosphorylation reaction was characterized by the loss of Pi (-80 Da). The conversion of phosphopeptide to its product was quantified by integration of the corresponding peaks and was given by Activity = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 % where AUC referred to the area under the curve.

In Fig. 2-3A, the described method was used to assay PTP activities in *E. coli.* lysates. The lysate concentrations were normalized using Bio-Rad protein (Bradford) assay. In spectra of WT and E76K, mass shifts of -80 Da were observed and the activities were found to be 32 % and 93 % whereas in control (*E. coli.* not induced with IPTG) and C463S, less than 5% activities from endogenous bacterial phosphatases were observed. In Fig. 2-3B, a peptide containing both pT and pY was used to demonstrate the selectivity of the enzymes. Both WT and E76K SHP2 showed reactivities specific to pY because only the mono-dephosphorylated product was observed but no corresponded to further pT dephosphorylation.

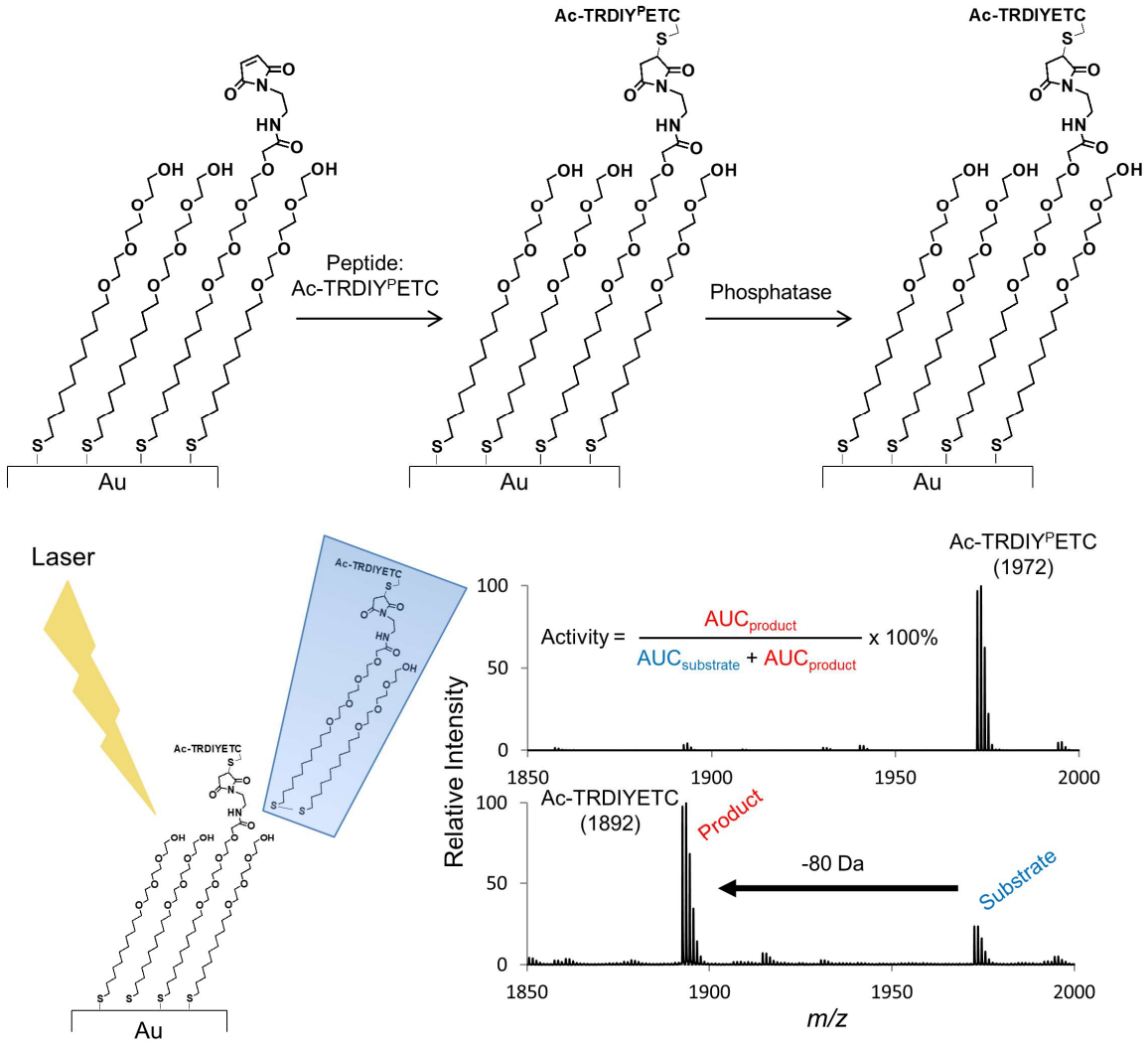



**Figure 2-2 SAMDI-MS phosphatase assay.**

A pY-containing peptide (Ac-TRDIpYETC) is immobilized to the SAMs and treated with enzyme solutions. The surface is analyzed with SAMDI-MS and the activity is defined by comparing the substrate and product AUCs.

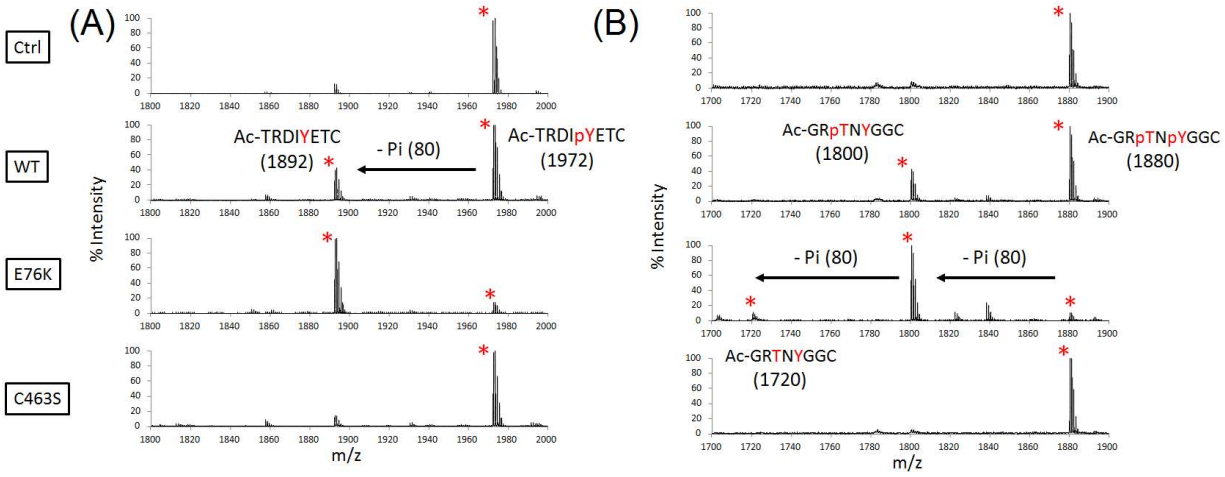



**Figure 2-3 Assaying SHP2 activities with SAMDI-MS.**

A) Assaying SHP2 activities in *E. Coli* lysates. Consistent with the results of PNPP assay, E76K is the more active than WT. B) Verification of SHP2 substrate selectivity using a peptide containing pT and pY. Only one dephosphorylation (-pY) is observed.

### *2.2.3. SHP2 Inhibitor Screening*

The inhibitor screening was conducted with 200 nM SHP2 WT at 37 °C for 30 min where the screening Z'-factor was found to be 0.86 (Fig. 2-4A). Screening Z'-factor is an indication of screening quality defined as:

$$Z' = 1 - [\frac{3\,(\sigma_P + \sigma_N)}{|\mu_P + \mu_N|}]$$

P and N denote the positive and negative controls. Positive control is defined as the observed enzyme activity in the absence of inhibitor; negative control is defined as the observed enzyme activity in the presence of a known inhibitor, bpV(phen) at 10 μM.[31] μ denotes the average activity and σ denotes the standard deviation. A Z'-factor > 0.7 is considered adequate for a high-throughput screening assay.[95] An FDA-approved small molecule library consisted of 1018

compounds from Selleckchem (Houston, TX) was screened against SHP2 WT at 10 μM small molecule concentration. Each compound was incubated with the enzyme for 30 min before treating the solution to the pY-peptide presenting SAMs. The screening result was plotted, and 27 compounds were found to have > 20% inhibition (Figure 2-4B). These 27 hits were verified at 100 μM (Fig. 2-4C) and 10 μM (Fig. 2-4D) to identify the 4 most potent SHP2 inhibitors: adapalene, verteporfin, closantel sodium and thonzonium bromide.

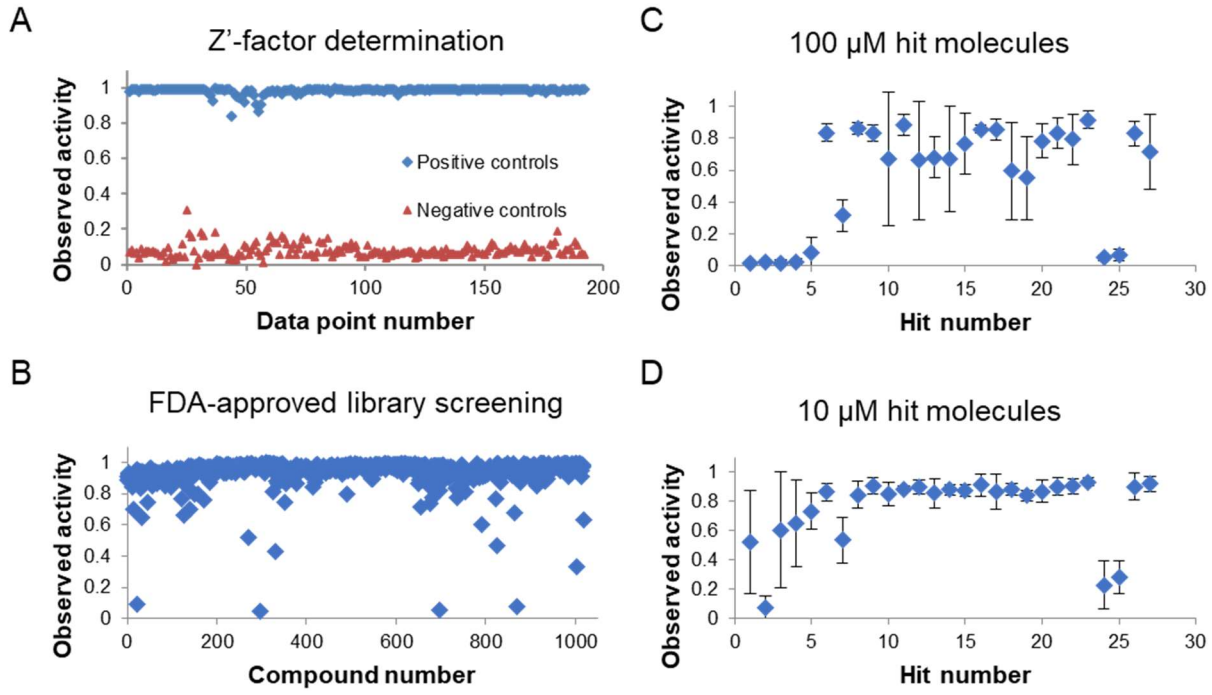



**Figure 2-4 Using SAMDI-MS for high-throughput SHP2 inhibitor screening.**

A) Using positive and negative controls to determine Z'-factor for the screening. B) SHP2 inhibitor screening results using a Selleckchem FDA approved compound library. C, D) Hits were verified at 100 μM and 10 μM.

### *2.2.4. Determination of $IC_{50}$ Using SAMDI-MS*

To determine the $IC_{50}$s of the inhibitors, the compounds were incubated at different concentrations ranging from 1 nM to 400 μM with enzymes (SHP2 WT, SHP2 E76K and SHP1) for 30 min before treating the solution to the pY-peptide presenting SAMs as previously described in section 2.2.3. The activities were measured by SAMDI-MS and converted to inhibitions, which is given by Inhibition = 1 - Activity. Inhibitions were plotted against inhibitor concentrations in log scales, and fitted with dose-response curves in SigmaPlot software, which given by the Hill equation:

$$\frac{I}{I_{max}} = \frac{1}{1 + \left(\frac{IC_{50}}{[Conc.]}\right)^{N}}$$

I: inhibition, $I_{max}$: maximum inhibition, [Conc.]: inhibitor concentration, N: Hill coefficient.

The structures of all 4 inhibitors and their IC50 plots are presented in Fig. 2-5. All exhibit inhibition to SHP2 WT and E76K with $IC_{50}$s at μM level. Hill coefficient describes the binding cooperativity between the inhibitor and the enzyme. Ideally, a noncooperative inhibitor that bind to the enzyme with a 1:1 ratio would have a Hill coefficient of 1; N > 1 indicates positively cooperative binding, which could often be non-specific.[96] For example, from the structure of thonzonium bromide and its high Hill coefficient, we can reason that its long alkyl chain might interact with the enzyme through non-specific hydrophobic interaction, making it less likely to be a potent drug. SHP2 inhibitors are often tested for SHP1 inhibition to demonstrate its specificity since SHP2 and SHP1 share 59 % sequence similarity.[97] We identified that adapalene had potent inhibition to both SHP2 WT and E76K, exhibited some specificity for SHP2 over

SHP1 (6.7-fold $IC_{50}$ ratio), and its structure was relatively simple. Therefore, we decided to move forward with adapalene for the following studies.

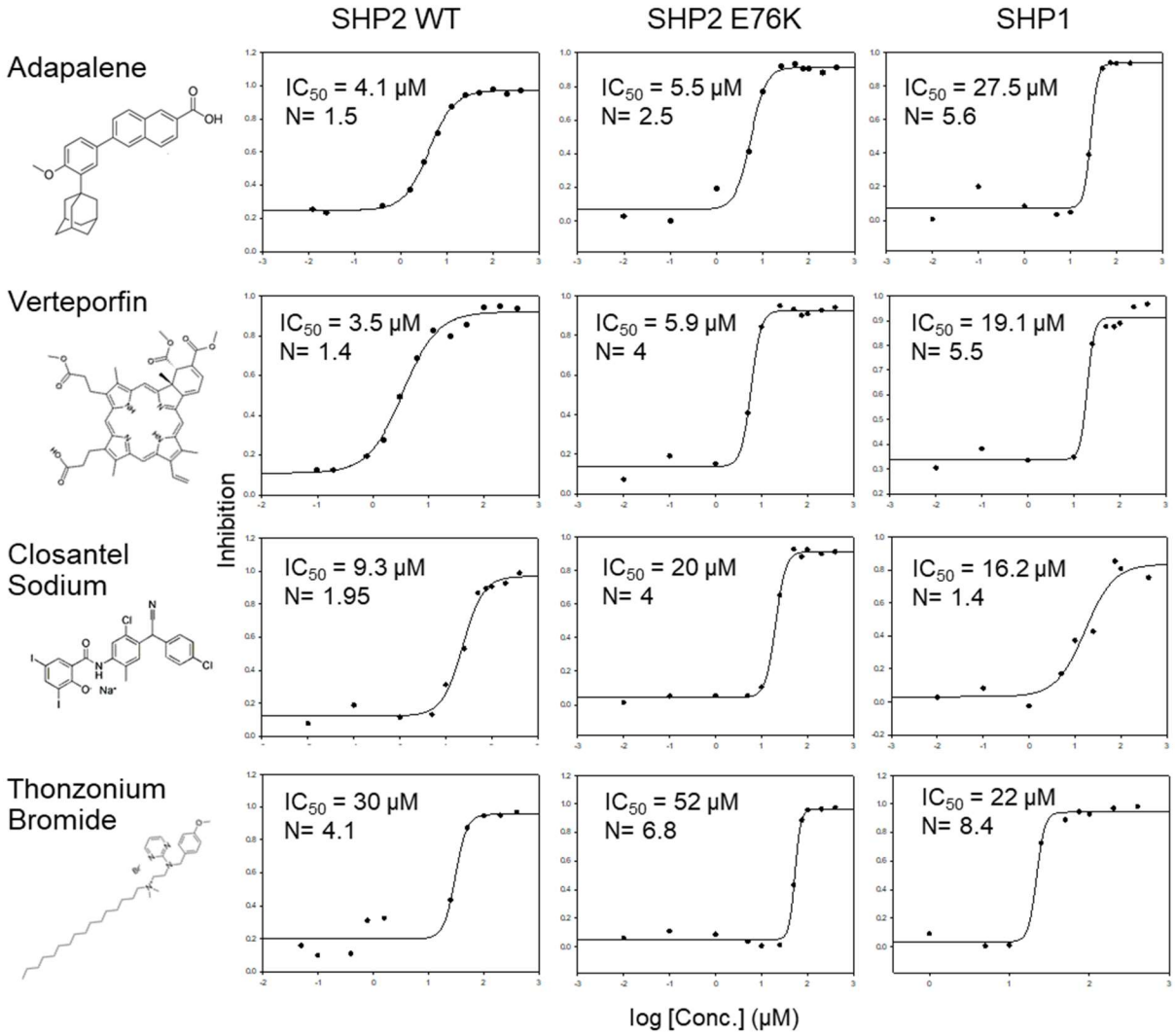



**Figure 2-5 $IC_{50}$ measurements of 4 SHP2 inhibitors.**

The $IC_{50}$s of 4 identified SHP2 inhibitors against WT, E76K SHP2 and SHP1 were measured by SAMDI-MS. The data were fitted with dose-response curves in SigmaPlot software. N denotes Hill coefficient.

### *2.2.5. Molecular Modeling and Modification of Adapalene*

*This work was performed in collaboration with Dr. Rama K Mishra.

We next used molecular modeling to study the interaction between the inhibitor adapalene and SHP2. The molecule was docked in the SHP2 catalytic site in the PTP-domain (a.a. 251-313; 325-525; 4PVG.pdb) and allosteric site in the cleft between N-SH2 and PTP domains (full length; 5EHP.pdb) with Induced Fit Docking (IFD). The docking in the catalytic site predicted a pKd (dissociation constant) of 6.18 with a penalty score -1.28; and pKd in the allosteric site of 8.48 with a penalty score -2.25. The calculated Kd in the catalytic site is near μM in magnitude ($10^{-6.18}$ M = 6.6 x $10^{-7}$ M = 0.66 μM), closer to the measured IC50 of adapalene (4.1 μM). Therefore, the catalytic site docking is a better model of its binding to SHP2 than the allosteric site docking (Fig. 2-6A).

The docking pose of adapalene in the PTP-domain near the catalytic site (C459) reveals specific electrostatic interaction between the carboxylic acid and the basic residue R465 in the pocket (Fig 2-6B). There is also a potential hydrogen bonding with the protein backbone. However, the adamantyl group is flanking outside of the hydrophilic pocket. We wondered if it plays any particular role in the binding. Therefore, we used 6-(4-methoxyphenyl)-2-naphthoic acid **2**, an adapalene precursor that did not have the adamantyl moiety, to investigate its interaction with SHP2.[98] At concentrations ranging from 0.01 μM to 400 μM, the molecule was incubated with the enzyme before treated to the pY-peptide presenting SAMs for SAMDI-MS PTP assay. We observed no concentration dependency in the SHP2 activity, in contrast with the positive control adapalene where enzyme activity increased as the inhibitor concentration was decreased (Fig. 2-6C). Without the adamantyl group, the inhibitor completely lost its potency.

This result suggests that the adamantyl group is critical in the inhibition despite that it is flanking outside of the catalytic pocket. It may potentially interact with other domains that keep SHP2 structure in closed form, which limits the PTP activity.

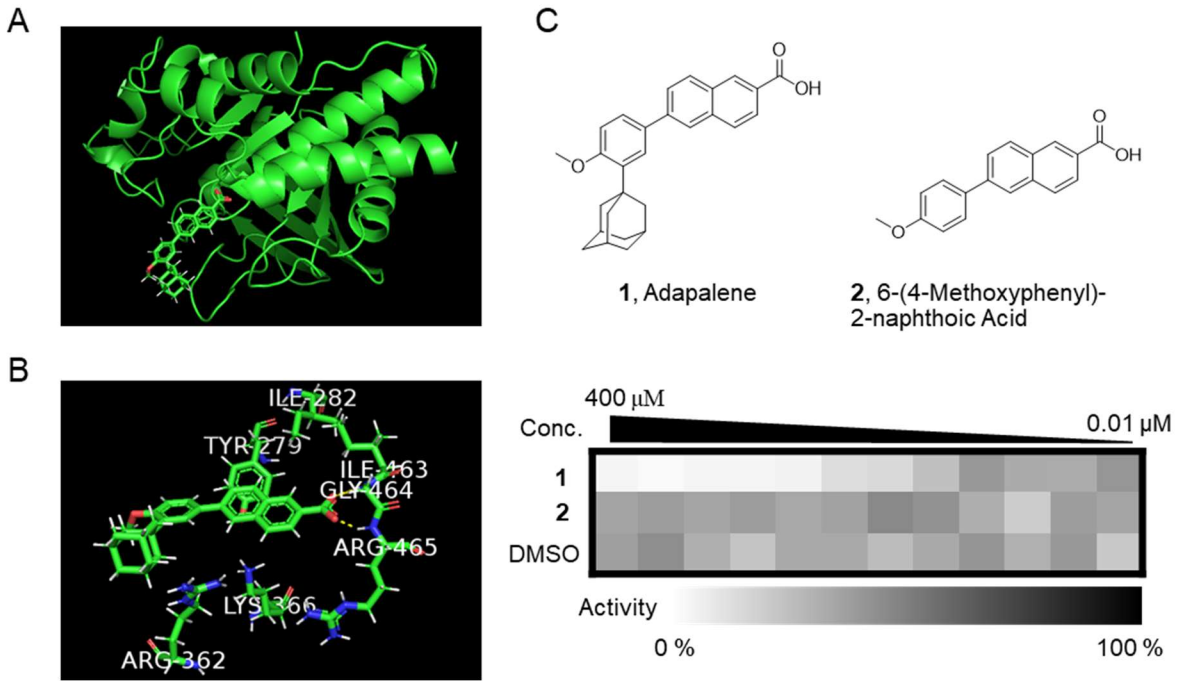



**Figure 2-6 Molecular docking of adapalene and SHP2 PTP-domain.**

A) Docking adapalene to the SHP2 PTP-domain found a calculated pKd of 6.18. B) The carboxylic acid in adapalene contributes to main inhibitory interactions while the function of adamantyl is unclear. C) Removal of the adamantyl group completely diminish the molecule's capacity to inhibit SHP2. Activities were measured my SAMDI-MS in the presence of each molecule ranging from 0.01 µM to 400 µM. DMSO is a negative control.

## *2.3. Conclusion*

In this chapter, we describe a SAMDI-MS-based PTP assay to measure SHP2 activity. It is label free, substrate specific, and capable of measuring phosphatase activity in complex samples such as cell lysates. We also demonstrate its high-throughput measurement capacity by conducting a SHP2 inhibitor screening. The assay is reproducible with an excellent screening Z'-factor of 0.86. We identify several hits from the FDA-approved compound library. Their $IC_{50}$s can be determined with SAMDI-MS as the assay is quantitative. The molecular modeling of adapalene docking in the SHP2 catalytic site reveals that its carboxylic acid interacts with the basic residues. While the adamantyl is flanking outside of the pocket, its removal diminishes the inhibition of SHP2, suggesting that the adamantyl is responsible for crucial protein-drug interaction. This work provides important techniques to drive the development SHP2 inhibitors as well as new insights in understanding SHP2 active site.

## *2.4. Methods*

**General**. Laboratory chemicals and reagents were purchased from MilliporeSigma and used without additional purification unless specified. Peptide synthesis reagents, including Fmoc amino acids and Rink-amide resin, were purchased from Anaspec. SAMDI-MS was performed on a 5800 MALDI TOF/TOF mass spectrometer (AbSciex) using either manual or automated protocols. A detailed protocol of monolayer plate preparation, peptide synthesis and phosphatase assay can be found in a published method paper.[99]

**Plasmids.** SHP2 WT/pGEX1, E76K/pcDNA1 (constitutively active) and C463S/pcDNA1 (global deactivated) plasmids were generous gifts from Dr. Elizabeth A Eklund. The full length SHP2 genes were amplified using polymerase chain reaction (PCR) with the primers 5'-CTAGCTAGCCACCATCACCATCACCATCACCATATGACATCGCGGAGATGGTT-3' (forward) and 5'-CCCAAGCTTCTAAGTGCCTAGCCCTTCCA-3' (reverse). All genes were subcloned to pET-21d(+) vectors using NheI and HindIII restriction enzyme sites and verified by Sanger sequencing.

**Protein Expression and Purification.** To express SHP2 WT, E76K and C463S, the plasmids were transformed to BL21(DE3) *E. coli* (New England Biolabs) using heatshock method. 2xYT media (5 mL) with carbenicillin was seeded with SHP2 BL21(DE3) *E. Coli* and allowed to grow overnight at 30 °C while shaking at 240 rpm. The next morning, the cultures were added to 2xYT media (500 mL) supplemented with carbenicillin and grown at 30 °C while shaking until the OD600 reached 0.5. The cultures were induced with 1 mM IPTG for 3 h under shaking at 30 °C. The bacteria were pelleted by centrifugation and lysed in buffer (100 mM Tris

pH 7.5, 50 mM NaCl, 5 mM TCEP, 10 % glycerol (v/v), 0.1% Triton X-100, 40 mL) containing one cOmplete-Mini protease inhibitor tablet. The overexpressed SHP2s contained an N'-terminal His-tags and were purified using HisPur™ cobalt resin (ThermoFisher). Fractions were eluted with 300 mM imidazole in buffer (100 mM Tris pH 7.5, 50 mM NaCl, 5 mM TCEP), combined, concentrated using Amicon 50 kDa cutoff centrifugal filter units and stored at -80 °C in 50% glycerol.

**Peptide Synthesis.** Cysteine-containing peptides were synthesized using solid-phase peptide synthesis on Fmoc-Rink Amide MBHA resins (Anaspec). The resin (100 mg) was swelled in dimethylformamide (DMF) for 30 min, deprotected with 20% piperidine in DMF for 20 min and rinsed with DMF by vacuum filtration. Amino acids were coupled to the resin with PyBop and N-methylmorpholine (NMM) in DMF at a 4:4:8 molar excess, respectively, to the resin for 20 min and washed with DMF. The deprotection, wash, coupling, and wash steps were repeated until the last amino acid was coupled and deprotected. The N-terminus was acetylated with 10% acetic anhydride in DMF for 60 minutes. The resin was washed with DMF and dichloromethane and dried in a vacuum desiccator. The peptides were cleaved off the resin with a solution of 95% trifluoroacetic acid (TFA), 2.5 % triethylsilane, and 2.5 % water overnight, precipitated in ice-cold ethyl ether, centrifuged, and dried under vacuum. The peptides were purified with HPLC, lyophilized, and resuspended in 0.1 % TFA in water to a peptide concentration of 1 mM.

**Peptide Immobilization.** Steel array plates evaporated with 384 gold spots (with a diameter of 3.0 mm) were soaked at 4 ºC in a solution of 1 mM total disulfide with 0.8 mM tri(ethylene glycol)-terminated $C_{11}$-alkane disulfide and 0.2 mM $C_{11}$-alkane disulfide with one terminal tri(ethylene glycol) and one terminal maleimide in ethanol overnight to allow assembly of the

monolayer (10% maleimide coverage). A pY-peptide (Ac-TRDIpYETC) was diluted to a final concentration of 50 μM with 100 mM Tris buffer (pH 7.5) and 0.5 mM TCEP. Peptide solution (2 μL) was transferred to each gold spot on the monolayer-presenting plate and incubated at 37 ºC in a humidified chamber for 1 h to allow peptide immobilization.

**SAMDI-MS PTP Assay.** SHP2s were diluted in PTP buffer (100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and 2 μL was applied to each gold spot on a pY-peptide (Ac-TRDIpYETC) presenting SAMDI plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated for 30 min at 37 ºC in a humidified chamber. After the reactions were complete, the plate was rinsed with water and ethanol, treated with 1 μL matrix (10 mg/mL THAP, 5 mg/mL ammonium citrate dibasic in 50% acetonitrile, 50% water and 0.1 % phosphoric acid) to each spot and dried in air for 20 min. The spots were analyzed by MALDI-TOF MS to obtain a mass spectrum for each reaction. Enzymatic activities were quantified by measuring the areas under the curve (AUCs) for the dephosphorylate product peak and the substrate peak and determining the activity (%) = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %.

**SHP2 Inhibitor Screening.** FDA-approved compound library (Selleckchem) was diluted to 100 μM in DMSO and transferred to 384-well plates. Using a TECAN robot, SHP2 WT (9 μL, 222 nM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) was mix with each compound solution in the library (1 μL, 100 μM) to a final concentration of 200 nM SHP2 WT and 10 μM molecule and incubated at room temperature for 30 min. The solution was treated to a 384 SAMDI plate presenting pY-peptide (Ac-TRDIpYETC). The reactions were incubated for 30 min at 37 ºC in a humidified chamber, rinsed and analyzed by SAMDI-MS as described above.

**$IC_{50}$ measurement.** SHP2 inhibitors were diluted to concentrations ranging from 1 nM to 400 μL in a SHP2 WT solution (200 nM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and incubated at room temperature for 30 min. The solution was treated to a 384 SAMDI plate presenting pY-peptide (Ac-TRDIpYETC). The reactions were incubated for 30 min at 37 ºC in a humidified chamber, rinsed and analyzed by SAMDI-MS as described above. The activities were converted to inhibitions, which is given by Inhibition = 1 - Activity. Inhibitions were plotted against inhibitor concentrations in log scales and fitted with dose-response curves (Hill equation) in SigmaPlot software.

**Molecular Modeling.** Induced Fit Docking (IFD) was conducted by Dr. Rama K Mishra in Northwestern Center for Molecular Innovation and Drug Discovery (CMIDD). Detail protocol is available through Northwestern ChemCore: https://chemcore.northwestern.edu/.

# Chapter 3. Combining SAMDI-MS and Peptide Arrays to Profile SHP2 Disease Mutants

## *3.1. Introduction*

Somatic mutations in PTPN11, the gene encoding the protein tyrosine phosphatase SHP2 contribute to leukemogenesis including juvenile myelomonocytic leukemia (JMML), acute myeloid leukemia (AML) and Acute Lymphocytic Leukemia (ALL), whereas germline mutations cause Noonan syndrome (NS) and the clinically related LEOPARD syndrome (LS).[100] Missense mutations in SHP2 are also associated with lung cancer, colon cancer and neuroblastoma.[24]

It is still unclear how different SHP2 mutants develop into a wide spectrum of disease. Earlier studies focus on a "activity-centric" model where cancer-associated mutations cause constitutive activation and are more active than those associated with NS.[23] However, as more mutants are discovered and linked to more diverse diseases, the model becomes insufficient to explain the behaviors of all pathogenic SHP2 mutations, especially to those mutations map away from the N-SH2/PTP domain interface. Structural and mathematical modeling analysis shows that these mutations can not only affect basal activation, but also switch SH2 domain-phosphopeptide affinity and substrate specificity to varying degrees.[100]

Assays of phosphatase activity are quite challenging, and largely not well-suited to the direct determination of phosphatase specificity. Our laboratory has developed SAMDI-MS, a label-free, high-throughput analytical method for measuring enzyme activities, including phosphatase activity demonstrated in chapter 2. We have combined SAMDI-MS and peptide arrays to profile enzyme substrate specificities. These examples include profiling glycotransferases, bacterial outer-membrane protease OmpT, as well as lysine deacetylase (KDAC) specificities.[72, 75-76] Additionally, SAMDI-MS and peptide arrays are also capable of detecting changes in substrate

specificity, as shown in the study observing distinct OmpT wildtype and mutant activity profiles.[75]

In this chapter, we describe the use of SAMDI-MS and a peptide array based on a sequence previously used in earlier studies of SHP2 activity: Ac-TRDXpYZTC-$NH_2$ where the X and Z positions are variable.[94] We profiled WT and 11 full length SHP2 mutants as well as 19 $\triangle$104 SHP2 mutants where the N-SH2 domains were removed. We also compared their basal PTP activities. Our results reveal that some mutants particularly prefer aromatic residues at X position. These data are consistent with previous prediction of the mutations' effect on SHP2 function, and we discuss their implications in pathology.

## *3.2. Results and Discussion*

### *3.2.1. Profiling SHP2 Activities With SAMDI-MS and Peptide Arrays*

We synthesized a pY-peptide array using standard protocols with FMOC-protected amino acids. The array contained 361 unique peptides with the sequence Ac-TRDXpYZTC-$NH_2$. The X and Z positions comprise each of the 19 natural amino acids except for cysteine. To an array plate having 384 gold spots arranged in the geometry of the common microwell plate, we modified each gold spot with a self-assembled monolayer presenting maleimide groups at a density of 10% against a background of tri(ethylene glycol) groups as described previously.[99] The peptides were diluted in Tris buffer (pH 7.5) and transferred to the monolayer array plate using a robotic liquid handler, where each peptide underwent immobilization to the monolayer in its spot via conjugate addition of cysteine thiol to the maleimide group (Fig. 3-1A).

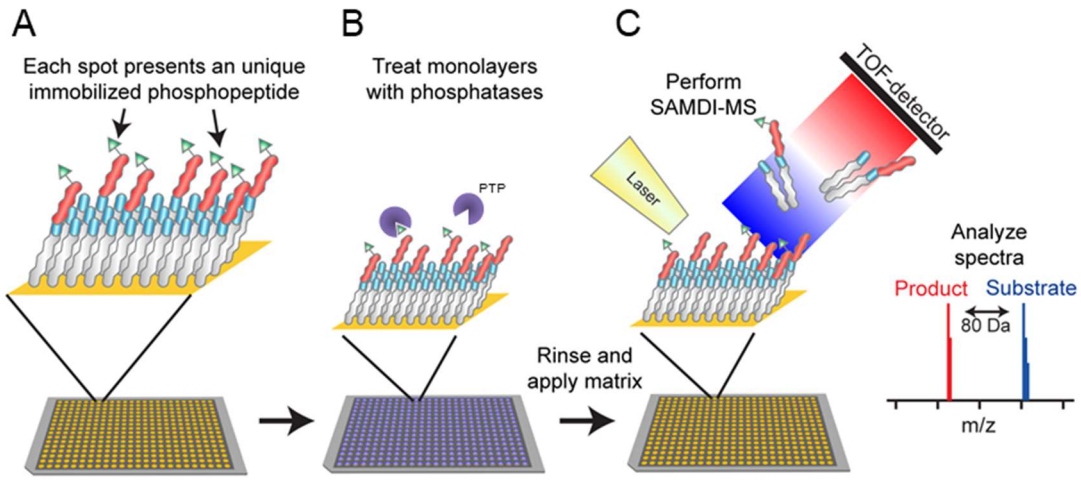



**Figure 3-1 Profiling PTP activities using SAMDI-MS and peptide array.**
A) Peptides in a pY-peptide library are immobilized on SAMs. Each spot presents one unique peptide. B) The surface is treated with a PTP solution. C) Each spot is analyzed with SAMDI-MS to determine the PTP activity on the peptide.

Then, each spot on the array was treated with a PTP solution and the reaction was carried out in a humidified chamber (Fig. 3-1B). Finally, the plate was rinsed, treated with THAP matrix and analyzed with SAMDI-MS (Fig. 3-1C). Conversion of phosphopeptide to its product was characterized by integration of the corresponding peaks, where Activity = Area Under the Curve $(AUC)_{product}$ / $(AUC_{substrate} + AUC_{product})$ x 100 %. Activities for each peptide sequence were represented in a 19 x 19 heatmap where each column defined the amino acid in the X (-1) position and each row defined the amino acid in the Z position (+1). Activity was represented by a color scale with white corresponding to 0 % and black as 100 % (Fig. 3-2).

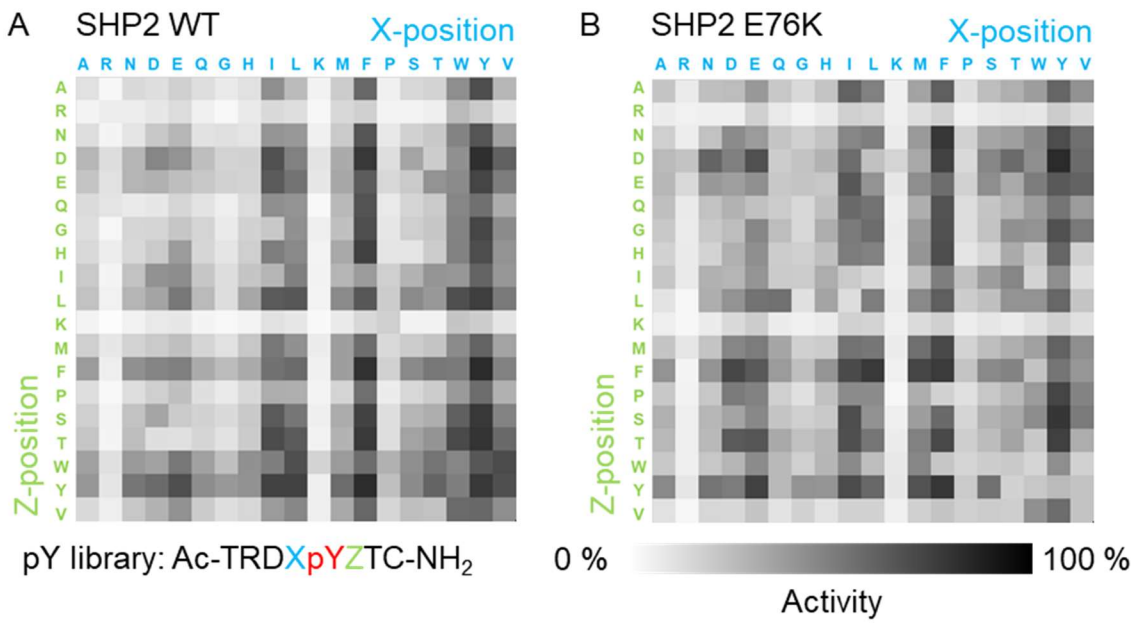



**Figure 3-2 Profiling SHP2 WT and E76K activities.**

A) SHP2 WT activity profile. B) SHP2 E76K activity profile.

### *3.2.2. Comparing SHP2 WT and Constitutively Active Mutant E76K*

The PTP activity profile heatmaps of WT and E76K SHP2 (Fig. 3-2) show very similar patterns: when X and Z positions are positively charged basic residues (R and K), the enzyme activities are restricted; as the local charges are neutralized by acidic D and E residues, the enzyme activities are restored. Negatively charged D and E residues promote both WT and E76K SHP2 activities. The heatmaps also show that hydrophobic residues (I, L, V) and aromatic residues (F, Y, W) are particularly favored by the enzymes at X position and slightly preferred at Z position. Proline (P) is disfavored at both positions. No significant difference is observed in the two maps; therefore, we conclude that WT and E76K SHP2 have the same substrate selectivity.

This result is consistent with the mechanism of action in the SHP2 E76K mutation, which impairs the inhibitory interaction between the N-SH2 domain and PTP domain, switching the SHP2 closed conformation to its open form, leading to increased basal PTP activity and increased affinity of the SH2 domains for pY ligands. Since E76K mutation is located in the N-SH2 domain and does not participate in the binding of the PTP domain and its substrate, it is less likely to alter the substrate selectivity of the SHP2 phosphatase.

### *3.2.3. Exploring Activities and Substrate Selectivities of SHP2 Mutants*

Disease-related SHP2 mutations spread across the entire protein. Their substrate specificities remain largely unexplored. To study the mutation-function relationship, we selected 11 SHP2 mutants: 2 JMML mutants (D61G, E76K), 8 NS mutants (E76D, Q79R, Q256R, N308D, T468M, R501K, G503E) and 2 related to both disease (G503R, Q506P).[101] The vectors containing the mutants were generate by site-directed mutagenesis, transformed to *E coli.*, and induced with IPTG for full-length SHP2 expression (Fig. 3-3A). The lysate concentrations were normalized with Bradford assays and the SHP2 activities in the lysates were compared using a general peptide substrate (Ac-IpYERC-$NH_2$) measured by SAMDI-MS (Fig. 3-3B). The control experiment showed that without IPTG induction, *E coli.* had minimal background phosphatase activity. Consistent with the "activity-centric" model, we observed that JMML-associated N-SH2 domain mutations caused constitutive activation and were more active than PTP-domain mutations associated with NS.[23] Although G503 residue was in the PTP-domain, its mutation could disrupt the key interaction between E76-S502 and lead to constitutive activation. Therefore, we observed elevated G503E and G503R activities. However, the model could not explain the disease differentiation of the mutations at 503 and 506 positions because they were shared by both JMML and NS.

We then profiled all the mutants with the peptide library Ac-TRDXpYZTC-$NH_2$ and SAMDI-MS as previously described (Fig. 3-3C). To our surprise, we found that all SHP2 mutant profiles remained largely the same, suggesting no specificity changes. This appeared inconsistent with the structural predictions and the biological outcomes of these mutants. We reasoned that the N-SH2 domain in the full-length SHP2 protein might hinder the PTP-domain and the nuance

changes in specificity its mutations caused. Therefore, we next sought to modify the construct to remove the N-SH2 domain for further studies.

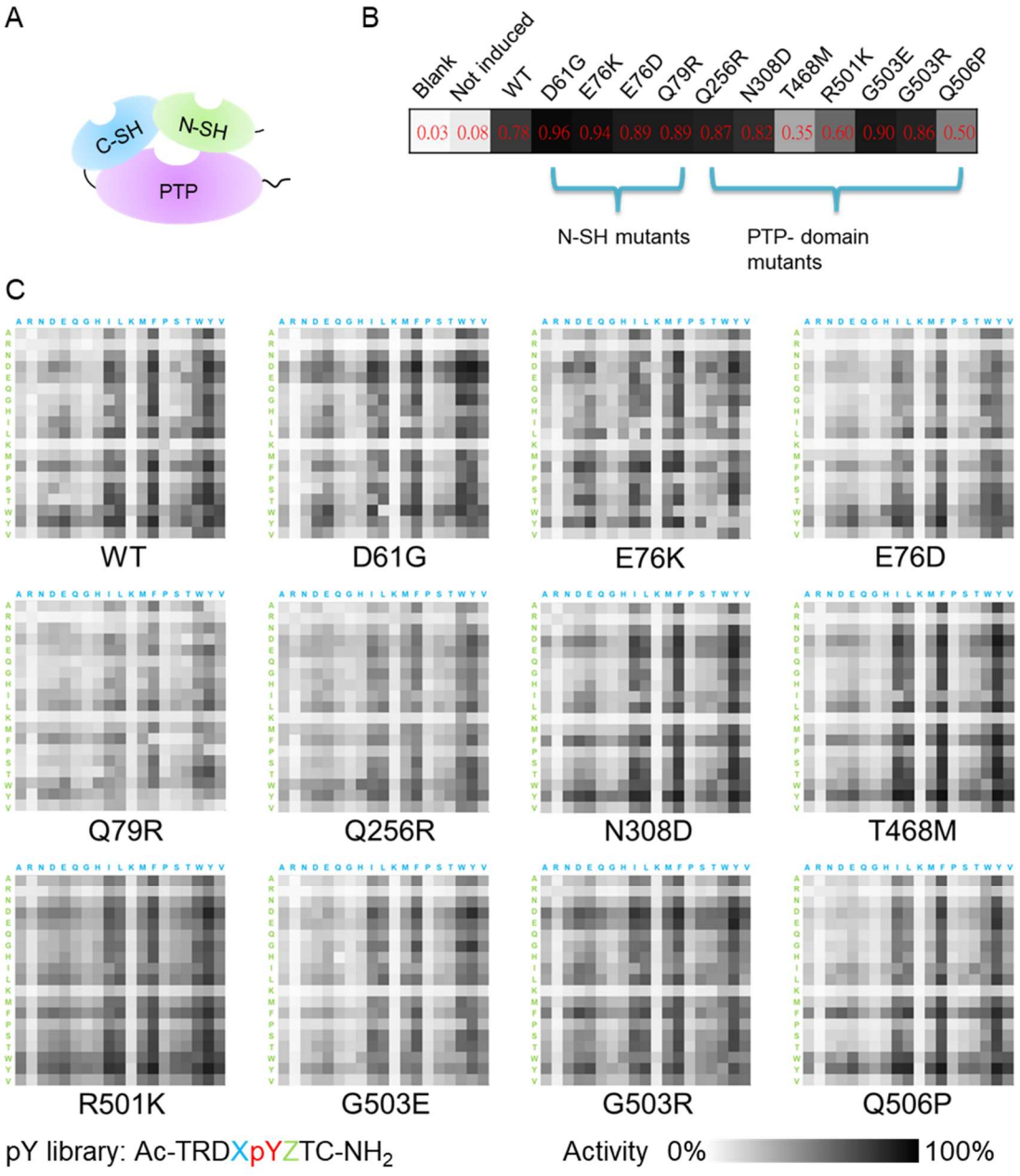



**Figure 3-3 Profiling full-length SHP2 mutant activities.**

A) Full-length SHP2 structure. B) Activities of SHP2 and mutants on peptide Ac-IpYERC-$NH_2$ measured by SAMDI-MS. C) Profiles of SHP2 WT and 11 mutants.

### *3.2.4. Using SHP2Δ104 Constructs to Exam Changes in Substrate Selectivity*

*This work was performed in collaboration with Dr. Simone Martinelli.

A plasmid encoding the amino acid sequence corresponding to the isolated PTP domain preceded by the C-SH2 domain (SHP2Δ104, residue 105-541) was used to generated 19 disease related SHP2 mutants via site-directed mutagenesis. The identities of the mutants and their predicted effect on SHP2 function are summarized in Table 3-1.[100] The plasmids were transformed to *E coli.(DE3) Rosetta 2* competent cells for expression. The poly-his-tagged SHP2Δ104 proteins were purified using cobalt resins.

**Table 3-1 SHP2Δ104 mutants and their predicted effect on SHP2 function.**

| # | Mutant | Disease | Mutation Group | Predicted Effect on SHP2 Function[100] |
|---|---|---|---|---|
| **1** | WT | N/A | N/A | N/A |
| **2** | E110A | NS | VI | SH2 orientation or mobility |
| **3** | Q256R | NS | III | A/I switching and specificity[a] |
| **4** | Q256K | NS | III | A/I switching and specificity |
| **5** | L261H | NS | IV | A/I switching and/or catalysis |
| **6** | L261F | NS | IV | A/I switching and/or catalysis |
| **7** | G268S | NS | IV | A/I switching and/or catalysis |
| **8** | F285L | NS | IV | A/I switching and/or catalysis |
| **9** | F285S | NS | IV | A/I switching and/or catalysis |
| **10** | N308D | NS | IV | A/I switching and/or catalysis |
| **11** | N308S | NS | IV | A/I switching and/or catalysis |
| **12** | P491S | NS | IV | A/I switching and/or catalysis |
| **13** | S502L | NS/Leukemia | III | A/I switching and specificity |
| **14** | S502P | NS/Leukemia | III | A/I switching and specificity |
| **15** | S502T | NS/Leukemia | III | A/I switching and specificity |
| **16** | G503A | NS/Leukemia | III | A/I switching and specificity |
| **17** | G503E | NS/Leukemia | III | A/I switching and specificity |

| | | | | |
|---|---|---|---|---|
| **18** | G503R | NS/Leukemia | III | A/I switching and specificity |
| **19** | N504V | NS | IV | A/I switching and specificity |
| **20** | E507K | Leukemia | II | A/I switching and catalysis |

[a]A/I = Active/inactive conformation

We next profiled all SHP2Δ104 mutants with the peptide library Ac-TRDXpYZTC-$NH_2$ and SAMDI-MS as previously described (Fig. 3-4A and B). Their activities were also compared at the same concentration using a control peptide Ac-IpYERC-$NH_2$ (Fig. 3-4C). The activities indicated a disruptive effect of all mutations at 502, 503, 504 and 507 residues. From the profiles, we noted some different behavior of these mutants regarding aromatic residues (F, Y, and W), which might suggest changes in substrate specificity. To quantitate the influence that each residue has on activity of the substrate for each mutant, we conducted the following analysis.

For each amino acid in the X position, we determined the average activity of all peptides in the row corresponding to that residue in the heatmap. We also determined the average activity for all peptides in the array (global average, GA). The ratio of the difference (Δ) of the average activity for the particular residue ($AA^x$) and the global average (GA) was then determine according to the equation $\Delta = (AA^x - GA)/GA \times 100$ %. We similarly repeated this analysis for residues in the Z position ($AA^z$). The numerical values of these ratios for all 38 rows and columns (19 amino acids at either the X or Z position) for each of the SHP2Δ104 and 19 mutants are shown in Fig. 3-5. This analysis gives insight into the sequence determinants of activity.

First, the matrices show a consistent inhibitory effect (around -60 %) of basic residues at both X and Z positions across all mutants. This trend may be explained by electrostatic interactions in the SHP2 pY binding pocket. SHP2 catalytic site has the signature PTP-family motif $HC(X)^5R$

where a cysteine residue acts as the nucleophile and the conserved arginine residue folds back toward the phosphate-binding pocket to assist in substrate binding and catalysis.[15] Therefore, the pocket is highly positively charged to accommodate the negatively charged phosphate. This gives explanation to the consistent disfavor in basic residues and some preference in acidic residues on either side of pY.

Second, we find that aromatic (F, Y, W) and most hydrophobic (V, I, L, excluding A) residues are usually favored by SHP2. A reported crystal structure provides some structural insight into the preference for aromatic residues.[102] In SHP2, two aromatic groups (Y279, H426) occupy either side of the PTP active site, providing π-π interactions to pY and its neighboring residues on the substrate. However, the matrices indicate that there is no strong correlation between preferences in the X and Z positions. For example, we find that preference for aromatic is more critical in the X position.

Lastly, we report evidence changes in specificity of some mutants. While most amino acids have a consistent Δ value across WT and mutants in Fig. 3-5, we notice some changes regarding aromatic residues at X position ($F^X$, $Y^X$, $W^X$). E110A, F285S, P491S, S502L and S502P are all significantly more prefer aromatic residues than WT. Also, G503A, G503E and G503R exhibit a different degree of preference in aromatic residue. These observations suggest that mutations proximal to the active site pocket might alter the way SHP2 interact with its substrates, consistent with previous predictions.[100] In addition, changes in specificity give us an alternative way to explain the complexity in development and differentiation of disease, particularly to those mutations surrounding residue 502 and 503, associated with both NS and leukemia.

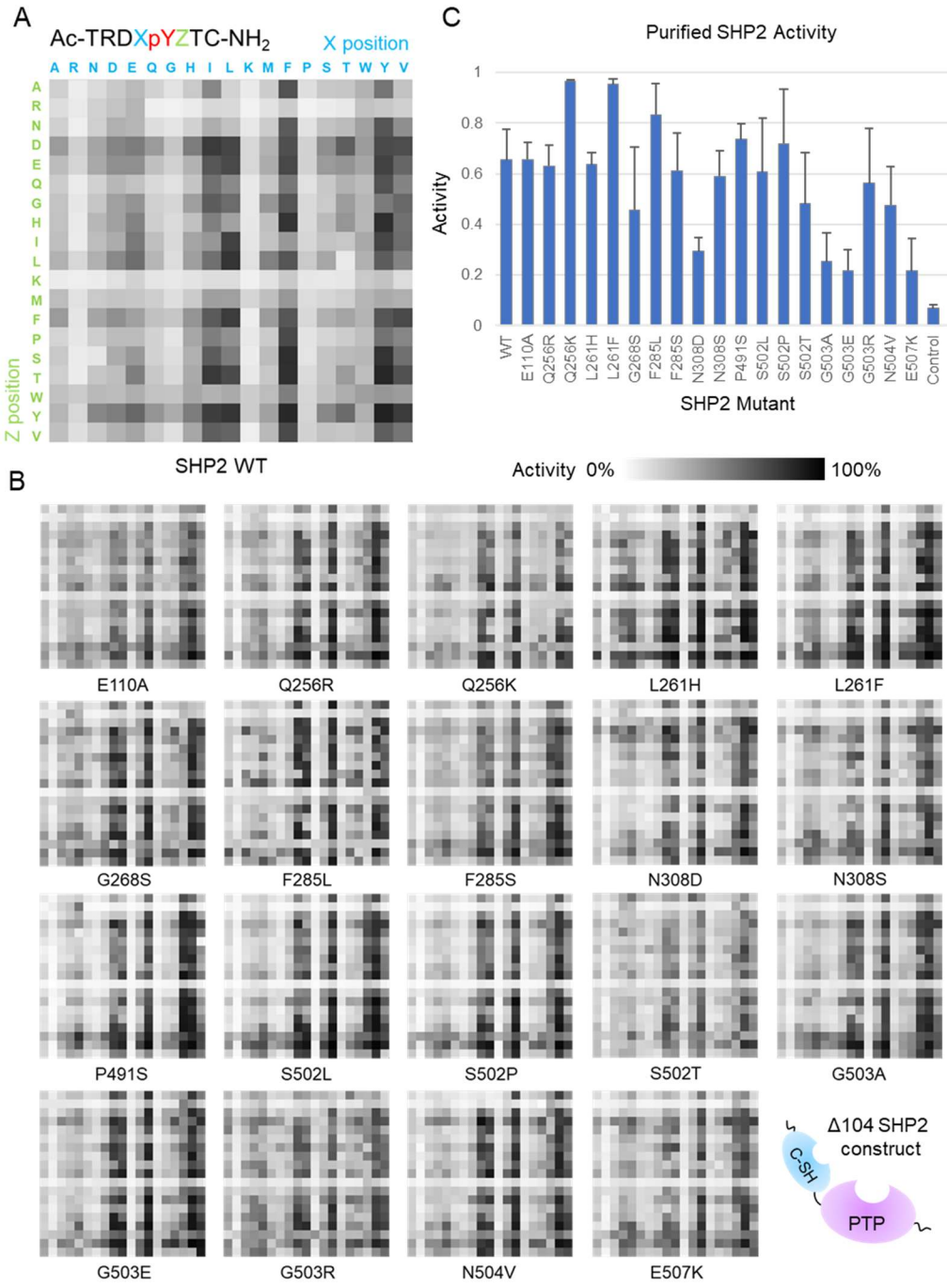



**Figure 3-4 Profiling SHP2Δ104 mutant activities.**

A) Profile of SHP2Δ104 WT. B) Profiles of 19 disease associated SHP2Δ104 mutants. C) Activities of purified SHP2Δ104 and mutants on peptide Ac-IpYERC-$NH_2$ measured by SAMDI-MS.

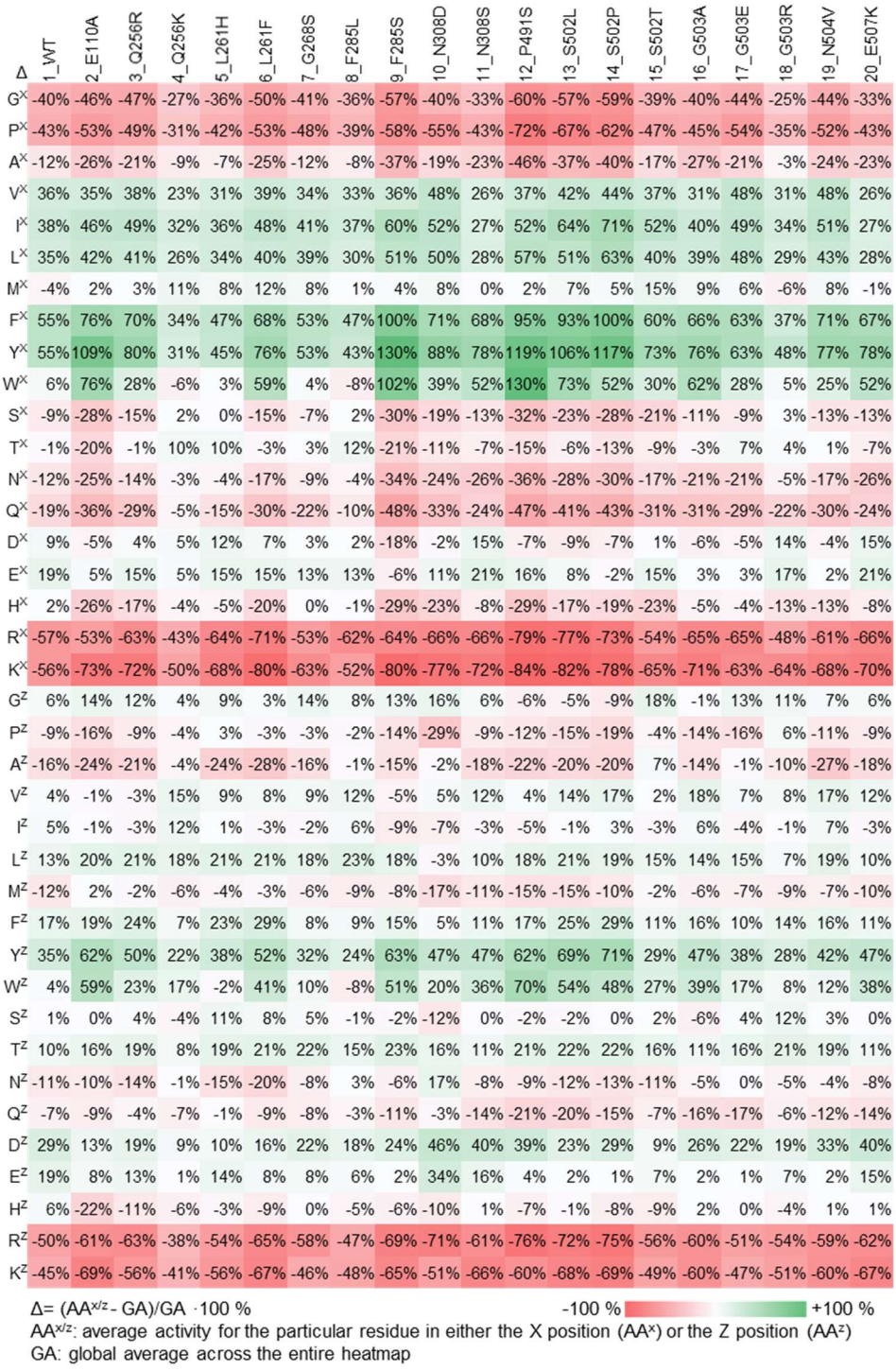



**Figure 3-5 Analysis in amino acid specificity of SHP2Δ104 and mutants.**

Heatmap illustrating the average role of an amino acid residue in either the X ($AA^X$) or Z ($AA^Z$) position has on the activity of the substrate for each phosphatase enzyme.

## *3.3. Conclusion*

This work addresses the long-standing challenges in characterizing SHP2 mutant specificity. Our use of peptide arrays and SAMDI-MS is the first approach to directly determine the specificity profiles of SHP2 and its mutants using hundreds of peptide substrates. By applying this method to WT SHP2, 11 full-length mutants and 19 Δ104 mutants, we observed conserved trends in sequence-dependent activity that are important for SHP2 to accommodate its substrate in the active site. We also report some changes in profiles caused by disease related mutations previously predicted to have effect on SHP2 specificity. These mutants exhibit a wide degree of preference to aromatic residues. We expect these data will be valuable in many more pathological contexts, providing now insights on the development and differentiation of Noonan syndrome, LEOPARD syndrome, leukemia, and other SHP2-related cancers. Equally important, we believe this approach will lead to more complete understandings of the regulatory roles of phosphatases.

## *3.4. Methods*

**General**. Laboratory chemicals and reagents were purchased from MilliporeSigma and used without additional purification unless specified. Peptide synthesis reagents, including Fmoc amino acids and Rink-amide resin, were purchased from Anaspec. SAMDI-MS was performed on a 5800 MALDI TOF/TOF mass spectrometer (AbSciex) using either manual or automated protocols. A detailed protocol of monolayer plate preparation, peptide synthesis and phosphatase assay can be found in a published method paper.[99]

**Plasmids.** SHP2 WT/pET-21d(+), E76K/pET-21d(+) were cloned as previously described in chapter 2. Other SHP2 full-length mutants were generated using the primers listed in Table 3-2 and site-directed mutagenesis kits (New England Biolabs) following the manufacturer's protocol. All plasmids were verified by Sanger sequencing. SHP2Δ104/pET-26b and mutants were cloned by Dr. Simone Martinelli.

**Table 3-2 Primer list for full length SHP2 mutants.**

| Mutant | Forward Primer (5' to 3') | Reverse Primer (5' to 3') |
|---|---|---|
| **D61G** | AACACTGGGGGCTACTATGAC | CTGAATCTTGATGTGGGTAAC |
| **E76D** | CTTTGGCTGACCTGGTTCAGTATTAC | TGGCAAACTTCTCCCCAC |
| **Q79R** | TGAACTGGTTAGGTATTACATGGAAC | GCCAAAGTGGCAAACTTC |
| **Q256R** | GACGCTCCAGAGACAGGAATGC | TCAAACTCTTCCCAAAAGC |
| **N308D** | CATTAATGCAGACATCATCATGC | TAATCAGAAACAGGCTCATTG |
| **T472M** | CGGACAGGAATGTTCATTGTGATTGACATCCTTATTGACATCATTCGAG | GCCAATCCCAGCGCTGCA |
| **R505K** | CGGTCCCAGAAGTCGGGGATG | CACCATCTGAATGGTTTTAGGAACG |
| **G507E** | CAGAGGTCGGAGATGGTCCAG | GGACCGCACCATCTGAAT |

| | | |
|---|---|---|
| **G507R** | CCAGAGGTCGAGGATGGTCCA | GACCGCACCATCTGAATG |
| **Q510P** | GGGATGGTCCCGACAGAAGCACAG | CGACCTCTGGGACCGCAC |

**Protein Expression and Purification.** Full length SHP2/pET-21d(+) plasmids were transformed to BL21(DE3) *E. coli* (New England Biolabs) and expressed as described in chapter 2. SHP2Δ104/pET-26b plasmids were transformed to *E coli.(DE3) Rosetta 2* competent cells (Novagen). 2xYT media (5 mL) with kanamycin (50 μg/mL) and chloramphenicol (34 μg/mL) was seeded with SHP2Δ104 *E. Coli* and allowed to grow overnight at 30 °C while shaking at 240 rpm. The next morning, the cultures were added to 2xYT media (500 mL) supplemented with kanamycin and chloramphenicol and grown at 30 °C while shaking until the OD600 reached 0.5. The cultures were induced with 1 mM IPTG for 3 h under shaking at 30 °C. The bacteria were pelleted by centrifugation and lysed in buffer (100 mM Tris pH 7.5, 50 mM NaCl, 5 mM TCEP, 10 % glycerol (v/v), 0.1% Triton X-100, 40 mL) containing one cOmplete-Mini protease inhibitor tablet. The overexpressed SHP2Δ104 were purified using HisPur™ cobalt resin (ThermoFisher). Fractions were eluted with 300 mM imidazole in buffer (100 mM Tris pH 7.5, 50 mM NaCl, 5 mM TCEP), combined, concentrated using Amicon 30 kDa cutoff centrifugal filter units and stored at -80 °C in 50% glycerol.

**Peptide Synthesis.** Solid-phase peptide synthesis was performed on Rink-amide resin (10 mg) housed in 96-well filter plates. N-terminal fluorenylmethyloxycarbonyl (Fmoc) protecting groups were deprotected with 20% piperidine in dimethylformamide (DMF) at room temperature for 20 min. The resin was filtered and rinsed 5 times with DMF on a multiscreen vacuum manifold. Amino acids were coupled to the resin with PyBop and N-methylmorpholine (NMM)

in DMF at a 4:4:8 molar excess for 30 minutes, twice. The deprotection and coupling were repeated for each residue. Following deprotection of the final residue, the N-terminus was acetylated with 10% acetic anhydride in DMF for 60 minutes. Peptides were cleaved and deprotected in 95% trifluoroacetic acid (TFA), 2.5% triethylsilane (TES), and 2.5% water for 16 h. The cleavage solution was evaporated with $N_2$ gas flow. The peptides were resuspended in 0.1% TFA in water, lyophilized and resuspended again in 0.1% TFA in water to a final concentration of 200 μM and stored at -80 ºC.

**Preparing Peptide Arrays on SAMDI Plates.** Steel array plates evaporated with 384 gold spots (with a diameter of 3.0 mm) were soaked at 4 ºC in a solution of 1 mM total disulfide with 0.8 mM tri(ethylene glycol)-terminated $C_{11}$-alkane disulfide and 0.2 mM $C_{11}$-alkane disulfide with one terminal tri(ethylene glycol) and one terminal maleimide in ethanol for 2 days to allow assembly of the monolayer (10% maleimide coverage). Peptides were diluted to a final concentration of 20 μM with 100 mM Tris buffer (pH 7.5) in a 384-well plate with 5 μL TCEP beads (Thermo Scientific) in each well. Using a TECAN robotic liquid handler, 2 μL of peptide solution from each well was transferred on to the corresponding gold spot on the monolayer-presenting plate and incubated at 37 ºC in a humidified chamber for 1 h to allow peptide immobilization.

**Profiling SHP2 activities by SAMDI-MS.** SHP2s were diluted in PTP buffer (100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and 2 μL was applied to each gold spot on the peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated at 37 ºC in a humidified chamber. Concentration and incubation time for each mutant were adjusted so that the average activity across the array is ~40%. After the reaction was complete, the plate was

rinsed with water and ethanol, treated with 1 μL matrix (10 mg/mL THAP, 5 mg/mL ammonium citrate dibasic in 50% acetonitrile, 50% water and 0.1 % phosphoric acid) to each spot and dried in air for 20 min. The spots were analyzed by MALDI-TOF MS to obtain a mass spectrum for each reaction. Enzymatic activities were quantified by measuring the areas under the curve (AUCs) for the dephosphorylate product peak and the substrate peak and determining the activity (%) = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %. Activity heatmaps were generated by Microsoft Excel.

# Chapter 4. Profiling Protein Tyrosine Phosphatase Specificity with SAMDI-MS and Peptide Arrays

*This chapter is reproduced in part with permission from Huang, C.-F.; Mrksich, M. Profiling Protein Tyrosine Phosphatase Specificity with Self-Assembled Monolayers for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry and Peptide Arrays. *ACS Comb. Sci.* **2019,** *21*, 760-769. 

## *4.1. Introduction*

While the phosphorylated states of proteins are determined by the balance of opposing kinase and phosphatase activities, the overwhelming majority of work addressed the roles of kinases and their substrates in regulating phosphorylation, and generally assumed that phosphatases serve a non-regulatory housekeeper role.[6] However, this assumption lacks justification and appears inconsistent with the roughly equal numbers of PTK and PTP in the human proteome (90 PTKs and 107 PTPs).[4-5] Further, recent work has illustrated a regulatory role for PTPs and sophisticated modes of regulation.[7-10] Their dysregulated activities have also been directly linked to disease and cancer; SHP2, for example, was identified as the first oncogenic phosphatase.[11-13] Advancing our understanding of the roles that PTPs play in signaling would benefit from determining the substrate specificities of PTPs. Here, we use peptide arrays and SAMDI-MS to profile 22 phosphatases, and we report distinct classes of substrate specificities for PTPs.

Assays of phosphatase activity are quite challenging, and largely not well-suited to the direct determination of phosphatase specificity. One approach uses bottom-up proteomics or ELISA (enzyme-linked immunosorbent assay) to observe dephosphorylation of a sample that has first been enriched in phosphoproteins.[103-104] Approaches for directly assaying enzymatic phosphatase activities frequently use generic and non-specific substrates—commonly, PNPP or DIFMUP—which release products that can be measured with absorbance or fluorescence spectroscopy. Alternatively, the malachite green assay measures the phosphate by-product, but is difficult to apply to cell lysates that have significant background levels of phosphate ions.[86-88] Hence, these assays involve tedious sample preparation, are not high-throughput, and do not easily permit the use of a large number of substrates; these limitations have hindered studies of

PTP specificity, particularly as compared to numerous studies of kinase specificity and activity.[105-108] MS-based phosphatase activity assays have gradually become popular in the field for their easier workflow and the high-throughput nature. MALDI-MS has been used in several recent examples to screen for PTP inhibitors.[109-111]

Combinatorial libraries and peptide arrays are powerful tools for studying the substrate specificities of a variety of enzymes.[65, 112] To determine the specificity of PTPs, Pei and co-workers used a combinatorial bead-based peptide library that was prepared by split-pool synthesis. They developed a colorimetric labeling method to identify non-phosphorylated tyrosine residues that result from PTP activity and could sequence those beads to identify active substrates for the PTPs.[113-114] Their work showed acidic residues are favored over basic residues in the substrates for 14 PTPs, and basic residues decrease activity.[115] However, the methods have the limitation that they can isolate and sequence a small fraction of the peptides in the pool, and therefore while they provide general trends for activity, they do not give a nuanced understanding of specificity. Cesareni and co-workers prepared an array of phosphotyrosine-containing peptides. Because they were unable to directly measure the extent of dephosphorylation of each peptide, they instead used a mutant PTP that could bind the substrate but was catalytically impaired and used this binding activity as a proxy for enzyme activity. They measured ‘activity’ profiles for sixteen PTPs using an assay where the PTP was conjugated to GST, which was then labeled with an anti-GST-Cy5 conjugate. However, there remains the possibility that the mutant enzyme has an altered substrate-specificity.[75, 116-118] Other work has used phosphopeptide arrays to profiles PTPs, with an anti-phosphotyrosine antibody to identify active substrates.[119]

Our development of the SAMDI-MS method provides a label-free and high-throughput assay for measuring a broad range of enzyme activities.[58, 60-63] This method uses SAMs that present the peptide substrate against a background of tri(ethylene glycol) groups, which are effective at preventing the non-specific adsorption of proteins and play an important role in biological assays.[42] Treatment of the SAMs with a enzyme solution may lead to a post-translational modification of the substrate, accompanied by a corresponding change in mass. The product can then be quantitated with MALDI-MS, which reveals peaks corresponding to both the substrate and product (and any intermediates or additional products). The SAMDI-MS method is compatible with the common 384 and 1,536 spot formats and has been used to profile enzymes with peptide arrays.[65, 71-73, 75-76, 120] We also recently demonstrated SAMDI could be used to profile the activity of a PTP on a phosphopeptide array.[99] Becker and coworkers' recent advance in studying protein-protein interactions using protein arrays and MALDI-MS also demonstrates the power of combining these technologies.[121-122]

In this chapter, we describe the use of a peptide array based on a sequence previously used in earlier studies of SHP2 activity: Ac-TRDXpYZTC-$NH_2$, and where the X and Z positions are variable.[94] We profiled 22 phosphatases and generally found that the specificities of those investigated previously were consistent with earlier reports, but our studies also revealed that many of the PTPs have unique and highly selective activities. Our data provides general rules of how charge, steric bulk, and hydrophobic character affect enzyme activity for the various PTPs. Finally, our data confirm that all of the phosphatases lack activity towards substrates that have an arginine or lysine residue to either side of the phosphotyrosine, and we discuss the possible relevance of this dependence in molecular mechanisms of diabetes.

## *4.2. Results and Discussion*

### *4.2.1. Preparation of Phosphopeptide Array*

We prepared an array having 361 peptides with the sequence Ac-TRDXpYZTC-$NH_2$, where the X and Z positions comprise each of the 19 canonical amino acids except for cysteine. In this way, we can determine which residues, when present adjacent to the phosphotyrosine, promote and inhibit activity. We used an array plate having 384 gold spots arranged in the geometry of the common microwell plate, where each was modified with a self-assembled monolayer presenting maleimide groups at a density of 10% against a background of tri(ethylene glycol) groups as described previously.[99] The peptides were synthesized using standard protocols with FMOC-protected amino acids and then stored in a 384 multi-well plate as described previously.[99] The peptides were transferred to the monolayer array plate using a robotic liquid handler, where each peptide underwent immobilization to the monolayer in its spot via conjugate addition of cysteine thiol to the maleimide group (Fig. 4-1).

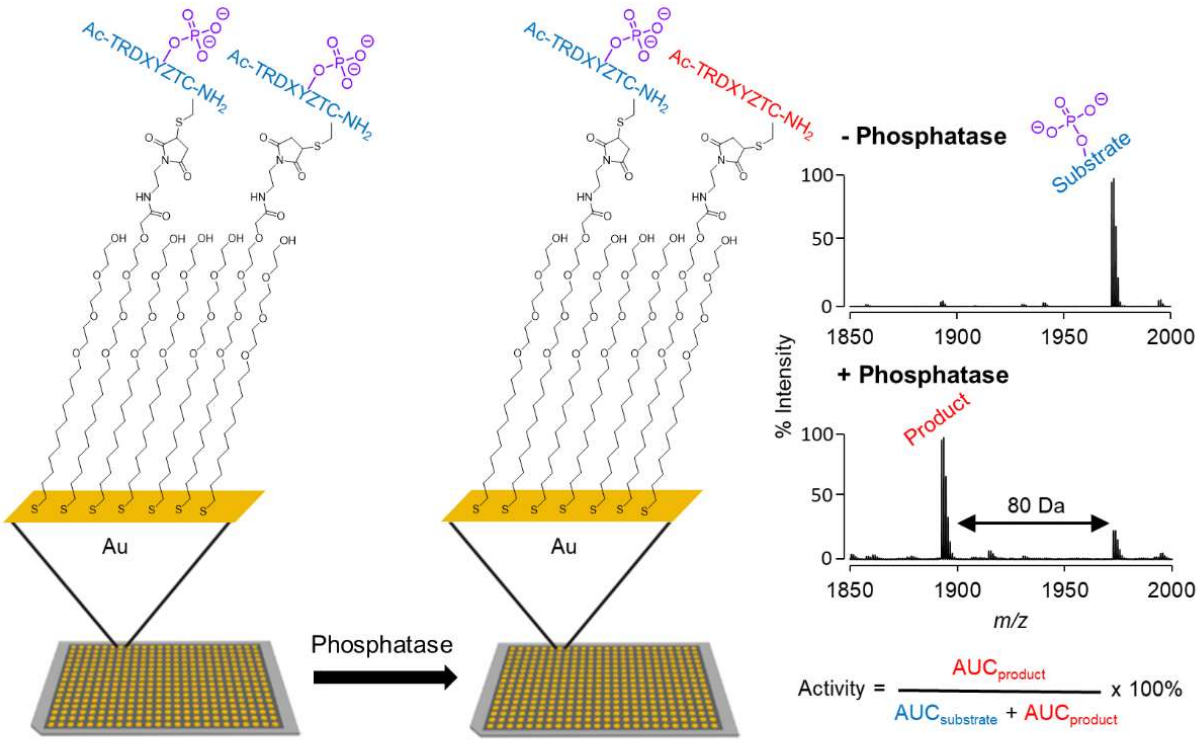



**Figure 4-1 Profiling phosphatase activities using peptide arrays and SAMDI-MS.**

(Left) Peptides are immobilized on a self-assembled monolayer surface presenting 10% maleimide against a background of tri(ethylene glycol) gruops. (Middle) The array is treated with a phosphatase and the extent of dephosphorylation of each peptide is analyzed with a MALDI-TOF mass spectrometer. (Right) A SAMDI spectrum of the initial monolayer has a peak at *m/z* = 1972 corresponding to the phosphotyrosine peptide−alkyldisulfide conjugate and a spectrum of the monolayer after treatment with a phosphate reveals a new peak at 80 Da lower mass, which corresponds to the dephosphorylated product.

### *4.2.2. Profiling Activities of DEP1 (PTPRJ)*

We first describe an experiment to profile the specificity of the transcriptional regulatory phosphatase DEP1 on the peptide array. We prepared a solution of the phosphatase (1.2 nM in 100 mM Tris, pH 7.5, 50 mM NaCl and 100 μM TCEP) and used a robotic liquid dispenser to rapidly apply 2 μL of this solution to each spot on the array plate. The array was placed in a humidified chamber at 37ºC for one hour and then rinsed first with water and then ethanol, and finally treated with THAP (2,4,6-trihydroxyacetophenone) matrix. The plate was analyzed using an AbSciex 5800 MALDI-TOF mass spectrometer to acquire mass spectra for each spot, which revealed separate peaks corresponding to the substrate and product of the reaction. The conversion of phosphopeptide to its product was characterized by integration of the corresponding peaks and is given by Activity = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 % where AUC refers to the area under the curve (Fig. 4-1). The ionization efficiencies of the substrate and product are not identical and therefore these nominal conversions are not calibrated, but the quantities do provide a relative measure of activity and therefore are useful in the following studies.

The activities for each peptide sequence are represented in a 19 x 19 heatmap where each row defines the amino acid in the Z position (+1), and each column defines the amino acid in the X (-1) position. The percent dephosphorylation is represented in greyscale with white corresponding to 0% activity and black to 100% activity. The heatmap of DEP1 (Fig. 4-2, upper left) shows that peptides containing a glycine in the Z position have higher activity, and similarly those having the aromatic residues phenylalanine and tyrosine, and to a lesser extent the hydrophobic residues isoleucine and leucine, in the X position are more active.

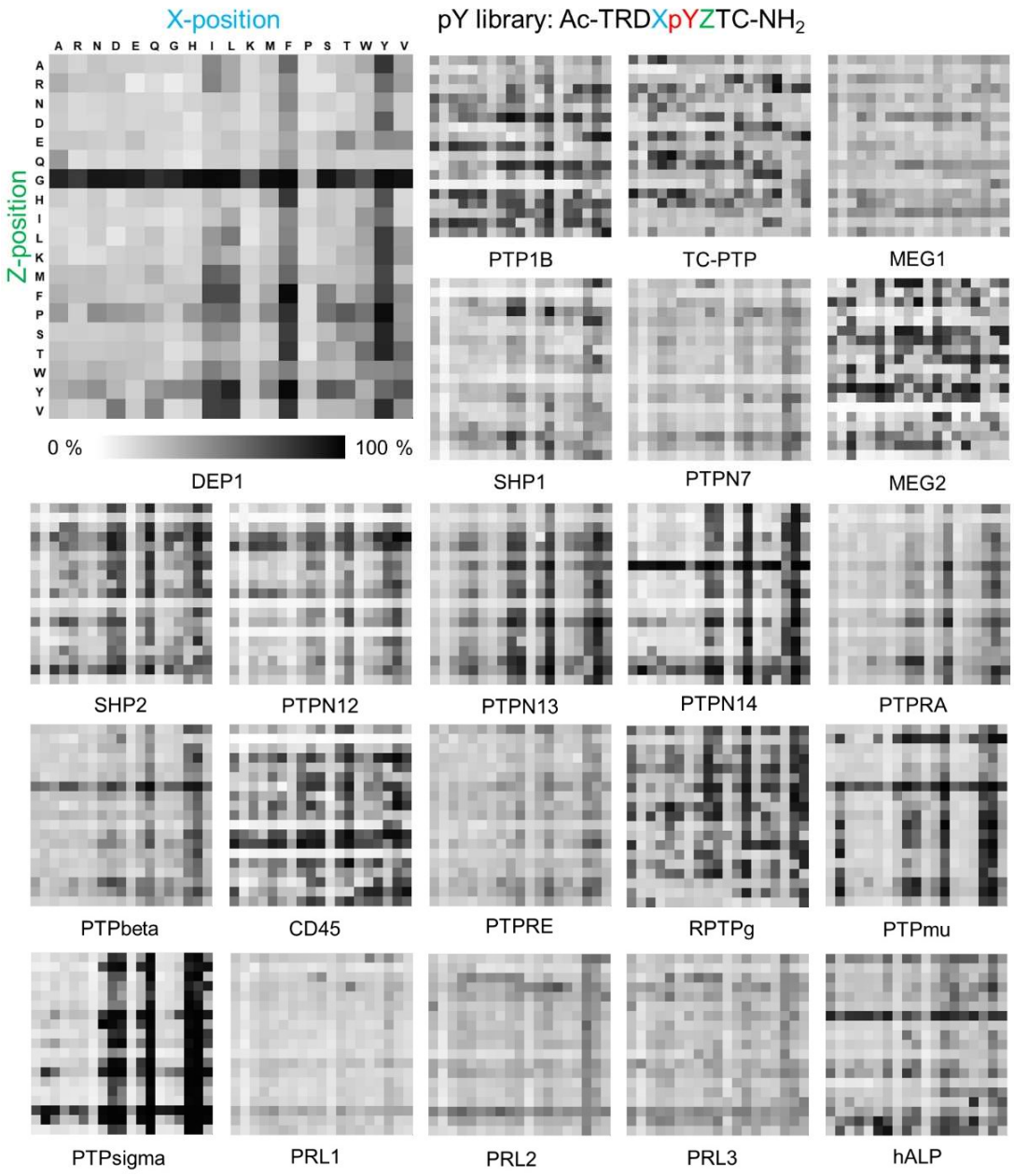



**Figure 4-2 Activity heatmaps for each of 22 phosphatases profiled on the peptide array having 361 sequences of form Ac-TRDXpYZTC-NH2.**

The columns represent 19 different amino acids at X position and the rows represent the amino acids at the Z position. The activities are represented in grey scale where white corresponds to 0 % activity and black to 100 %.

### *4.2.3. Profiling Activities of 22 Phosphatases.*

We repeated the experiment described above for 21 additional phosphatases—ten non-receptor protein tyrosine phosphatases (NRPTPs, classical PTPs), eight receptor protein tyrosine phosphatases (RPTPs, classical PTPs), three regenerating liver phosphatases (PRLs, VH1-like PTPs) and one alkaline phosphatase (ALP). The heatmaps for each of these phosphatases are shown in Fig. 4-2. Initial inspection of the heatmaps reveals several significant observations. First, the phosphatases show a range in specificity, with some having a clear preference for a smaller number of peptides within the array. Second, there are multiple classes of specificity that are shared by a subset of the enzymes. Third, there are positions in the substrate where certain residues affect the activity of the substrate in a similar way for all of the enzymes, and other specificities that are shared by a subset of the enzymes.

### *4.2.4. Quantitative Analysis of Heatmaps.*

We first analyzed the heatmaps to quantitate the influence that a particular residue has on activity of the substrate for each phosphatase. For each amino acid in the X position, we determined the average activity of all peptides in the row corresponding to that residue in the heatmap. We also determined the average activity for all peptides in the array (which we refer to as the global average). The ratio of the difference (Δ) of the average activity for the particular residue ($AA^x$) and the global average (GA) was then determine according to the equation $\Delta = (AA^x - GA)/GA \times 100\ \%$. We similarly repeated this analysis for residues in the Z position ($AA^z$). The numerical values of these ratios for all 38 rows and columns (19 amino acids at either the X or Z position) for each of the 22 phosphatases are shown in Fig. 4-3. This analysis gives insight

into the sequence determinants of activity. The residues that contribute to more than 10% activation or inhibition at X (-1) and Z (+1) positions are summarized in Table 4-1. The Table reveals that certain residues consistently activate or inhibit activity of the substrate across the phosphatases we profiled. Finally, we also generated histograms to identify the residues that activate and inhibit phosphatase activities at X and Z positions (Fig. 4-4). The x-axis represents amino acids and the y-axis is the total number of PTPs found from Table 4-1 that activates/inhibits dephosphorylation activities at either X or Z positions. We found that the determinants of sequence selectivity depend significantly on the chemical properties of the amino acid including charge, steric bulk, hydrophobic character as well as the PTP active site structure, and we discuss several of these observations below.

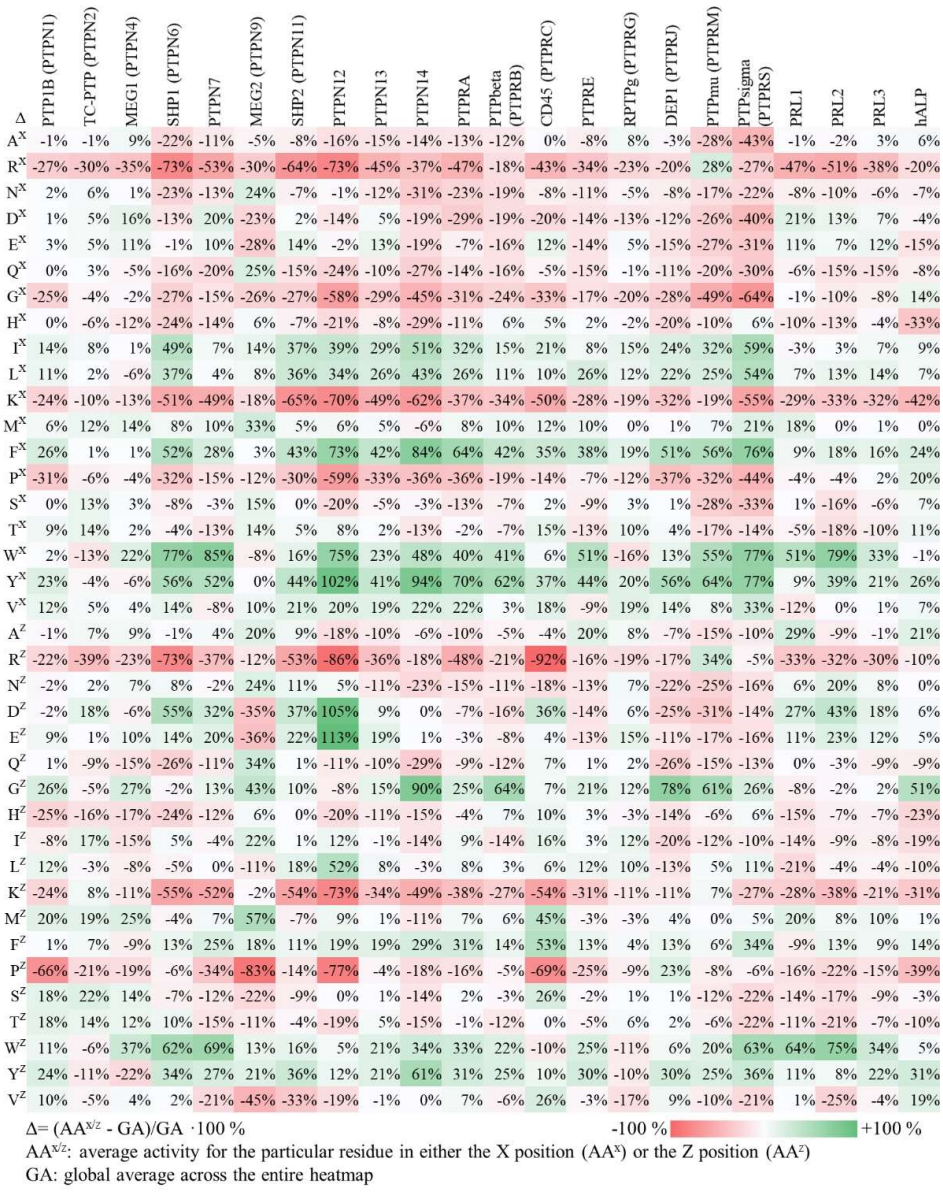



**Figure 4-3 The ratio of the difference (Δ) heatmap of 22 PTPs.**

Heatmap illustrating the average role than an amino acid residue in either the X ($AA^X$) or Z ($AA^Z$) position has on the activity of the substrate for each phosphatase enzyme. For each amino acid and position, the ratio of the difference of the average activity of peptides having that residue and the average of all peptides over the latter were determined and represented both in a red-to-green color scale and with the percent. Green indicates that the residue on average increases the phosphatase activity and red indicates a reduction.

**Table 4-1 Summary of PTP selectivity.**

| Phosphatase | activate X | | | inhibit X | | | activate Z | | | inhibit Z | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 10 % | 20 % | >30 % | 10 % | 20 % | >30 % | 10 % | 20 % | >30 % | 10 % | 20 % | >30 % |
| PTP1B (PTPN1) | I,L,Y,V | F | | | R,G,K | P | E,L,S,T,W,V | G,M,Y | | | R,H,K | P |
| TC-PTP (PTPN2) | M,S,T | | | K,W | | R | D,I,M,T | S | | H,Y | P | R |
| MEG1 (PTPN4) | D,E,M | W | | H,K | | R | S,T | G,M | W | Q,H,I,K,P | R,Y | |
| SHP1 (PTPN6) | V | | I,L,F,W,Y | D,Q | A,N,G,H | R,K,P | Y | | D,E,W | | Q,H | R,K |
| PTPN7 | E,M | D,F | W,Y | A,N,G,H,P,T | Q | R,K | G | E,F,Y | D,W | Q,H,S,T | V | R,K,P |
| MEG2 (PTPN9) | I,S,T,V | N,Q | M | K,P | D,E,G | R | F,W | A,N,I,Y | Q,G,M | R,L,T | S | D,E,P,V |
| SHP2 (PTPN11) | E,W | V | I,L,F,Y | Q | G | R,K,P | N,G,L,F,W | E | D,Y | P | | R,K,V |
| PTPN12 | | V | I,L,F,W,Y | A,D | Q,H,S | R,G,K,P | L,F,Y | | D,E | A,Q,T,V | H | R,K,P |
| PTPN13 | E,V | I,L,W | F,Y | A,N,Q | G | R,K,P | E,G,F | W,Y | | A,N,Q,H | | R,K |
| PTPN14 | | V | I,L,F,W,Y | A,D,E | Q,H | R,N,G,K,P | | F | G,W,Y | R,H,I,M,P,S,T | N,Q | K |
| PTPRA | | L,V | I,F,W,Y | A,Q,H,S | N,D | R,G,K,P | | G | F,W,Y | A,N,P | | R,K |
| PTPbeta (PTPRB) | I,L,M | | F,W,Y | A,R,N,D,E,Q,P | G | K | F | W,Y | G | N,D,Q,T | R,K | |
| CD45 (PTPRC) | E,L,M,T,V | I | F,Y | P | D | R,G,K,P | H,I,Y | S,V | D,M,F | N | | R,K,P |
| PTPRE | M | L | F,W,Y | N,D,E,Q,G,T | | R,K | L,F | A,G,W | Y | R,N,D,E | P | K |
| RPTPG (PTPRG) | I,L,F,T,V | Y | | D,K,P,W | R,G | | E,G,I,L | | | R,K,W,Y,V | | |
| DEP1 (PTPRJ) | W,V | I,L | F,Y | D,E,Q | R,G,H | K,P | F | P | G,Y | R,E,H,L,K | N,D,Q,I | |
| PTPmu (PTPRM) | | R,L | I,F,W,Y | N,H,K,T | A,D,E,Q,S | G,P | | W,Y | R,G | A,N,D,E,Q,I,S,V | | |
| PTPsigma (PTPRS) | | M | I,L,F,W,Y,V | T | R,N | A,D,E,Q,G,K,P,S | L | | G,F,W,Y | A,N,D,E,Q,I | K,S,T,V | |
| PRL1 | E,M | D | W | H,V | K | R | E,Y | A,D,M | W | H,I,P,S,T | L,K | R |
| PRL2 | D,L,F | | W,Y | N,Q,G,H,S,T | | R,K | F | N,E | D,W | S | P,T,V | R,K |
| PRL3 | E,L,F | Y | W | Q,T | | R,K | D,E,M | Y | W | P | K | R |
| hALP | G,T | F,P,Y | | E | R | H,K | F,V | A | G,Y | R,I,L,T | H | K,P |

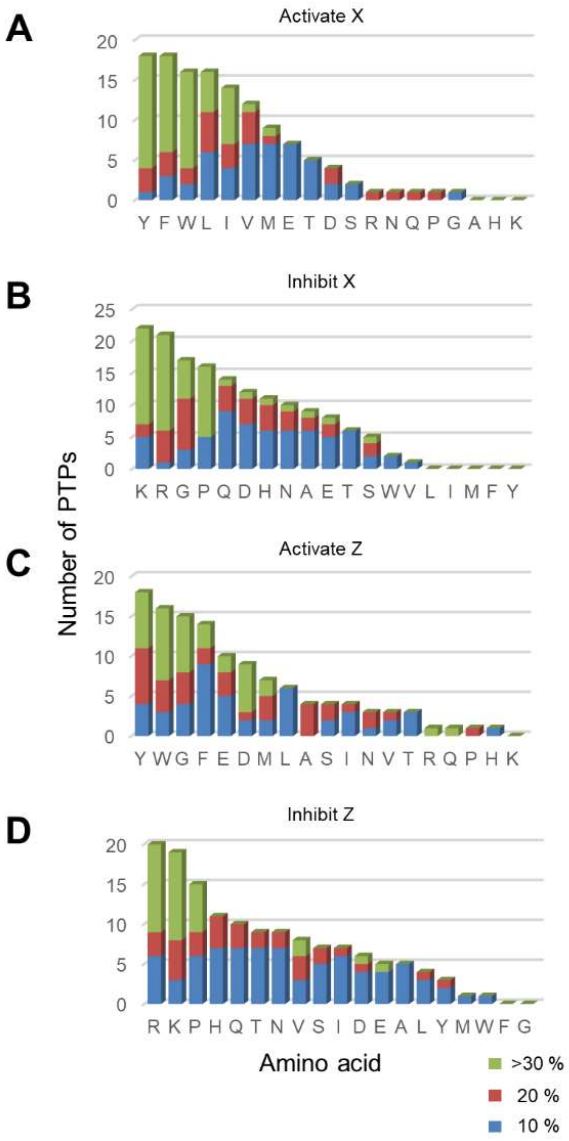



**Figure 4-4 Histograms showing the role of residues found in PTP specificities.**

The x-axis represents amino acids and the y-axis is the total number of PTPs that activates/inhibits dephosphorylation activities at either X or Z positions. The colors correspond to different degrees of activation/inhibition in activities assayed by SAMDI-MS. (A) Aromatic and hydrophobic residues are activating at the X position. (B) Basic residues, G and P are inhibitory at X position. (C) Aromatic residues are favored at the Z position while G is also favored in contrary to X position. (D) Basic residues and P, again, are inhibitory at the Z position.

### *4.2.5. Basic Residues Generally Reduce Substrate Activity*

Among the most consistent trends in our data is that basic residues (R, K, H) at either X or Z position very frequently decrease the activity of the substrate for all but one of the phosphatases studied here, consistent with a previous report.[115] Among 22 phosphatases, 21 exhibit a preference against basic residues adjacent to pY on their substrate. For example, the oncogenic phosphatase SHP2 shows an average decrease in activity of 64% and 53% when arginine is present at the X and Z positions, respectively. This trend may be explained by electrostatic interactions in the pY binding pocket. Most PTP catalytic sites have a signature motif HCXXGXXRS(T) or HC(X)$^5$R where a cysteine residue acts as the nucleophile and the conserved arginine residue folds back toward the phosphate-binding pocket to assist in substrate binding and catalysis.[15] The presence of a second positively-charged residue nearby may reduce the electrostatic attraction between the phosphatase and its substrate and would therefore lower the binding energy of the substrate. Our observation that arginine (R) and lysine (K) have a more significant effect in decreasing activity than does histidine (H) is consistent with their basicities, because the lower pKa of protonated imidazole means that it can be deprotonated with less energetic penalty. Below, we return to the significance of the basic residue in regulating phosphorylation states of proteins and possible relevance to disease.

We did find, however, one phosphatase, PTPmu, that favored an arginine residue in both the X and Z positions. This enzyme, though, does not tolerate a lysine residue in this position, suggesting that a specific hydrogen bonding interaction is involved and not just electrostatics. We studied the active site sequence using UniProtKB protein sequence alignment tool[123] and found that PTPmu has two PTP domains. The first one with catalytic motif HCSAGVGRT is

conserved across all assayed RPTPs. The second domain has a catalytic motif HCLNGGGRS that is different from others. This second catalytic domain might contribute to the arginine preference with extra hydrogen bonding interactions.

### *4.2.6. Acidic Residues Play a Complex Role*

Unlike the influence of basic residues described above, the presence of acidic residues (D and E) has varied effects on the activities of the peptide substrates for the phosphatases. Inclusion of an acidic residue can either increase, decrease or not affect the activity of the peptide, depending on the particular phosphatase. We found SHP1, SHP2, PTPN7, PTPN12, PTPN13 and PRLs showed increased activities for peptides having an acidic residue next to the pY substrate; PTPN14, MEG2, PTPRA, PTPRB, PTPRE, DEP1, PTPsigma, PTPmu showed lower activities on substrates having acidic residues in these positions; PTP1B, TC-PTP, MEG1, RPTPg and hALP were less affected by the presence of acidic residues in these positions. CD45 is an interesting case because an aspartic acid (D) at the X position reduces PTP activity but when present at the Z position, results in an increased activity. Our results reveal that acidic residues play a more complex role than commonly described as generally being favored.[115, 124-125] The quantitative SAMDI-MS assay allows us to see the whole picture of PTP activity profiles and understand how each amino acid contributes to the selectivities.

### *4.2.7. Glycine Has Different Effects at X and Z Positions*

Glycine (G) is a unique residue because of its wider conformational space in protein structure[126]. We found that glycine plays very different roles in affecting substrate activity when present at X (-1) and Z (+1) positions. At the Z position, glycine is often an activator (observed for 15 out of 22 phosphatases, Fig. 4-4C) and we find it never inhibits activity relative to the average activity for all peptides (Fig. 4-4D). This trend suggests that the binding pocket at the +1 position of the substrate does not participate in strong recognition of the sidechain. At the X position, in contrast, glycine frequently serves to decrease activity (observed for 16 out of 22 phosphatases, Fig. 4-4C). Clearly, side chain interactions of the substrate with the active site are important in this position.

### *4.2.8. Proline Is a Strong Inhibitor for PTPs*

Proline (P), because of the secondary amide, has a turn conformation that often disrupts binding of a substrate to the active sites of enzymes.[127] We found that in no case is the presence of a proline residue at either the X or the Z position favorable for PTP activity. All 21 PTPs disfavor proline residues at both X or Z positions on their substrates (Fig. 4-4B and 4-4D). ALP has a slight preference for a proline at the X position though we note it is not a classic cysteine-based protein tyrosine phosphatase and the active site structure is very different.[128]

### *4.2.9. Aromatic and Hydrophobic Residues Are General Activators*

We found that aromatic residues (Y, F, W) are among the most consistent activators for PTP activity (Fig. 4-4A and 4-4C). This observation is consistent with previous work that analyzed substrate selectivity of SHP1, SHP2, PTP1B and PTPRA[114]. We found the trend is universal to all 21 PTPs assayed. A crystal structure provides some structural insight into the preference for aromatic residues.[102] In SHP2, two aromatic groups (Y279, H426) occupy either side of the PTP active site, providing π-π interactions to pY and its neighboring residues on the substrate. The analogous residues in PTP1B are Y46 and F182. From sequence alignment of all the PTPs assayed, we found that the role for these two residues in substrate recognition are highly conserved. All PTPs have the Y residue (except PTPN14, which has an I) in the first position and either an H, Y or F in the second position (except RPTPg which has an M). These residues also provide hydrophobic interactions to the hydrophobic residues (A, I, L, V) while they are less activating compared to aromatic residues perhaps because the interaction is weaker and less specific.

### *4.2.10. Polar Neutral Residues Are Less Involved in PTP Specificities*

Unlike residues having side chains with charge, aromatic or steric non-polar groups, polar neutral residues (N, Q, S, T) have fewer features for molecular recognition and play a minor role in enhancing or decreasing the activities of the peptide substrates. Amides (N, Q) are more frequently found to be slightly disfavored, but even this trend is inconsistent across the phosphatases.

### *4.2.11. Comparisons to Prior Studies*

Pei and co-workers' pioneering studies of PTP specificity used combinatorial peptide libraries prepared by split-pool synthesis and had more than $10^6$ peptides.[115] Beads that had active peptide sequences for the PTPs were identified with a colorimetric labeling method and then sequenced by partial Edman degradation-mass spectrometry (PED-MS). This approach has the benefit that it rapidly identifies the most active sequences from a very large number of peptides, but the limited throughput of PED-MS only allows a very small fraction of the beads to be sequenced. In their experiments as low as 0.0008% of peptides were sequenced and therefore the overall specificity and the nuanced preferences for the PTPs are not revealed. The approach also relies on a qualitative isolation of beads based on intensely, medium and lightly red beads and does not directly provide a more quantitative ranking of activities. In these respects, our work with peptide arrays and mass spectrometry provides a complementary insight into the PTP specificity. We note that the expense associated with peptide synthesis currently limits our array sizes to hundreds (not thousands or millions) of peptides, but it does have the benefit of providing relative activities for each peptide substrate in the array.

Pei and co-workers first discovered that PTPs generally have a preference of substrates having an abundance of acidic residues and disfavor substrates with basic residues. Our work confirms this finding that basic residues are indeed the most disfavored for PTPs in their substrates, and reveals that acidic residues promote activity, though they play more complex roles when compared to all other amino acids. SHP1, SHP2, PTPN7, PTPN12, PTPN13 and PRLs prefer acidic residues whereas PTPN14, MEG2, PTPRA, PTPRB, PTPRE, DEP1, PTPsigma, PTPmu disfavor them; and they are not particularly preferred when compared to all

other amino acids for PTP1B, TC-PTP, MEG1, RPTPg and hALP. For those PTPs that prefer acidic residues, we found that the preference is more obvious at Z (+1) position. For example, a glutamic acid (E) at X (-1) leads to a small -2% activity decrease for PTPN12, however at Z (+1) it increases activity by +113% for this phosphatase, and was consistent with results reported by Pei and coworkers. Similarly, we also find that aromatic residues in the substrate promote activity for the PTPs. While the combinatorial screening approach taken by Pei and coworkers successfully captured the most obvious characteristics of PTP substrate selectivites, our use of peptide arrays gave a more complete understanding of specificity. We are able to better assess those residues that generally contribute to lower PTP activities and understand their roles. We also found that glycine has an opposite effect when present in the X and Z positions; prolines are disfavored in general; non-aromatic hydrophobic residues are also activators and polar residues are less involved in PTP specificities.

### *4.2.12. Categorizing PTPs According to Their Substrate Selectivities.*

We next asked whether the different specificities of the PTPs are related to the phylogenetic tree and sequence similarity of the enzymes. In Fig. 4-5, we show how the presence or absence of five amino acids (R, D, E, G, P) can be used as filters to partition the PTPs into distinct classes. Starting at the left of the tree we divide the enzymes into groups according to each of eight features (listed at the top of the figure), where the upper group has the feature and the lower group does not. We started by selecting the enzymes that favor a P residue in the X position and followed by those that favor an R in either X or Z positions; these rules isolated a single phosphatase (hALP). We then selected PTPs that favor a R residue in both the X and Z positions

(again, isolating one enzyme PTPmu). Because acidic residues play a complex role in PTP selectivity—where they can either increase, decrease or not affect the activity on the peptide—we then separated PTPs that either favor D and E in both the X and Y positions ($D,E^{X,Y}$) or disfavor these residues. Finally, noting that glycine often promotes activity when present in the Z position but decreases activity in X position, we used this filter to further separate the PTPs. Finally, we separated the PTPs according to role of proline in decreasing activity.

We used the UniProtKB tool to compare and align the PTP sequences and we generally found that PTPs that are evolutionarily related have similar specificities. For example, SHP1 and SHP2 share >55% sequence identity (60.8 similarity in PTP domain), a common backbone fold, and a common regulatory mechanism.[129] We found that their substrate specificities were indistinguishable by the criteria we used in preparing the selectivity tree (Fig. 4-5). Similarly, PTPN13 shares a common substrate specificity with SHP1/2 in our selectivity tree and also has a high sequence similarity to SHP1 and SHP2 in the protein sequence. Three PRLs are very similar in both sequence and substrate selectivity and yet different from classic PTPs. The three phosphatases PTPRA, DEP1 and PTPbeta all belong to the same receptor-like PTP family and exhibited a similar specificity. The enzyme hALP is serine-based phosphatase that uses $Zn^{2+}$ and $Mg^{2+}$ activation, and its catalytic mechanism is very different from the cysteine-based PTPs. The sequence is completely unique, and this is the only enzyme we find that prefers a P residue in the X position. We also note that these observations agree with an evolutionary tree analysis previously reported by Cesareni and co-workers.[18]

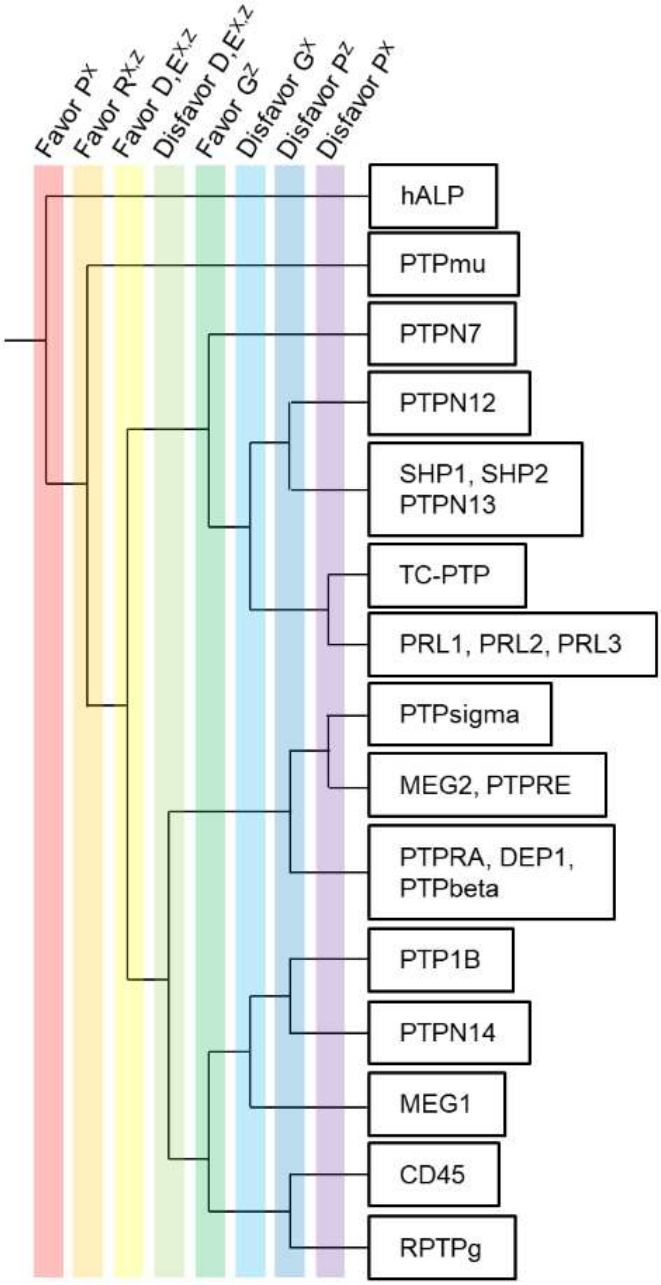



**Figure 4-5 Specificity tree of the 22 phosphatases profiled.**

We use specificity features based on five amino acid residues (P, R, D, E, G) at X and Z position to differentiate the phosphatases. A well-separated tree indicates that phosphatases are unique in their substrate specificity.

### *4.2.13. A Potential Mutation/Modification Crosstalk Involving PTP1B*

The phosphatase specificities determined in this work can contribute to our understanding of the mechanisms by which mutations contribute to pathological phenotypes. For example, Beguinot and coworkers reported a mutant of the insulin receptor (INSR, R1152Q) that was found in patients with non-insulin-dependent diabetes (Fig. 4-6A).[130] This mutant is interesting because it is adjacent to a tyrosine that is known to be autophosphorylated ($Y^{1151}$) by the insulin-bound receptor and that participates in the tyrosine triplet that activates the receptor tyrosine kinase (RTK) activity of INSR and regulates the downstream cellular signal transduction.[131] This study also reported that the mutant had a higher level of basal kinase activity on various endogenous and exogenous substrates, even though phosphorylation of $Y^{1151}$ is significantly lower than that in the wild-type receptor. The phosphorylation of the mutant by the wild-type receptor was also impaired, leading to the conclusion that the mutation led to poor phosphorylation of $Y^{1151}$ because it is a less active substrate for its own kinase domain, and in turn had a lower sensitivity to insulin, which in turn contributes to diabetes. It is interesting that this study did not consider the possibility that the mutation made the site more active for its PTP; because the level of phosphorylation depends on the dynamic balance of two opposing activities, it can be decreased either by a lower kinase activity or a higher phosphatase activity. In this example, PTP1B is specifically responsible for dephosphorylating this site.[132]

Our profiling experiments revealed that replacing an arginine with a glutamine on a peptide substrate leads to an increase in PTP1B activity. This insight would suggest that mutant INSR shows less phosphorylation at $Y^{1151}$ both because it is a poorer substrate for the kinase and because the corresponding phosphorylated peptide is a better substrate for PTP1B. We gained

support for this possibility by comparing the relative activities of PTP1B on 2 peptides—Ac-TRDYpYRTC-$NH_2$ (YpYR) and Ac-TRDYpYQTC-$NH_2$ (YpYQ)—where we found the latter was significantly more active as a substrate for PTP1B. A crystal structure of PTP1B and the triple-phosphorylated INSR peptide reveals how the arginine residue is repelled by this highly basic triple-pY-binding pocket.[133]

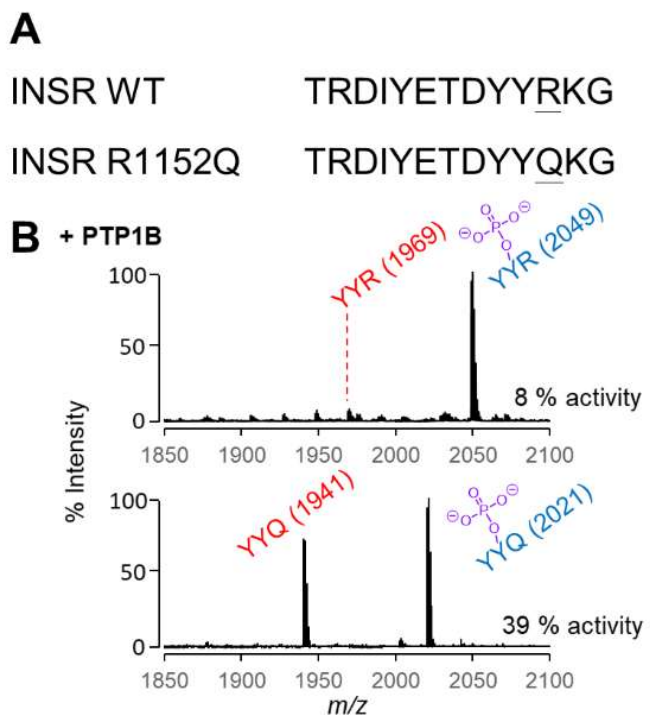



**Figure 4-6 Dephosphoraylation of peptides corresponding to WT and mutant INSR by the PTP1B phosphatase.**

A) The mutation R1152Q occurs in the triple phosphorylation site of INSR and is adjacent to the $Y^{1151}$ phosphorylation site. B) SAMDI spectra following treatment of each peptide with PTP1B shows significantly more activity on the peptide corresponding to the mutant.

## *4.3. Conclusion*

This work addresses the long-standing challenges in characterizing phosphatase activities. Our use of peptide arrays and SAMDI-MS is the first approach to rapidly determine the specificity profiles of phosphatases using hundreds of peptide substrates. By applying this method to 22 phosphatases, we observed trends in sequence-dependent activity that will be useful in developing hypotheses for phosphorylation-dependent signaling activities in cells; many of these trends agreed with previous work based on combinatorial screening, but gave a more complete assessment of specificity. The most general trend—that the presence of a basic residue next to the phosphotyrosine site decreases PTP activity on the substrate—will likely have broad relevance (such as our description of a mutant in non-insulin dependent diabetes), and we expect these data will be valuable in many more contexts. Equally important, we believe this approach will lead to more complete understandings of the biochemistry and biology of phosphatase enzymes.

## *4.4. Methods*

**General**. Laboratory chemicals and reagents were purchased from MilliporeSigma and used without additional purification unless specified. Peptide synthesis reagents, including Fmoc amino acids and Rink-amide resin, were purchased from Anaspec. Phosphatases were purchased from MilliporeSigma. Self-assembled monolayers with matrix-assisted laser desorption-ionization (SAMDI) mass spectrometry were performed on a 5800 MALDI TOF/TOF mass spectrometer (AbSciex) using either manual or automated protocols. A detailed protocol of monolayer plate preparation, peptide synthesis and phosphatase assay can be found in a previously published method paper.[99]

**Peptide Synthesis.** Solid-phase peptide synthesis was performed on Rink-amide resin (10 mg) housed in 96-well filter plates. N-terminal fluorenylmethyloxycarbonyl (Fmoc) protecting groups were deprotected with 20% piperidine in dimethylformamide (DMF) at room temperature for 20 min. The resin was filtered and rinsed 5 times with DMF on a multiscreen vacuum manifold. Amino acids were coupled to the resin with PyBop and N-methylmorpholine (NMM) in DMF at a 4:4:8 molar excess for 30 minutes, twice. The deprotection and coupling were repeated for each residue. Following deprotection of the final residue, the N-terminus was acetylated with 10% acetic anhydride in DMF for 60 minutes. Peptides were cleaved and deprotected in 95% trifluoroacetic acid (TFA), 2.5% triethylsilane (TES), and 2.5% water for 16 h. The cleavage solution was evaporated with $N_2$ gas flow. The peptides were resuspended in 0.1% TFA in water, lyophilized and resuspended again in 0.1%TFA in water to a final concentration of 200 μM and stored at -80 $^{o}$C.

**Preparing Peptide Arrays on SAMDI Plates.** Steel array plates evaporated with 384 gold spots (with a diameter of 3.0 mm) were soaked at 4 ºC in a solution of 1 mM total disulfide with 0.8 mM tri(ethylene glycol)-terminated $C_{11}$-alkane disulfide and 0.2 mM $C_{11}$-alkane disulfide with one terminal tri(ethylene glycol) and one terminal maleimide in ethanol for 2 days to allow assembly of the monolayer (10% maleimide coverage). Peptides were diluted to a final concentration of 20 μM with 100 mM Tris buffer (pH 7.5) in a 384-well plate with 5 μL TCEP beads (Thermo Scientific) in each well. Using a TECAN robotic liquid handler, 2 μL of peptide solution from each well was transferred on to the corresponding gold spot on the monolayer-presenting plate and incubated at 37 ºC in a humidified chamber for 1 h to allow peptide immobilization.

**Phosphatase Assays by SAMDI Mass Spectrometry.** Phosphatases were diluted in PTP buffer (100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and 2 μL was applied to each gold spot on the peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated at 37 ºC in a humidified chamber. (See Table 4-2 for concentration and incubation time for each phosphatase.) After the reaction was complete, the plate was rinsed with water and ethanol, treated with 1 μL matrix (10 mg/mL THAP, 5 mg/mL ammonium citrate dibasic in 50% acetonitrile, 50% water and 0.1 % phosphoric acid) to each spot and dried in air for 20 min. The spots were analyzed by MALDI-TOF MS to obtain a mass spectrum for each reaction. Enzymatic activities were quantified by measuring the areas under the curve (AUCs) for the dephosphorylate product peak and the substrate peak and determining the activity (%) = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %. Activity heatmaps were generated by Microsoft Excel.

**Table 4-2 Reaction conditions for profiling phosphatase activities.**

| Phosphatase | MilloporeSigma Cat. # | Concentration (nM) | time | Temperature |
|---|---|---|---|---|
| hALP[a] | 524604 | 1.4 U/mL[b] | 1 h | 37 ºC |
| PTP1B (PTPN1) | P6244 | 6.7 | 1 h | 37 ºC |
| TC-PTP (PTPN2) | SRP0218 | 9.3 | 1 h | 37 ºC |
| MEG1 (PTPN4) | SRP0205 | 0.34 | 1 h | 37 ºC |
| SHP1 (PTPN6) | SRP5076 | 14 | 4 h | 37 ºC |
| PTPN7 | SRP5077 | 93 | overnight[c] | 37 ºC |
| MEG2 (PTPN9) | SRP0206 | 4.2 | 1 h | 37 ºC |
| SHP2 (PTPN11) | SRP0217 | 54 | 1 h | 37 ºC |
| PTPN12 | SRP5073 | 25 | overnight | 37 ºC |
| PTPN13 | SRP5074 | 20 | overnight | 37 ºC |
| PTPN14 | SRP0211 | 164 | overnight | 37 ºC |
| PTPRA | SRP5078 | 12.5 | overnight | 37 ºC |
| PTPbeta (PTPRB) | P9864 | 2.2 | 1 h | 37 ºC |
| CD45 (PTPRC) | SRP0219 | 15 | 1 h | 37 ºC |
| PTPRE | SRP5333 | 34 | overnight | 37 ºC |
| RPTPg (PTPRG) | SRP0223 | 10 | 1 h | 37 ºC |
| DEP1 (PTPRJ) | SRP0220 | 1.2 | 1 h | 37 ºC |
| PTPmu (PTPRM) | SRP0222 | 78 | 2 h | 37 ºC |
| PTPsigma (PTPRS) | SRP0224 | 20 | overnight | 37 ºC |
| PRL1 | SRP0208 | 1683 | overnight | 37 ºC |
| PRL2 | SRP0209 | 1745 | overnight | 37 ºC |
| PRL3 | SRP0210 | 1710 | overnight | 37 ºC |

[a]Buffer pH adjusted to 8.0, supplied with 2.5 mM $Mg^{2+}$ and 0.5 mM $Zn^{2+}$

[b]Molar concentration not provided by the vendor

[c]Overnight = 16-20 h

# Chapter 5. Tyrosine Phosphatase Activity Is Restricted by Basic Charge Substituting Mutation of Substrates: Implications for Missense Mutations Associated with Disease

*This chapter is reproduced in part from Huang, C.-F.; Gottardi, C. J.; Mrksich, M. Tyrosine Phosphatase Activity Is Restricted by Basic Charge Substituting Mutation of Substrates: Implications for Missense Mutations Associated with Disease. *Manuscript in prep.* 

## *5.1. Introduction*

While phosphoproteome equilibrium states are determined by competing kinase and phosphatase activities, most studies overwhelmingly address the role of kinases and assume that phosphatases are non-regulatory.[6] However, while sharing the same phosphorylation sites, the numbers of protein tyrosine kinases (PTK) and protein tyrosine phosphatases (PTP) in the human proteome are roughly equal (90 PTKs and 107 PTPs).[4-5] The "housekeeper" assumption of phosphatases lacks justification and appears inconsistent. Emerging work further demonstrate sophisticated regulatory roles for PTPs, [7-10] where dysregulated PTP activity is linked directly to disease. For example, PTP1B contributes to the development of insulin resistance and type 2 diabetes[134] and SHP2 is identified as the first oncogenic phosphatase.[11-13]

Understanding general rules for PTP-substrate specificity would advance our understanding of their roles in signaling.[135-136] Phosphatase activity assays are challenging and largely not well-suited to direct determination of phosphatase specificity. Our group developed SAMDI-MS (self-assembled monolayers for matrix-assisted laser desorption/ionization mass spectrometry), a label-free technology to assay enzyme activities.[58] Combined with peptide arrays, we profiled a series of enzymes—including acetyltransferases, glycotransferases, proteases, kinases and phosphatases—determined their specificities, studied the nature of protein-substrate interaction, improved the design of mutants and probes for synthetic biology purposes and understood the underlying mechanism of disease development.[63, 65, 73, 75-76, 99] In our recent study, we profiled 22 PTPs and found most lack activity towards substrates that have an arginine (R) or lysine (K) residue to either side of the phosphotyrosine (pY).[137] PTP selectivity against basic residues agreed with those in the prior studies.[115-116] Using 5 different cell lysates, we confirmed that PTP

activity disfavors a basic residue adjacent to the pY site.[138] This finding raises the possibility that charge-substitution mutations proximal to pY sites could impact a range of signaling proteins and disease pathology. For example, the insulin receptor R1152Q mutant found in several type 2 diabetes patients manifested lower Y-phosphorylation levels, previously attributed to reduced kinase activity.[130] However, our work suggests that loss of the positive charge at R1152Q might increase accessibility of phosphatases towards pY, leading to less Y1151 phosphorylation.[137] In the study that follows, we aim to validate our hypothesis that charge-substitution missense mutations proximal to Y residues can affect pY abundance and signaling *in vivo*, using the established pY modified adhesion-signaling protein, β-catenin (β-cat).

β-cat is a canonical junction-nuclear signaling protein. As a physical link between cell surface cadherins and the actin-binding protein α-catenin, β-cat plays a critical role in cell-cell cohesion. Outside of the cadherin-catenin complex, β-cat also functions as essential effector of the Wnt pathway, pairing with DNA-binding factors to activate transcription of genes that drive cell fate decisions.[139] The degree to which β-cat participates in these two functions is determined by the availability and relative affinity of β-cat binding with its competing partners.[140]

One modification that shifts β-cat from its cell-cell adhesive function in favor of nuclear signaling is its phosphorylation at Y654 by Src family kinases (SFKs) where this modification reduces β-cat binding to the cadherin cytoplasmic domain[141-142] and also antagonizes its degradation by the β-cat phospho-destruction complex[143-144], leading to additional modifications that enhance transcriptional activity.[145] Interestingly, forward genetic screens in mice generated a missense mutation that created a basic residue adjacent to Y654 (T653K), leading to intellectual disability and syndromic features including a broaden face (batface, *Bfc*).[146] While the *Bfc*

mutation reduced β-cat binding to N-cadherin and mimicked Wnt-pathway gain-of-function phenotypes during early neuro-development, how the T653K mutant leads to elevated β-cat signaling remains unknown.

In this chapter, we establish the broad principle that basic amino acid missense mutations proximal to neighboring phosphotyrosine residues restrict PTP-mediated dephosphorylation. Through a database search, we identify 6,000+ cancer mutations involving basic residues being added or removed adjacent to the phosphotyrosine sites.[147-148] The β-cat T653K *Bfc* mutant and its other basic variation T653R are both found in the cancer genome.[149] We demonstrate that introducing a basic residue adjacent to β-cat Y654 antagonizes its dephosphorylation by SHP1, favoring persistent phosphorylation and elevation of the canonical β-cat target, AXIN2. This mechanism rationalizes neuro-developmental phenotypic similarities between T653K *Bfc* and Y654E phospho-mimic mutant mice[145-146], as well as how β-cat missense mutations proximal to Y654 may drive epithelial cancers.[149]

## *5.2. Results*

### *5.2.1. Profiling Activities of SHP1*

We sought to profile SHP1 (PTPN6) specificity on a peptide array with the sequence Ac-TRDXpYZTC-$NH_2$, where the X and Z positions comprise each of the 19 canonical amino acids except cysteine. The array was synthesized and immobilized to a 384-gold-spot plate as previously described.[99] A solution of the SHP1 (14 nM in 100 mM Tris, pH 7.5, 50 mM NaCl and 100 μM TCEP) was prepared and dispensed to each spot on the array plate (Fig. 5-1, left, 2 μL). The reaction was carried out in a humidified chamber at 37 $^{o}$C for four hours and then rinsed with water and ethanol. Each spot was treated with THAP (2,4,6-trihydroxyacetophenone) matrix and analyzed using a MALDI-TOF mass spectrometer. Conversion of phosphopeptide to its product was characterized by integration of the corresponding peaks, where Activity = Area Under the Curve $(AUC)_{product}$ / $(AUC_{substrate} + AUC_{product})$ x 100 %. Activities for each peptide sequence are represented in a 19 x 19 heatmap (Fig. 1, right) where each column defines the amino acid in the X (-1) position and each row defines the amino acid in the Z position (+1). Activity is represented by a color scale with white corresponding to 0 % and purple as 100 %.

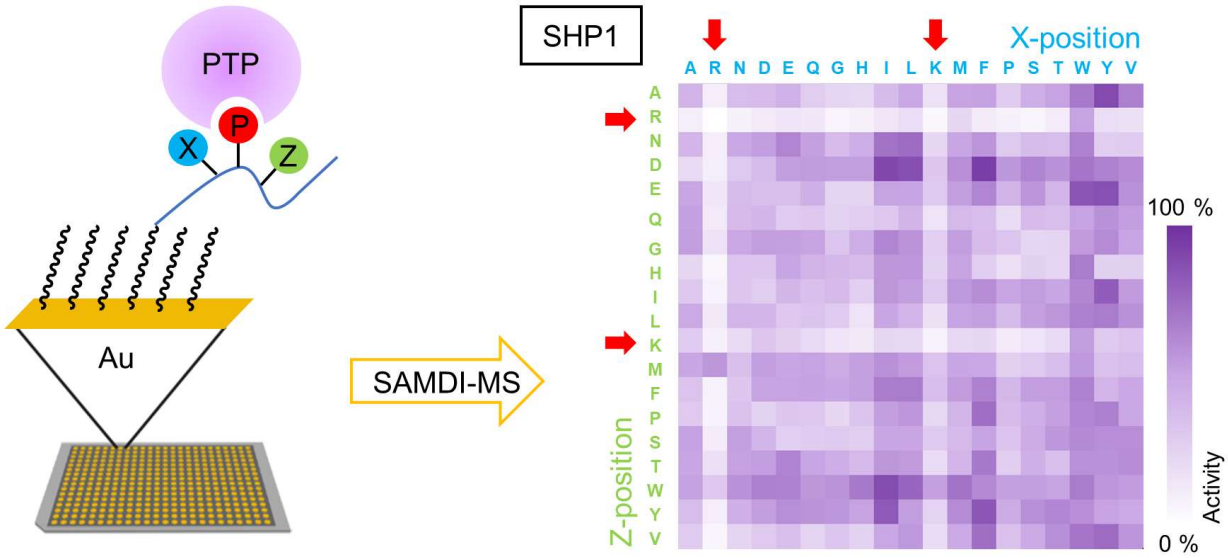
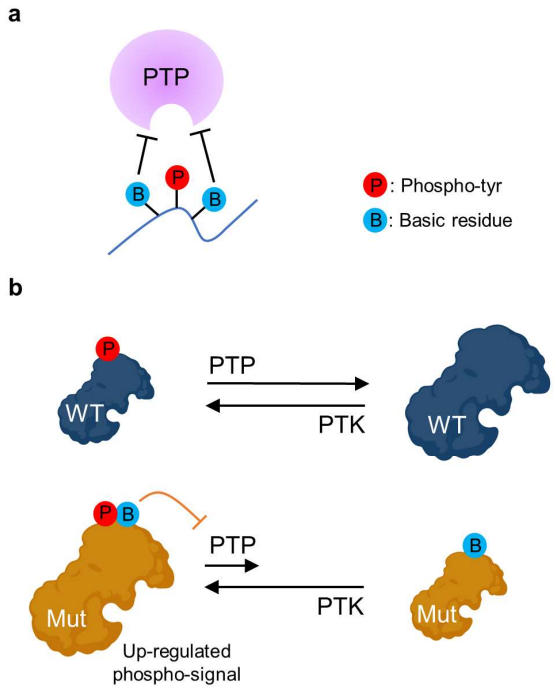



**Figure 5-1 Phosphatase activity profiling using peptide arrays and SAMDI-MS.**
(Left) Peptides immobilized on a self-assembled monolayer surface presenting a background of tri(ethylene glycol) groups. Array is treated with PTP and dephosphorylation of each peptide is analyzed with SAMDI-MS. (Right) Heatmap for SHP1 shows basic residues adjacent to phosphotyrosine inhibit enzymatic activity.

### *5.2.2. Basic Residues Inhibit PTP Activity*

The SHP1 heatmap (Fig. 5-1, right) clearly shows peptides containing basic residues (R or K) at either X or Z position exhibit lower dephosphorylation activity. In addition to SHP1, we found almost all 22 profiled PTPs disfavor basic residues.[137] Consistent with results from individual purified PTPs, we found that collective PTP activities also disfavor basic residues adjacent to the phosphotyrosine across five different cell lines.[138] Since classic PTPs have a signature catalytic site with the motif HCXXGXXRS(T) or HC(X)$^5$R where a cysteine residue acts as the nucleophile and the conserved arginine residue assists in substrate binding and

catalysis[15], we propose that basic residues adjacent to phosphotyrosine can universally decrease PTP activity, because they disrupt electrostatic interactions in the catalytic pocket (Fig. 5-2a).




c

| Gene site | WT aa | Mut aa | pY site | Cancer type |
|---|---|---|---|---|
| ABL1.254 | G | R | Y253 | uterine cancer, leukemia |
| BCR.277 | Q | R | Y276 | melanoma |
| CBLB.134 | E | K | Y133 | lung cancer |
| CTNNB1.653 | T | R | Y654 | germ cell cancer, ovarian cancer |
| CTNNB1.653 | T | K | Y654 | stomach cancer |
| EGFR.868 | E | K | Y869 | breast cancer, lung cancer |
| EGFR.977 | R | C, L, H | Y978 | uterine cancer, liver cancer, breast cancer |
| ERBB2.1111 | R | W | Y1112 | pancreatic cancer |
| HRAS.31 | E | K | Y32 | breast cancer, uterine cancer |
| KRAS.63 | E | K | Y64 | melanoma, brain cancer, germ cell cancer |
| LCK.393 | E | K | Y394 | melanoma |
| TP53.221 | E | K | Y220 | esophageal, vulvar, parotid gland |

**Figure 5-2 Basic residues antagonize PTP activity.**

**a**, Basic residues inhibit PTP–positively-charged pocket from associating with phosphotyrosine substrates. **b**, Charge-mutation/PTP-restriction mechanism: Mutant with a basic residue adjacent to the phosphotyrosine antagonizes dephosphorylation, leading to persistent phospho-signal compared to wild-type. **c**, Basic residue-substitution missense mutations identified for select cancer-relevant genes.

### *5.2.3. A Charge-Mutation/PTP-Restriction Mechanism*

Given the general trend that PTPs disfavor basic residues adjacent to their phosphotyrosine substrates, we reasoned that charge-substitution mutations proximal to phosphotyrosines could impact pY steady-state levels due to reduced PTP accessibility, leading to mis-regulated

phospho-signal across a range of proteins and disease pathology. Fig. 5-2b shows one scenario where a basic residue created adjacent to phosphotyrosine in the missense mutant prevents PTP from efficiently dephosphorylating the substrate. Consequently, steady-state concentration of the phospho-protein increases and up-regulates downstream signaling.

We observed an opposite scenario for the diabetes-related R1152Q insulin receptor missense mutation, which substituted a basic residue with a neutral amino acid.[137] This mutation facilitated dephosphorylation of the neighboring pY, leading to reduced steady-state concentration of pY and signals that ultimately promote diabetes development. While the previous model rationalized reduced kinase activity towards the insulin receptor R1152Q mutant as responsible for its lower pY levels,[130] we wondered if perturbed phosphatase accessibility could play a wider role in missense mutations associated with diseases. Indeed, an appeal of our model is that it generally applies to all members of the PTP family (Fig. 5-2b).

### *5.2.4. Database Search Identifies 6,000+ Relevant Cancer Mutants.*

*This work was performed in collaboration with Dr. Lihua Zou.

To identify cancer mutants with basic residue substitutions adjacent to phosphotyrosines, we employed two datasets: the protein post-translational modification (PTM) database PhosphoSitePlus and single-nucleotide variations (SNV) cancer database BioMuta.[147-148] We used PhosphoSitePlus to identify all reported phosphotyrosine sites and found corresponding cancer mutations at their +1 and -1 positions in BioMuta. We then filtered these sites with amino acid R and K appearing on wildtype or mutant at the +1 and -1 positions to identify base-

substitution missense cancer mutations including both base additions and removals. To our surprise, we identified 6,000+ mutants, most with poorly defined contributions to protein function. Notable sites identified on oncogenes and the tumor suppressor p53 are listed in Fig. 5-2c, which might be relevant to the direct development of cancers. While these pY-proximal missense mutations appear to occur frequently, there has been no way to systematically study them, nor generalized rule for how such mutations might affect the phospho-status of adjacent sites.

Towards this end, we became interested in the *CTNNB1* gene/β-cat protein, because T653R and T653K were identified in the above search and proximal to the well-studied Y654 phospho-site associated with reduced cell-cell adhesion and elevated nuclear signaling. Moreover, β-cat T653K was also identified in a screen for mutations affecting craniofacial development in mice, generating phenotypes that overlap with dominant human β-cat/*CTNNB1* mutations causing intellectual disability with similar syndromic features.[146] While the β-cat T653K mutant was rationalized to work as a mild, gain-of-function signaling form of β-cat due to phenotypic similarity with a β-cat Y654E phospho-mimic knock-in mice,[145] rigorous evidence that β-cat T653K leads to enhanced Wnt signaling through persistent phosphorylation of Y654 was lacking. We sought to address this hypothesis both as a means to understand the disease relevance the β-cat T653K mutation and to test the general rule that basic amino acid charge-substitution proximal to pY leads to sustained phospho-signaling.

### *5.2.5. β-cat T653 Basic Mutants Show Greater Phosphorylation at Y654.*

We obtained a FLAG-β-cat /pcDNA3 plasmid and generated 2 basic mutants T653R, T653K and a neutral one T653A as control. All 4 constructs express similarly in transfected HEK 293T cells. Consistent with previous studies[144], no Y654 phosphorylation was observed at baseline using an anti-β-cat pY654 site-specific phosphoantibody (Fig. 5-3a). Co-transfection with constitutively active Src kinase (Src/CA) led to robust detection of WT β-cat with the β-cat-pY654 specific antibody (Fig. 5-3b). Unfortunately, this antibody was unable to distinguish phosphorylation of our mutant β-cats where the adjacent T653-residue was altered, because the recognized epitope is not solely determined by pY, but by adjacent residues as well (A-**T**-pY-A-A). To get around issues with the β-cat-pY654 antibody, we sought to use a general phospho-tyrosine antibody (clone PY20). We found, however, that even a pan-pY antibody recognized WT and T653A mutant β-cat proteins better than the T653K/R basic mutants (Fig. 5-3c), consistent with evidence that the PY20 antibody also disfavors substrates with basic residues adjacent to pY.[150] Thus, despite being a pan-phosphotyrosine antibody, PY20 shows lower binding to pY residues preceded by basic amino acids, which means that estimation of pY abundance by immunoblotting under-represents the true level of phosphorylation for these protein sequences.

While the pan-PY20 antibody fails to recognize all pY sequences equally, it does manifest demonstrable affinity for all of our β-cat constructs. Thus, we reasoned we could use excess PY20 antibody to quantitatively immunoprecipitated most pY proteins (including exogenously expressed FLAG-tagged β-cats) and infer β-cat pY status by immunoblotting with the FLAG antibody, where the FLAG epitope is identical across all constructs (Fig. 5-3d). With this

approach, we found that β-cat T653R and T653K mutants show more phosphorylation by Src/CA than WT and T653A, consistent with our *in vitro* peptide library profiling results (Fig. 5-3e). To explore the role of phosphatase regulation of this site, we co-transfected SHP1, a known PTP to dephosphorylate β-cat pY654[144], along with β-cat constructs and Src/CA. We found that while the amount of phospho-β-cat was reduced, the basic mutants remained more phosphorylated than WT (Fig. 5-3f). Importantly, treatment with the phosphatase inhibitor, sodium pervanadate ($Na_3VO_4$), eliminated phosphorylation differences detected between WT and mutant β-cats (Fig. 5-3g). These data strongly suggest that it is differential phosphatase activity towards β-cat pY654 versus WT β-cat that drives phospho-abundance differences, where PTP activity can be restricted by presence of basic residues adjacent to the pY site.

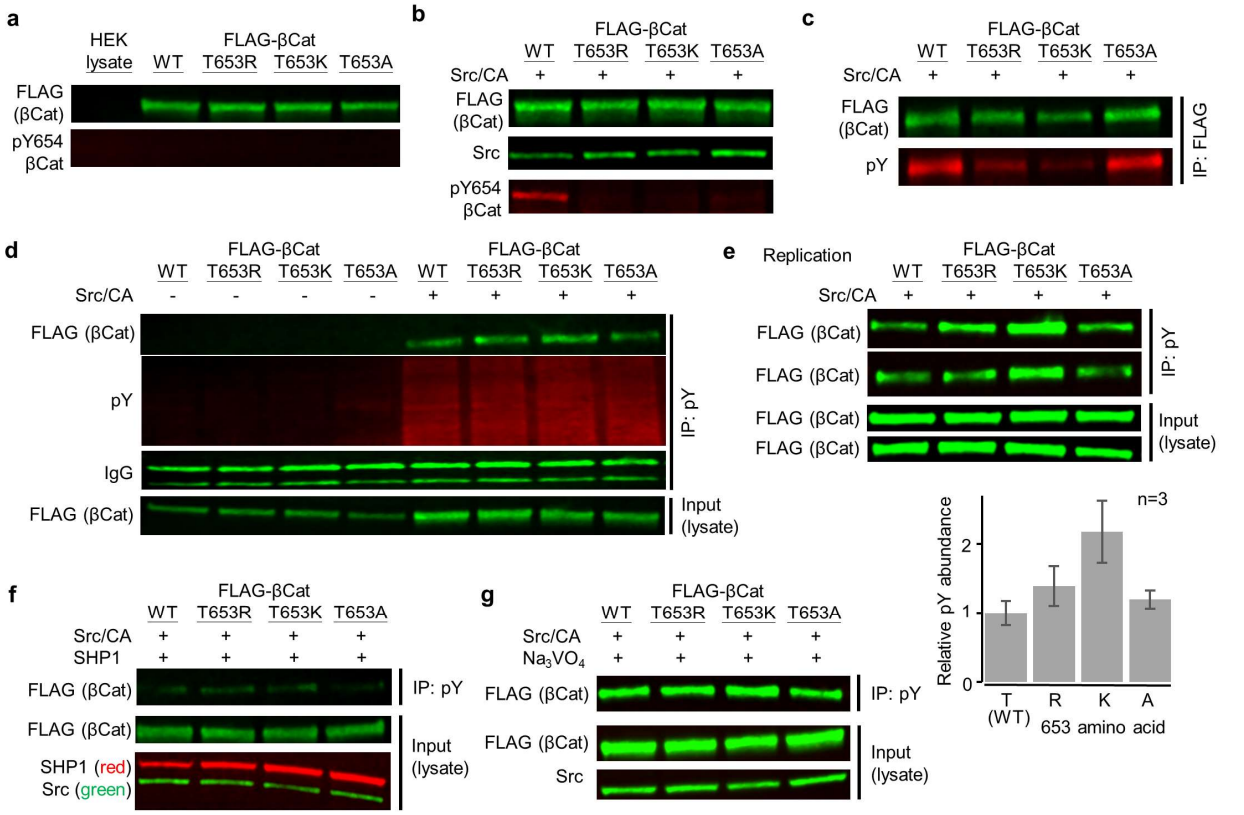



**Figure 5-3 β-cat phosphorylation at Y654 is enhanced by basic amino acid replacement at T653.**

**a**, Four β-cat constructs were transfected and expressed equally in HEK 293T cells. **b**, Co-transfection of Src/CA promotes β-cat Y654 phosphorylation. Site-specific antibody does not recognize mutants. **c**, Equal amount of lysates (1 mg) were immunoprecipitated with a FLAG antibody. Pan-phosphotyrosine antibody displays different binding affinity to WT and mutants. **d**, β-cats were immunoprecipitated with pY and blotted for FLAG to quantify their Y654 phosphorylation. **e**, In biological triplicate experiments, both basic β-cat T653 mutants are more phosphorylated than WT. **f**, β-cat pY654 is a substrate for SHP1. **f**, Inhibiting PTPs with pervanadate (100 μM) eliminates the difference in Y654 phosphorylation between WT and mutants.

### *5.2.6. β-cat Basic Charge Mutants Manifest Enhanced Wnt-Signaling.*

To address the role that β-cat T653K/R mutants play in Wnt-signaling, we quantified AXIN2 levels as an indicator, because it is a universal target gene and a negative-feedback regulator of the pathway.[151] We used the CRISPR-Cas9 strategy to generate a HEK β-cat$^{KO}$ cell line (See section 5-5 for the generation and characterization of the knockout cells) and selected clones lacking AXIN2 induction upon treatment with the Wnt pathway agonist and GSK3 inhibitor, LiCl (Fig 5-4a, clone 16).[151] FLAG-tagged constructs were transfected into the β-cat$^{KO}$ cells, serum-starved overnight and then fed with 5% FBS for 6 h before harvesting. We found AXIN2 was largely depleted after transfected cells were serum starved (Fig. 5-4b, first lane) and increased only after cells were supplied with FBS, allow us to effectively synchronize our cells. Using this method, we found that the T653R and T653K mutants expressed significantly more AXIN2 (Fig. 5-4b). These findings are consistent with the elevated pY654 phosphorylation observed in Fig. 5-3e, and demonstrate that the more phosphorylated basic mutants contribute to elevated β-cat signaling. To show that phosphatase activity was required for the different levels of β-cat signaling, we treated cells with 100 μM $Na_3VO_4$ to inhibit PTPs and found that WT β-cat induced AXIN2 as much as the basic mutants (Fig. 5-4c). Altogether, these experiments definitively demonstrate that disease-associated missense mutant forms of β-cat (T653K/R) display enhanced Wnt-signaling activity through a positive-charge gating (or steric hindrance) mechanism that restricts tyrosine phosphatase substrate selectivity at pY654.

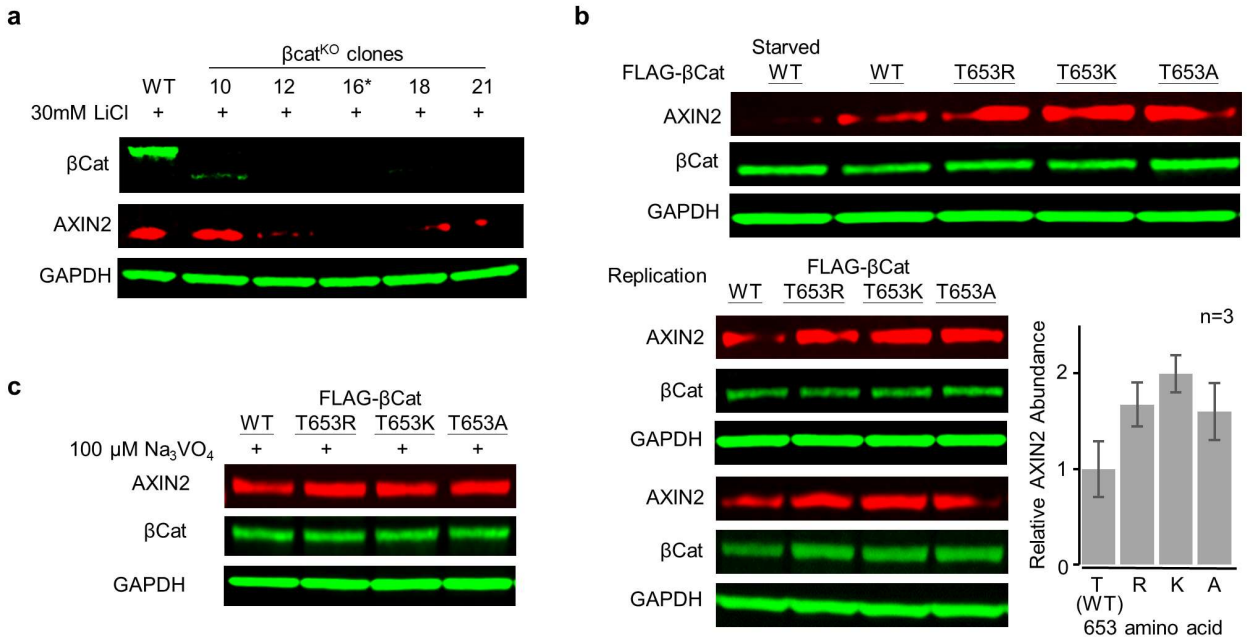



**Figure 5-4 β-cat T653 basic charge mutants display enhanced Wnt-signaling.**

**a,** HEK β-cat$^{KO}$ cells generated with CRISPR-Cas9 displays a lower AXIN2 expression level treated with LiCl. Clone 16 was used for the studies. **b**, In biological triplicate experiments, T653R and T653K β-cats induce more AXIN2 protein expression than WT β-cat. **c**, PTP inhibition with pervanadate (100 μM) eliminates differences in AXIN2 levels between WT and mutant β-cats.

## *5.3. Discussion and Conclusion*

This chapter describes a general substrate selectivity rule using peptide arrays and SAMDI-MS to discover that basic residues adjacent to pY antagonize PTP dephosphorylation. We applied this rule to the canonical signaling/adhesion protein, β-cat, where missense mutations associated with disease were known but incompletely understood. Specifically, we describe a charge mutation/PTP-restriction (or gating mechanism) surrounding β-cat Y654 and neighboring residues that contributes to Wnt-pathway gain-of-function phenotypes *in vivo*. β-cat basic mutants, T653R and T653K, are both found in cancer database and a mouse phenotype with intellectual disability caused by the T653K mutation was previously identified.[146] We showed that both basic mutants are more phosphorylated and lead to high levels of AXIN2, suggesting elevated Wnt signaling compared to WT β-cat (Fig. 5-5). Importantly, this difference is eliminated when cells are treated with a phosphatase inhibitor $Na_3VO_4$, suggesting that phosphatases differentially target mutant versus WT β-cat. Our model now explains the neuro-developmental phenotypic similarities between Y654E phospho-mimic and T653K *bfc* mutant mice.[145-146] As phosphatases are less efficient in dephosphorylating pY654 next to T653K basic residue, the mutant favors persistent phosphorylation and elevation of β-cat-mediated gene expression, rationalizing its resemblance to a permanent phospho-mimic in the developmental process. Our study also points to a potential SHP1 activation treatment strategy through induction of SHP1 expression or small molecule activators.[152]

This study also shows how SAMDI-MS combined with peptide arrays is a powerful and versatile platform to discover general rules for enzyme-substrate specificities and the findings are generally applicable to protein substrates. Remarkably, the rule we describe—where PTPs

disfavor basic residues proximal to pY—is surprisingly universal and generalizable. We validated this rule in two cases with β-cat (this work) and insulin receptor[137] where basic residues are added or removed in missense mutations proximal to pY to affect phosphorylation status and signals that drive disease development. Our database search identified over 6,000 relevant cancer mutations that could employ this missense mutation/PTP restriction mechanism. This opens up the possibility that phosphatases may play a much larger role in cancer development than that previously appreciated.

In summary, the phosphatase-centered basic mutation/dephosphorylation crosstalk mechanism provides a novel way to view the phospho-regulation of proteins. We expect this to enable new understanding of diseases and inspire more diverse treatment strategies.

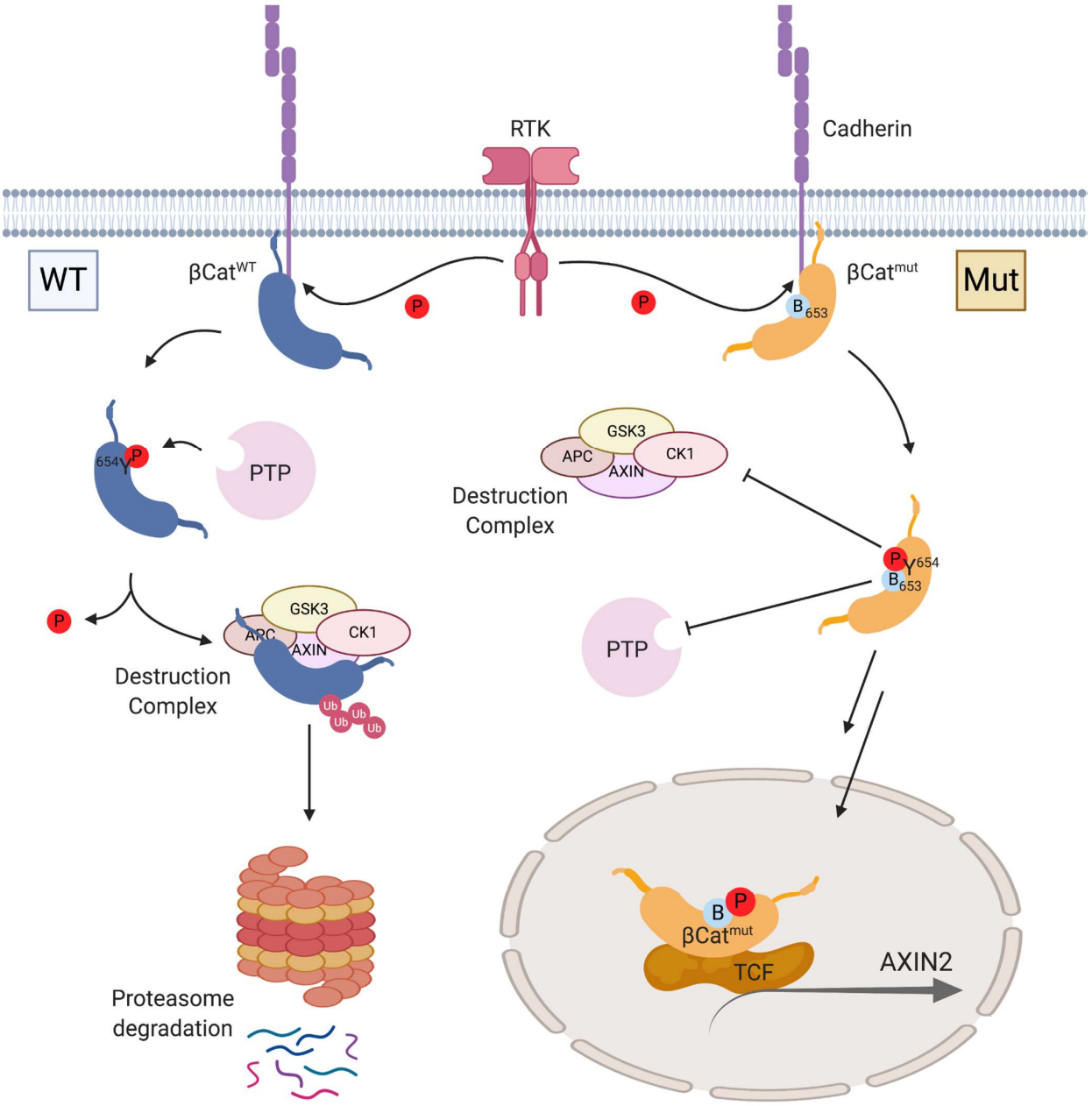



**Figure 5-5 β-cat charge-mutation/PTP-restriction mechanism that enhances Wnt signaling.** (Left) WT pY654 is dephosphorylated by PTPs and degraded. (Right) Basic mutants inhibit pY654 dephosphorylation. Sustained phosphorylation of β-cat reduces association with cadherins and leads to enhanced transcriptional co-activation with T-cell factor (TCF), revealed by the canonical target and negative regulator, *AXIN2*. Schematic created with Biorender.com.

## *5.4. Methods*

**Profiling SHP1 activity.** Peptide library and SAMDI plates were prepared and immobilized as previously described.[99] SHP1 (SRP5076, MilliporeSigma) was diluted to 14 nM in PTP buffer (100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and 2 μL was applied to each gold spot on the peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated for 4 h at 37 ºC in a humidified chamber. After the reaction was complete, the plate was rinsed with water and ethanol, treated with 1 μL matrix (10 mg/mL THAP, 5 mg/mL ammonium citrate dibasic in 50% acetonitrile, 50% water and 0.1 % phosphoric acid) to each spot and dried in air for 20 min. The spots were analyzed by MALDI-TOF MS (AbSciex 5800) to obtain a mass spectrum for each reaction. Enzymatic activities were quantified by measuring the areas under the curve (AUCs) for the dephosphorylate product peak and the substrate peak and determining the activity (%) = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %. Activity heatmaps were generated by Microsoft Excel.[137]

**Database search.** Phosphotyrosine sites were downloaded from PhosphoSitePlus and single-nucleotide variations (SNVs) in cancer were downloaded from BioMuta.[147-148] The cancer mutations at +1 and -1 positions to a phosphotryrine that added or removed a basic residue (R or K) were extracted. Database search was performed by Dr. Lihua Zou.

**Plasmids and site-directed mutagenesis.** β-cat/pcDNA3 plasmid was a generous gift from Dr. Eric Fearon. Site-directed mutagenesis kits (New England Biolabs) were used to generate β-cat/pcDNA3 T653R, T653K and T653A mutants following the manufacturers protocol. Primers [5'-AGG TGT GGC GAG ATA TGC AG CTG CTG-3' (R forward), 5'-AGG TGT GGC GAA

ATA TGC AGC TGC TG-3' (K forward), 5'-AGG TGT GGC GGC ATA TGC AGC TGC TG-3' (A forward), and 5'-TCA TTC CTA GAG TGA AGT AAC TCT GTC AGA G-3' (reverse)] were purchased from Integrated DNA Technologies. Src/CA/pDEST40 (#124659) and SHP1/pJ3 (#8572) plasmids were purchased from addgene. All plasmids were transformed to DH5α *E. coli.* (New England Biolabs), amplified and purified with DNA maxiprep kits (Invitrogen). All sequences were confirmed by Sanger Sequencing (ACGT, Inc.).

**Antibodies.** The following primary antibodies were used in this study: mouse anti-FLAG (F3165, MilliporeSigma), rabbit anti-FLAG (F7425, MilliporeSigma), rabbit anti-pY654-β-cat (ab59430, abcam), rabbit anti-Src (701396, Invitrogen), mouse anti-phosphotyrosine (clone PY20, P4110, MilliporeSigma), mouse anti-SHP1 (MA5-11669, Invitrogen), rabbit anti-β-cat (clone RM276, SAB5600086, MilliporeSigma), mouse anti-β-cat (610153, BD Bioscience), mouse anti-γ-catenin (610253, BD Bioscience), rabbit anti-AXIN2 (PA5-25331, Invitrogen) , mouse anti-GAPDH (CB1001, MilliporeSigma)

**Cell culture.** HEK 293T cells were obtained from American Type Culture Collection (ATCC) and maintained in Dulbecco's Modified Eagle's Medium (DMEM, Corning), containing 10% fetal bovine serum (FBS, Atlanta Biologicals), 100 U/ml penicillin and 100 mg/ml streptomycin (Corning) at 37 ºC and 5% $CO_2$.

**Transfection and cell lysis.** For one reaction, plasmids (1 μg each) and Lipofectamine 2000 (3 uL, Invitrogen) were diluted in 125 μL Opti-MEM (Corning) respectively and mixed. The mixture was incubated for 30 min at room temperature. HEK cells were plated in 6 well-plates at a seeding density of 1.0 x $10^6$ cells in 1.25 mL Opti-MEM. The reaction mixture was

added to the cells and incubated for 3 h at 37 ºC and 5% $CO_2$. DMEM completed with FBS and antibiotics (1.5 mL) were added. The cells were incubated for two day, washed with PBS and harvested in cell lysis buffer [50 mM Tris, 150 mM NaCl, 5 mM EDTA, 1% Triton X100, pH 7.5, cOmplete™ Protease Inhibitor (1 tablet/ 10 mL, Roche), PhosSTOP™ (1 tablet/ 10 mL, Roche), 200 uM $Na_3VO_4$.]. Lysates were sonicated and centrifuged at 14,000 rpm for 10 min to remove debris. The concentrations were determined by Bradford assay (5000006, Bio-Rad). For AXIN2 immunoblotting, the transfected cells were serum-starved overnight and incubated with DMEM complete for 6 h before lysis.

**Immunoprecipitation.** Primary antibodies (10 μg) were added to 1 mg cell lysate diluted to 0.5 mL with cell lysis buffer and incubated overnight at 4 ºC under agitation. Protein A-agarose beads (~50 μL, MilliporeSigma) were added and incubated for 1 h at 4 ºC under agitation. Immunocomplexes were collected by centrifugation and washed four times with ice-cold cell lysis buffer. Proteins were solubilized with 30 uL 3x SDS loading buffer (150 mM Tris, 4.5% SDS, 240 mM DTT, 0.12% Orange G, 30% glycerol, pH 7.5) and separated by SDS-PAGE.

**Immunoblotting.** Lysates or IP eluates were resolved by 4-20% Criterion TGX Midi gels (Bio-Rad) and transferred to nitrocellulose membranes (GE) following the manufacturer's protocol. Membranes were blocked in Intercept (TBS) blocking buffer (LI-COR) for 1 h at room temperature. Membranes were incubated overnight with primary antibodies at 4 ºC and then with fluorophore-conjugated secondary antibodies (LI-COR) for 1 h at room temperature. All antibodies were used at the dilutions suggested by the manufacturers. The bands were visualized by LI-COR Odyssey Fc Imaging System. The images were processed with ImageJ.

**HEK *CTNNB1* Knockout with CRISPR-Cas9.** Guide RNA (gRNA) targeting different human *CTNNB1* exons were designed with CHOPCHOP online tools.[153] The sequences of oligonucleotides were as follows: 5'-TGAGTAGCCATTGTCCACGC-3' (exon 1), 5'-CTAACAGCCGCTTTTCTGTC-3' (exon 2), 5'-CTGTCTTTTCGCCGACAATC-3' (nonsense), and 5'- CAACAGTCTTACCTGGACTC-3' (exon 3). Single guide RNA (sgRNA), Cas9 nuclease (HiFi) and duplex buffer were purchased from Integrated DNA Technologies. RNAs were reconstituted and diluted to 5 μM with duplex buffer; Cas9 protein to 10 μM with PBS. HEK cells were plated in 12-well plates at a seeding density of 1.0 x $10^5$ cells in 1 mL DMEM-complete a day before. For one reaction, sgRNA (20 μL), Cas9 (15 μL), DMEM (30 μL, no serum and antibiotics) and Lipofectamine RNAiMax (4 μL, Invitrogen) were mixed and incubated at room temperature for 20 min. The complex was treated to the cells in complete medium and incubated overnight at 37 ºC and 5% $CO_2$. The medium was changed and the cells were allowed to recover for one day. Cells were split to maintain 30-50% confluency. The sgRNA-Cas9 treatment was repeated 3 times. Cells were expanded and sorted with a flow cytometer (FACSMelody, BD) to 96-well plates to grow single cell colonies. The colonies were screened for low (< 5 ng/well at confluent) β-cat expression using a β-cat ELISA kit (Invitrogen). The selected colonies were verified with western blot. The knockout cell line was maintained in the same conditions as WT.

**Immunocytochemistry.** Microscope cover glasses (12 mm, Fisher) were flamed and coated with 1 mL poly-D-lysine (0.1 mg/mL, Gibco) in a 12 well plate for 1 h at room temperature. HEK cells (WT, β-cat$^{KO}$ or 1:4 mixture) were seeded at at a seeding density of 1.0 x $10^5$ cells and incubated for 3 days. Cells were fixed in 4% paraformaldehyde (PFA) in PBS for 30 min,

quenched with 100 mM glycine, permeabilized with 0.3% Triton X-100, and blocked with 3% normal goat serum (MilliporeSigma). Primary (1:100, rabbit anti-β-cat, clone RM276, MilliporeSigma) and secondary (1:300, Alexa Fluor 568 goat anti-rabbit, MilliporeSigma) antibody incubations were performed at room temperature for 1 h, Hoechst (1:10,000, ThermoFisher) was incubated for 5 min, interspaced by multiple washes in PBS, and followed by mounting coverslips in ProLong Gold fixative (Life Technologies). Fixed cells were imaged on an Axioplan 2 microscope (Carl Zeiss) at room temperature using a x40 objective lens. Captured images were processed in ImageJ.

## *5.5. Creating β-cat$^{KO}$ HEK 293T cell lines using CRISPR-Cas9*

*This work was performed in collaboration with Dr. Alex Yemelyanov and Annette S. Flozak.

Guide RNAs targeting 3 different human *CTNNB1* exons (guide 1, 2, 4, Fig. 5-6a) as well as a nonsense sequence (guide 3, reverse of guide 2) were designed using CHOPCHOP online tools.[153] Wildtype HEK 293T cells were transfected with each sgRNA-Cas9 complex overnight 4 times and the bulk cells were expanded and blotted for β-cat (Fig. 5-6b). Guide 2 and 4 showed about 50% knockout in bulk cells whereas those treated with exon 1 and the nonsense guides had β-cat expressed like WT. The result was consistent with previously reported efforts creating β-cat$^{KO}$ cell using the same gRNA sequences.[154] Cells treated with gRNA 2 and 4 were then sorted to 96-wells plates (1 cell/well) with a flow-cytometer, expanded and primarily screened with ELISA for low-β-cat expressing colonies. These colonies were further expanded and blotted for β-cat (Fig. 5-6c). Several colonies from both guides showed a complete knockout on western blots (rabbit anti-β-cat, clone RM276). We also observed that clone 16 and 18 had a more significant upregulation of plakoglobin (γCat), which was previously identified to rescue the junction in β-cat$^{KO}$ cells.[155]

We next used an AXIN2 antibody to verify the β-cat$^{KO}$ cells. To avoid species overlap, we switched to a mouse β-cat antibody (BD). To our surprise, some clones previously appeared to be β-cat null (anti-β-cat-RM276) showed a fragment below 75 kD using the BD antibody (Fig. 5-6d). This was most likely a c-terminus fragment as the antibody was generated with the immunogen β-cat a.a. 571-781. The fragment was also active to turn on Wnt pathway and induce AXIN2 expression with or without the Wnt pathway agonist and GSK3 inhibitor, LiCl.[151] We

identified that clone 16 was a complete β-cat$^{KO}$ cell line without c-terminus fragments and background AXIN2 expression (Fig. 5-6e). The cell line was further verified for knock by immunocytochemistry. A separate culture of WT and KO cells and staining with a β-cat antibody showed a distinct difference where β-cat was observed at the junctions in WT but not in KO (Fig. 5-6f, first two images). A co-culture of WT and KO cells (Fig. 5-6f, last image) clearly demonstrated on the same slide that β-cat only distributed at the junctions in WT (upper right area) but not KO (lower left). HEK 293T β-cat$^{KO}$ clone 16 was used for the further studies in this chapter.

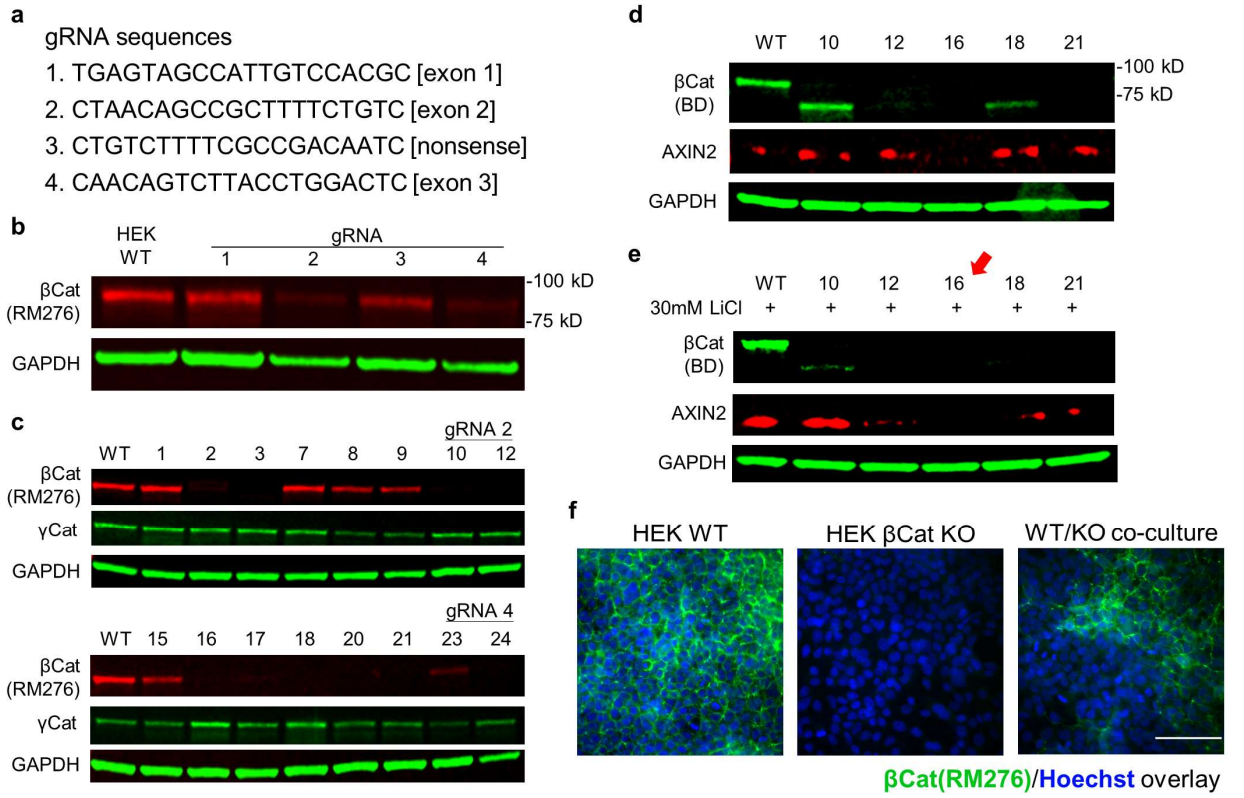



**Figure 5-6 Creating β-cat$^{KO}$ HEK 293T cell lines and their characterization.**

**a**, gRNA sequences used for CRISPR-Cas9 β-cat knockout. **b**, Bulk cells blotted for β-cat after sgRNA-Cas9 complex treatments. **c**, Validation of knockout by single cell colonies. **d**, Validation using a c-terminus β-cat and AXIN2 antibodies. **e**, Screening for AXIN2-null cells with LiCl (30 mM) treatment. **f**, Separate and co-culture of WT and β-cat $^{KO}$ HEK 293T cells. Scale bar: 50 μm.

# Chapter 6. A Novel Citrullination/Phosphorylation Crosstalk via Protein Tyrosine Phosphatases

*This work was performed in collaboration with Dr. Jennifer Grant. This chapter is reproduced in part from Huang, C.-F.; Grant, J.; Mrksich, M. A Novel Citrullination/Phosphorylation Crosstalk via Protein Tyrosine Phosphatases. *Manuscript in prep.*

## *6.1. Introduction*

Protein tyrosine phosphatases (PTPs) and protein tyrosine kinases (PTKs) regulate phosphorylation levels by removing or adding phosphate group on proteins. The number of human PTPs and PTKs are roughly equal (90 PTKs and 107 PTPs), however, phosphatases are often stigmatized as non-regulatory "housekeeping enzymes" and receive disproportional research attention compared to kinases.[4-6] Emerging work further demonstrates that PTP dysregulation is linked directly to several diseases, which emphasize the importance of studying PTP activity.[7-10] For example, SHP2 is identified as the first oncogenic phosphatase[11-13], and PTP1B contributes to the development of insulin resistance and type 2 diabetes.[134] These discovery draw new momentum to phosphatase drug development and signaling studies[30], including how PTPs can mediate modification crosstalk between different residues described in this work.

Despite these advancements in phosphatase biology, there are still limited assays to directly monitor PTP activities, making it difficult to discover novel crosstalk mechanisms. Most phosphatase assays often rely on non-specific small molecule substrates[87-88], which are not effective mimics of native protein PTP substrates and prevent studies of structure-activity relationships. Peptide probes are gaining popularity for their sequence specificity and ability to profile combinatorial arrays of unique sequences, providing insights for PTP-substrate specificity.[115-116, 135-136] We developed a label-free approach using SAMDI-MS (self-assembled monolayers for matrix-assisted laser desorption/ionization mass spectrometry) to measure PTP activities on peptides.[99] This assay is extremely sensitive that allows single-cell measurements[156], and high-throughput (~$10^5$ samples/day).[157] Combined with peptide arrays, we profiled 22 PTPs

and found most lack activity towards substrates that have an arginine (R) or lysine (K) residue directly adjacent to either side of the phosphotyrosine (pY).[137] This substrate specificity was confirmed in lysates from five different cell lines.[138] We recently demonstrated that charge-substitution mutations proximal to pY sites, could impact a range of signaling proteins and disease pathology through a crosstalk mechanism. In one example[137], our work suggested that neutralizing the positive charge on R1152 of an insulin receptor reduces phosphorylation of Y1151, which may contribute to type 2 diabetes.[130] In chapter 5, we showed that β-catenin (β-cat) cancer mutations introducing a basic residue adjacent to β-cat Y654 (T653R and T653K) antagonizes its dephosphorylation by SHP1, favoring persistent phosphorylation and elevation of the canonical β-cat-mediated Wnt-signaling, which ultimately leads to cancers.[145-146, 149] Therefore, the crosstalk mechanism that we reported appears to be ubiquitous, and may have undiscovered roles in cell signaling and other pathologies. These examples also demonstrate that SAMDI-MS and peptide arrays together are platforms that can discover crosstalk relationships between residues that are derived from protein substrates.[62, 75-76]

Protein citrullination is a post-translational modification (PTM) that replaces the basic guanidino group on an arginine residue with a neutral uriedo functionality.[158] Altering local amino acid charge states can induce different protein conformations, modulate protein-protein interactions, and send important cellular signals.[159] For example, negatively-charged actin filaments bind less to lipid vesicles as their negative surface charge increased due to phosphorylation or deimination of myelin basic protein (MBP), the most abundant protein components of the myelin membrane in the central nervous system (CNS).[160] Therefore,

understanding how phosphorylation and citrullination are cross-regulated would advance our insights in CNS-related disease such as multiple sclerosis.[161-162]

In this chapter, we describe a novel citrullination/phosphorylation crosstalk via PTPs. We demonstrate that an arginine residue proximal to a phosphotyrosine residue restricts PTP dephosphorylation. Removal of the basic residue by protein-arginine deiminases (PADs) citrullination restores PTP activity and shifts the substrate toward dephosphorylation. We report that citrullination at MBP R264 increases its neighboring pY261 dephosphorylation by 2.3-fold in $k_{cat}/K_M$ as a proof-of-concept where these modifications crosstalk and could generate a cascade of regulatory signals.

## *6.2. Results and Discussion*

### *6.2.1. Basic Residues Antagonize PTP Activity*

We previously profiled 22 PTP specificities on a peptide array with the sequence Ac-TRDXpYZTC-$NH_2$, where the X and Z positions comprise each of the 19 canonical amino acids except cysteine.[137] The array was synthesized and immobilized to a 384-gold-spot plate, treated with PTP solutions and analyzed with SAMDI-MS (Fig. 6-1A). Conversion of phosphopeptide to its product was characterized by integration of the corresponding peaks, where Activity = Area Under the Curve $(AUC)_{product}$ / $(AUC_{substrate} + AUC_{product})$ x 100 %. Activities for each peptide sequence were represented in a 19 x 19 heatmap where each column defined the amino acid in the X (-1) position and each row defined the amino acid in the Z position (+1). Activity was represented by a color scale with white corresponding to 0 % and purple as 100 %. A representative SHP2 activity heat map is shown in Fig. 6-1B, where basic residues (R and K) inhibit SHP2 activity at both X and Z positions, generating noticeable white lighter rows and columns of corresponding amino acids on the heatmap. We demonstrated that basic residues proximal to pY universally restrict all PTP dephosphorylation[137], and this discovery can be applied to protein substrates (chapter 5), making SAMDI-MS combined with peptide arrays an extremely powerful platform to discover crosstalk relationships between residues.

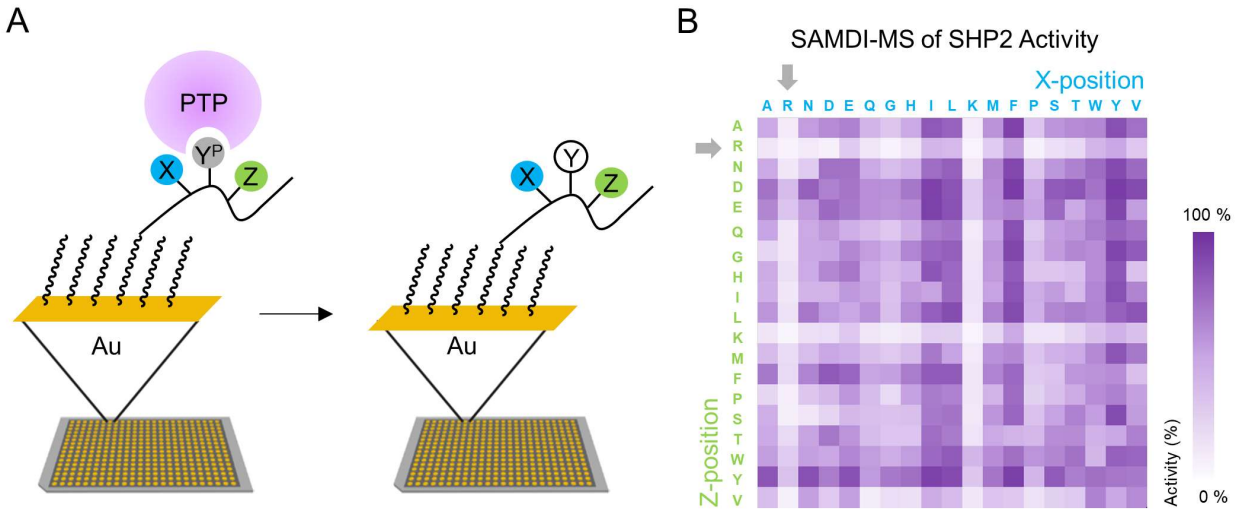



**Figure 6-1 Profiling phosphatase activity using peptide arrays and SAMDI-MS.**
A) Phospho-peptides immobilized on a self-assembled monolayer surface are treated with PTP and dephosphorylation of each peptide is analyzed with SAMDI-MS. B) Heatmap for SHP2 shows basic residues (R and K) adjacent to phosphotyrosine inhibit enzymatic activity. Activity measured by SAMDI-MS is defined as $(AUC)_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %.

### *6.2.2. Citrullination Restores PTP Activity.*

The SHP2 heatmap (Fig. 6-1B) clearly shows peptides containing a basic arginine (R) residue at either X or Z position exhibit lower dephosphorylation activity. Citrullination is a canonical PTM catalyzed by protein-arginine deiminases (PADs) which removes a basic R residue through guanidine deimination, leaving a neutral uriedo group (Fig. 6-2A). On substrates where an arginine is adjacent to pY, citrullination eliminates the local positive charge that is inhibitory to PTP activity and therefore allows more efficient PTP dephosphorylation, creating a crosstalk between two modifications (Fig. 6-2B). To demonstrate that citrullination restores PTP activity, we measured SHP2 activity on peptides with an R on either side of pY (Fig. 6-2C, blue) and compared to those with a citrulline (Cit) replacing the R. We found that either by

synthetically incorporating Cit (Fig. 6-2C, light green) or enzymatically modifying R to Cit by PAD1 (Fig. 6-2C, dark green), citrullination increases PTP dephosphorylation, confirming the crosstalk relationship. We note that this mechanism universally applies to most PTPs and not just specific to SHP2 because all PTPs share the evolutionary conserved active site structure and disfavor basic residues as we previously described[137], making the citrullination/phosphorylation crosstalk relationship via PTPs extremely generalizable and valuable in cellular phospho-regulation.

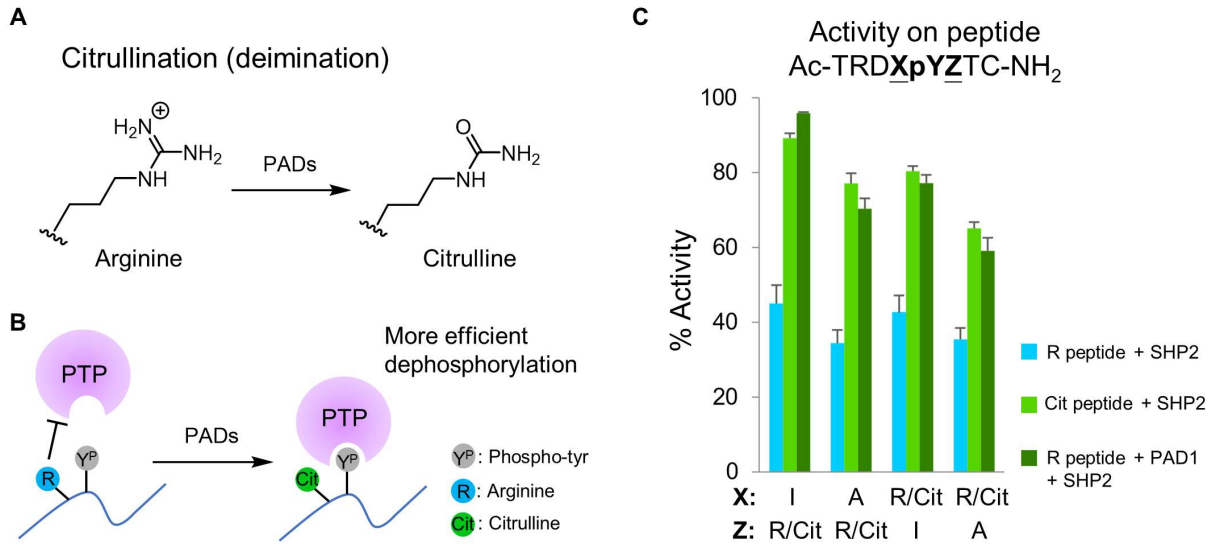



**Figure 6-2 Citrullination increases PTP activity.**

A) Citrullination by PADs removes the positive charge on arginine and replaces with a neutral uriedo group. B) Arginine inhibits PTP activity while citrulline allows more efficient dephosphorylation. C) Peptides containing a citrulline exhibit higher PTP dephosphorylation activity than those containing an arginine. Activity is measured by SAMDI-MS.

### *6.2.3. MBP Modification Crosstalk Between Y261 And R264.*

We next identified a protein substrate that could utilize the described crosstalk to regulate its phospho-state. Myelin basic protein (MBP) has been reported to have multiple phosphorylation and citrullination sites that alters its charge state and regulates lipid vesicles' affinity to actin-filament.[160] It has a phosphotyrosine site (Y261) and a citrullination site (R264) separated by two amino acids.[147] The pY261-phosphopeptides were synthesized with R and Cit at 264-residue and their SHP2 activities were compared (Fig. 6-3). We observed that Cit264 peptide is more prone to dephosphorylation, consistent with the proposed crosstalk mechanism and prior studies suggesting that long-range electrostatic interactions, which is responsible for PTP substrate selectivity against basic residues, could extend to 5 residues up/down-stream.[115]

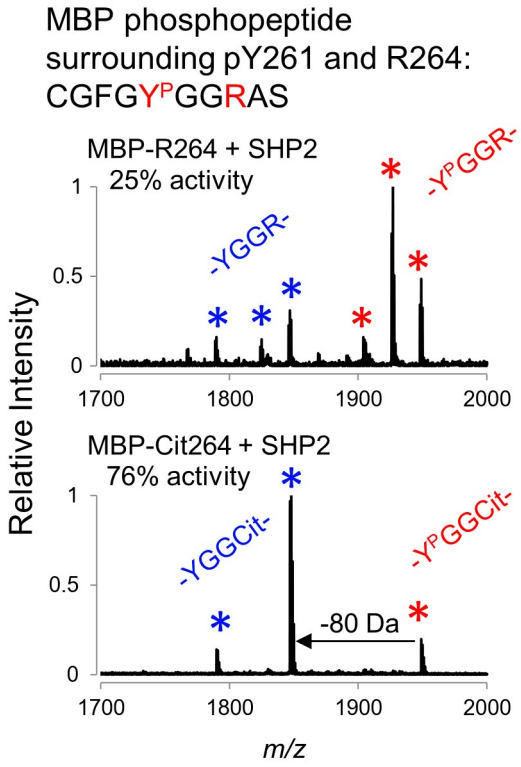



**Figure 6-3 Dephosphorylation of MBP peptides.**

MBP pY261 dephosphorylation is antagonized by R264 (25% activity, top spectrum) while the activity is restored by Cit264 (76% activity, bottom spectrum). On SAMDI-MS spectra, red asterisks indicate phosphopeptides and blue indicate corresponding dephosphorylated products separated by -80 Da.

### *6.2.4. Citrullination Increases PTP $k_{cat}/K_M$.*

We finally used SAMDI-MS to monitor SHP2 dephosphorylation reaction progression on MBP peptides to describe the crosstalk relationship in kinetic terms. We previously demonstrated that SAMDI-MS is great platform to study enzyme kinetics.[163-165] MBP peptides were incubated with SHP2 at different concentrations. The reactions were quenched at different time points by adding 100 μM bpV(Phen) and crude was immobilized to the SAM surface for analysis. The product concentrations were calculated using the activity. The ionization efficiency were accounted for by comparing the substrate and product peak to the tri(ethylene glycol)-terminated alkanethiolate (EG3) monolayer background peak as previously described.[75] We note that there is no drastic change in ionization efficiency in any modification in this work because peptides are covalently conjugated to the monolayer. The reaction traces for both peptides are showed in Figure 4A. The initial velocities ($V_0$) were calculated and plotted against the substrate concentrations (Figure 4B). We did not observe the saturation of the enzyme-substrate complexes and the convergence of $V_0$ to $V_{max}$ in our given substrate concentration range. We were limited by the solubility of the peptides and therefore could not calculated $k_{cat}$ and $K_M$ separately. However, we could deduct $k_{cat}/K_M$ directly from the slopes of the plots and found them in the $10^6$ $M^{-1}s^{-1}$ magnitude, consistent with the reported SHP2-peptide substrate $k_{cat}/K_M$ in prior studies.[115] We found that citrullination at MBP R264 increase its dephosphorylation $k_{cat}/K_M$ by 2.3-fold, confirming citrullination/phosphorylation crosstalk mechanism driven by PTP substrate selectivity.

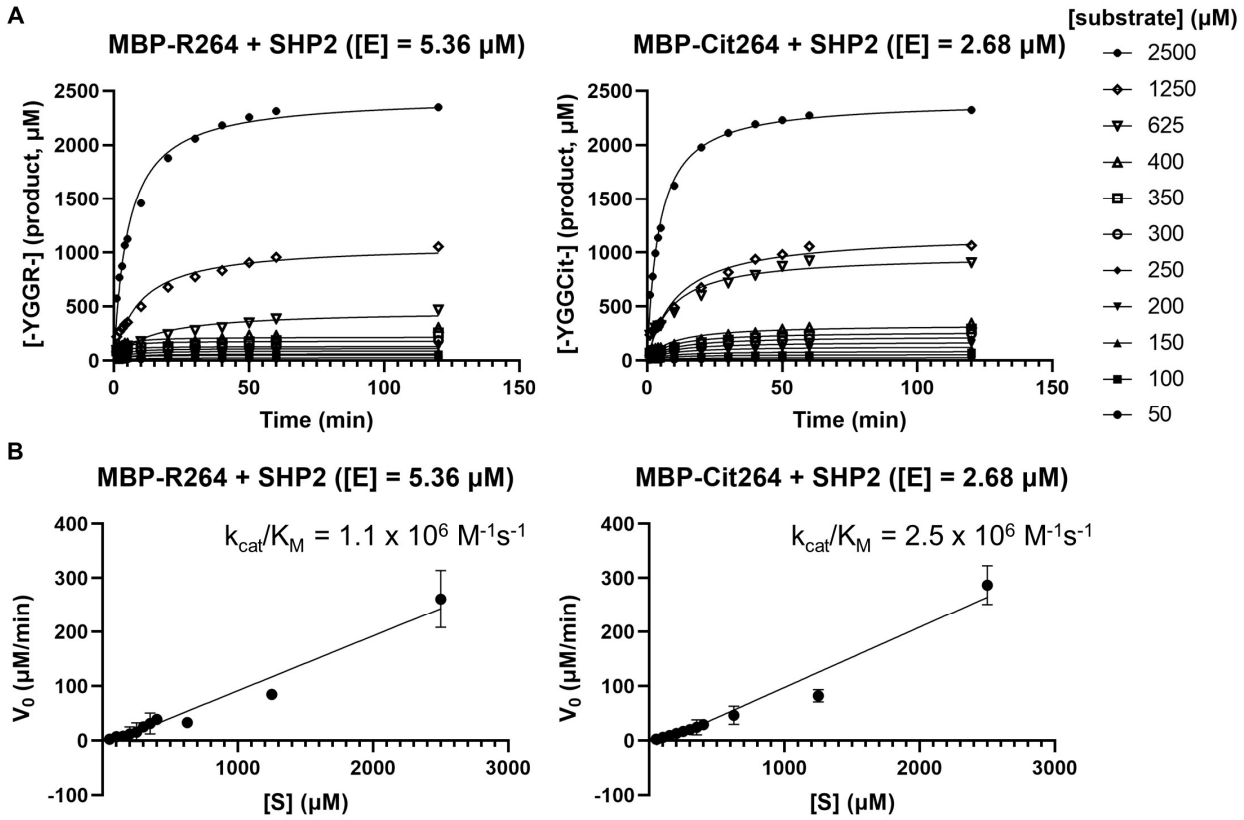



**Figure 6-4 Kinetic comparison of citrullinated/uncitrullinated MBP peptide dephosphorylation.**

A) Reaction curves monitored with SAMDI-MS. B) MBP citrullination increases SHP2 dephosphorylation $k_{cat}/K_M$ by 2.3-fold.

## *6.3. Conclusion*

This work illustrates how peptide arrays and SAMDI-MS can be used to discover crosstalk between amino acid residues through PTMs. We report a novel phosphorylation regulatory mechanism by citrullination of the neighboring arginine. Interestingly, it is phosphatases, not kinases, that play the key regulatory role because the PTP family universally disfavor basic residues nearby their substrates and can utilize this unique selectivity for control. Therefore, this mechanism is extremely generalizable to all PTPs despite most pY sites are unknown for their specific PTP. In addition, this work hints that other PTMs altering basic residues, for example, lysine acetylation, might also be able to regulate their neighboring pY through the same PTP-mediated mechanism. Studying indirect regulation of PTMs has also become an important task in cell signaling and development of disease treatment strategies. For example, a loss-of-function mutation in PAD could result in increasing MBP Y261 phosphorylation due to the R264 PTP inhibition according to our work. However, it could be mistaken as a gain-of-function mutation of its kinase and leads to a kinase inhibitor treatment that only cure the symptoms, not the disease. Together, SAMDI-MS and peptide arrays provide a rapid and reliable platform that would advance our understandings in this important area.

## *6.4. Methods*

**General**. Laboratory chemicals and reagents were purchased from MilliporeSigma and used without additional purification unless specified. Peptide synthesis reagents, including Fmoc amino acids and Rink-amide resin, were purchased from Anaspec. SAMDI-MS was performed on a 5800 MALDI TOF/TOF mass spectrometer (AbSciex) using either manual or automated protocols. A detailed protocol of monolayer plate preparation, peptide synthesis and phosphatase assay can be found in a previously published method paper.[99]

**Plasmids.** SHP2/pcDNA1 plasmid was a generous gift from Dr. Elizabeth A Eklund. The full length SHP2 gene was amplified using polymerase chain reaction (PCR) with the primers 5'-CTAGCTAGCCACCATCACCATCACCATCACCATATGACATCGCGGAGATGGTT-3' (forward) and 5'-CCCAAGCTTCTAAGTGCCTAGCCCTTCCA-3' (reverse) and subcloned to a pET-21d(+) vector using NheI and HindIII restriction enzyme sites. PAD1/ pET-16b plasmid was obtained from Thermo Fisher Scientific- GeneArt.

**Protein Expression and Purification.** To express SHP2, the plasmid was transformed to BL21(DE3) *E. coli* (New England Biolabs) using heatshock method. 2xYT media (5 mL) with carbenicillin was seeded with SHP2 BL21(DE3) *E. Coli* and allowed to grow overnight at 30 °C while shaking at 240 rpm. The next morning, the culture was added to 2xYT media (500 mL) supplemented with carbenicillin and grown at 30 °C while shaking until the OD600 reached 0.5. The culture was induced with 1 mM IPTG for 3 h under shaking at 30 °C. The bacteria were pelleted by centrifugation and lysed in buffer (100 mM Tris pH 7.5, 50 mM NaCl, 5 mM TCEP, 10 % glycerol (v/v), 0.1% Triton X-100, 40 mL) containing one cOmplete-Mini protease

inhibitor tablet. The overexpressed SHP2 contained an N'-terminal His-tag and was purified on a nickel-NTA column. Fractions were eluted with 300 mM imidazole in buffer (100 mM Tris pH 7.5, 50 mM NaCl, 5 mM TCEP), combined, concentrated using a Amicon 50 kDa cutoff centrifugal filter unit and stored at -80 °C in 50% glycerol.

To express PAD1, the plasmid was transformed into BL21DE3pLysS *E. coli* using chemical transformation methods. 2xYT media (5 mL) with carbenicillin and chloramphenicol was seeded with PAD1 BL21DE3pLysS *E. Coli* and allowed to grow overnight at 30 °C while shaking at 240 rpm. The next morning, the culture was added to 2xYT media (500 mL) supplemented with carbenicillin and chloramphenicol, and grown at 30 °C while shaking until the OD600 reached ~0.45. The culture was cooled at 4 °C and PAD1 expression was induced with 0.5 mM IPTG overnight under shaking at 25 °C. The bacteria were pelleted by centrifugation and lysed in buffer (50 mM sodium phosphate (7.5), 500 mM NaCl, 0.5 mM EDTA, 10 % glycerol (v/v), 0.1% Triton X-100, 50 mL) containing Bug Buster (10x, 2 mL) and one cOmplete-Mini protease inhibitor tablet. The overexpressed PAD1 contained an N'-terminal His tag and was purified on a nickel-NTA column. Fractions were eluted with a gradient of 5-500 mM imidazole in 50 mM sodium phosphate (7.5), 500 mM NaCl, 0.5 mM EDTA, 10 % glycerol (v/v), 0.1% Triton X-100. The eluted fractions were combined and concentrated to 1 mL in a 25 mL Amicon 30 kDa cutoff centrifugal filter unit. The PAD1 was further purified using size exclusion chromatography on an Akta FPLC (GE Healthcare) using running buffer (50 mM sodium phosphate (7.5), 500 mM NaCl). Fractions containing PAD1 were confirmed by SDS-PAGE, pooled and concentrated, and stored at 80 °C in 50% glycerol.

**Profiling SHP2 Activity.** SHP2 was diluted in PTP buffer (50 nM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and 2 μL was applied to each gold spot on a peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated at 37 ºC for 1 h in a humidified chamber. After the reaction was complete, the plate was rinsed with water and ethanol, treated with 1 μL matrix (10 mg/mL THAP, 5 mg/mL ammonium citrate dibasic in 50% acetonitrile, 50% water and 0.1 % phosphoric acid) to each spot and dried in air for 20 min. The spots were analyzed by MALDI-TOF MS to obtain a mass spectrum for each reaction. Enzymatic activities were quantified by measuring the areas under the curve (AUCs) for the dephosphorylate product peak and the substrate peak and determining the activity (%) = $AUC_{product}$ / ($AUC_{substrate}$ + $AUC_{product}$) x 100 %. Activity heatmaps were generated by Microsoft Excel.

**Sequential Enzymatic Reaction on the Surface.** Peptides (Ac-TRDXpYZTC-$NH_2$, X, Y = I, A or R, 100 μM in 100 mM Tris, pH 7.5 buffer) were immobilized to a SAMDI plates for 1 h at 37 °C. The surface was rinsed with water and ethanol, and each spot was treated with 2 μL PAD1 (875 nM in 100 mM Tris pH 8.0 with 10 mM $CaCl_2$ and 50 mM NaCl) and incubated at room temperature for 30 min in a humidified chamber. After the citrullination reaction was completed, the surface was rinsed with water and ethanol, and treated with 2 μL SHP2 (50 nM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) and incubated at 37 ºC for 1 h in a humidified chamber. Finally, the surface was rinsed, treated with matrix and analyzed by SAMDI-MS following the same procedure above.

**Kinetic Experiments.** MBP peptides (CGFGpYGGRAS and CGFGpYGGCitAS) were dilute to 11 different concentrations (100-5000 μM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100

μM TCEP, 10 uL each). Equal volume (10 uL) of SHP2 (10.72 μM or 5.36 μM in 100 mM Tris, pH 7.5, 50 mM NaCl, 100 μM TCEP) were added to the peptides. The reactions were terminated at different time points (1-120 min) by adding bpV(Phen) to a final concentration of 100 μM. After all reactions were carried out, 2 uL of the reaction crude was transferred to each spot on a SAMDI plate and incubated for 1 h at 37 ºC to allow peptide immobilized. The surface was then rinsed, treated with matrix and analyzed by SAMDI-MS following the same procedure above. AUC was used to calculate the conversion of phosphopeptides to their corresponding dephosphorylated products. GraphPad Prism 8 software was used to generate reaction trace plots and fit the data to a Michaelis-Menten model of the form:

$$V_0 = k_{cat}[E_0]\frac{[S]}{K_M + [S]}$$

Since we did not observe convergence of $V_0$ in the given substrate concentration range, we can assume $K_M >> [S]$ and use the slope of $V_0$-[S] plots to deduct $k_{cat}/K_M$.

# Chapter 7. Summary and Perspectives

This dissertation presents the combination of SAMDI-MS and peptide arrays to study protein tyrosine phosphatases. Before I joined the Mrksich group, there were already some preliminary works using SAMDI-MS to assay phosphatases. However, challenges including phosphopeptide synthesis and low ionization efficiency made it difficult to conduct high-throughput phosphatase studies. In this dissertation, I present my original work utilizing the high-throughput measurement capacity of SAMDI-MS in conjunction with multi-molecular libraries—including small molecule, peptide, and protein—to study the PTP family and their roles in cell signaling and disease development, an area that is often outshined by its kinase counterparts.

Chapter 2 illustrates the SAMDI-MS high-throughput PTP assay and its application in drug discovery for the oncogenic phosphatase SHP2. We used the assay to screen an FDA-approved compound library with 1018 unique molecules. We identified several potent SHP2 inhibitor and measured their $IC_{50}$s using SAMDI-MS. We applied molecular modeling to study one inhibitor, adapalene, and found that its adamantyl functional group was crucial for inhibition. This chapter demonstrates importance of a reliable, high-throughput and label-free assay like SAMDI in drug discovery, particularly for those enzymes that are notoriously hard to assay.

Chapter 3 addresses a long-standing topic in SHP2 oncology, which is whether SHP2 disease mutants alter their substrate specificity or merely affect their basal activity level. We combined SAMDI-MS with a pY peptide library to profile SHP2 and its mutants. This is the first approach to directly determine their specificity using hundreds of peptide substrates. We also reported changes in profiles caused by disease related mutations, including a wide range of preference to aromatic residues. The results would contribute to better understanding of SHP2 disease pathology.

Chapter 4 broadens the scope of study to profiling the PTP family in the human proteome. We used peptide arrays and SAMDI-MS to profile 22 PTPs and reported distinct PTP classes. We found the general rule that basic lysine or arginine residues on either side of the phosphotyrosine antagonize PTP activity. Using an R1152Q mutant in the insulin receptor as an example, we discussed a new mechanism on how mutations in PTP substrates could lead to disease. We reasoned that PTP activity may be restricted by basic amino acids—generated/removed by missense mutations—adjacent to phosphotyrosines leading to abnormal signaling and disease development.

Chapter 5 presents a cell model to validate this hypothesis. We first identified more than 6,000 cancer mutations involving basic residues adjacent to known phosphotyrosine sites. Using two β-catenin mutants associated with cancer (T653R/K) and a mouse model for intellectual disability (T653K), we showed that T653-basic mutant β-catenins are less efficiently dephosphorylated by SHP1 phosphatase, leading to sustained Y654 phosphorylation and elevated downstream Wnt signals, similar to those observed for Y654E phospho-mimic mutant mice. This model rationalized how basic mutations proximal to phosphotyrosines can restrict counter-regulation by phosphatases, providing new mechanismistic and treatment insights for 6,000+ potentially relevant cancer mutations.

Chapter 6 takes a turn from disease models and demonstrates how a similar mechanism can be applied by normal cells to regulate their phospho-states using other PTMs. We reported a novel citrullination/phosphorylation crosstalk via PTPs. We showed that citrullination of an arginine adjacent to a phosphotyrosine removes the PTP restriction by basic residues, resulting in further dephosphorylation. The crosstalk mechanism was showcased using peptides adapted

from myelin basic protein with known citrullination and phosphorylation sites proximal to each other. Citrullination increased the peptide dephosphorylation $k_{cat}/K_M$ by 2.3-fold. The results provide novel insights for phospho-regulation by the PTP family because they all share the similar active site structures and disfavor basic residues in their pockets.

In summary, my work successfully established a high-throughput platform to assay PTP activity with SAMDI-MS. Combined with peptide arrays, we could directly profile PTP specificities, which revealed their diverse roles in cell signaling and disease development. We used the assay to aid drug discover, understand SHP2 oncology and identify the most general substrate specificity rule among PTPs: basic residues antagonize PTP activities, leading to the discovery of new disease development and PTM crosstalk mechanisms. The results will have a broader significance in advancing the biochemistry of phosphatases and their role in signaling.

Finally, I would like to share some thoughts on future directions as the phosphatase field is thriving and our development of SAMDI-MS and peptide arrays has gone a long way beyond measuring enzyme activities. First, I think there are bioinformatic works to systematically analyze the basic mutants we identified in chapter 6. To date, very little is known about mutations on the substrate. It would be beneficial to understand what pathways, phosphatases and kinases are involved and whether there exist connections. Second, we discussed citrullination/phosphorylation crosstalk via PTP in chapter 7. The similar mechanism should be applicable to other charge-altering PTMs such as lysine acetylation. This would be very relevant to histone biology as phosphorylation and acetylation site are clustered. Third, recent advance in allosteric SHP2 inhibitor has drawn attention to the non-catalytic domains of PTPs to investigate the non-enzymatic interactions. Our group has developed the Pi-SAMDI technology to assay

domain-ligand interactions.[166] The field would benefit from profiling these domains and understanding how they are affected by small molecules or mutations. For example, we reported that some mutations in the PTP-domain alter SHP2 substrate selectivity. However, how N-SH2 or C-SH2 domain mutations affect their specificity remains largely unexplored. Lastly, the methodology described in this dissertation can be generalized for other phosphatases. For example, our group has used phosphoserine and phosphothreonine peptide arrays to profile S,T-phosphatases in cell lysates.[138] It would be interesting to study how their specificity are regulated by different subunits they are associated with.

# References


1. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M., Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006,** *127* (3), 635-48.

2. Hardman, G.; Perkins, S.; Brownridge, P. J.; Clarke, C. J.; Byrne, D. P.; Campbell, A. E.; Kalyuzhnyy, A.; Myall, A.; Eyers, P. A.; Jones, A. R.; Eyers, C. E., Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation. *EMBO J.* **2019,** *38* (21), e100847.

3. Pearlman, S. M.; Serber, Z.; Ferrell, J. E., Jr., A mechanism for the evolution of phosphorylation sites. *Cell* **2011,** *147* (4), 934-46.

4. Alonso, A.; Sasin, J.; Bottini, N.; Friedberg, I.; Friedberg, I.; Osterman, A.; Godzik, A.; Hunter, T.; Dixon, J.; Mustelin, T., Protein tyrosine phosphatases in the human genome. *Cell* **2004,** *117* (6), 699-711.

5. Beck, J. R.; Lawrence, A.; Tung, A. S.; Harris, E. N.; Stains, C. I., Interrogating Endogenous Protein Phosphatase Activity with Rationally Designed Chemosensors. *ACS Chem. Biol.* **2016,** *11* (1), 284-90.

6. Tonks, N. K.; Neel, B. G., Combinatorial control of the specificity of protein tyrosine phosphatases. *Curr. Opin. Cell Biol.* **2001,** *13* (2), 182-195.

7. Hale, A. J.; Ter Steege, E.; den Hertog, J., Recent advances in understanding the role of protein-tyrosine phosphatases in development and disease. *Dev. Biol.* **2017,** *428* (2), 283-292.

8. Hendriks, W. J.; Elson, A.; Harroch, S.; Pulido, R.; Stoker, A.; den Hertog, J., Protein tyrosine phosphatases in health and disease. *FEBS J.* **2013,** *280* (2), 708-30.

9. Tonks, N. K., Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* **2006,** *7* (11), 833-46.

10. Tonks, N. K., Protein tyrosine phosphatases--from housekeeping enzymes to master regulators of signal transduction. *FEBS J.* **2013,** *280* (2), 346-78.

11. Chan, G.; Kalaitzidis, D.; Neel, B. G., The tyrosine phosphatase Shp2 (PTPN11) in cancer. *Cancer Metastasis Rev.* **2008,** *27* (2), 179-92.

12. Loh, M. L.; Vattikuti, S.; Schubbert, S.; Reynolds, M. G.; Carlson, E.; Lieuw, K. H.; Cheng, J. W.; Lee, C. M.; Stokoe, D.; Bonifas, J. M.; Curtiss, N. P.; Gotlib, J.; Meshinchi, S.; Le Beau, M. M.; Emanuel, P. D.; Shannon, K. M., Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. *Blood* **2004,** *103* (6), 2325-31.

13.Tartaglia, M.; Niemeyer, C. M.; Fragale, A.; Song, X.; Buechner, J.; Jung, A.; Hahlen, K.; Hasle, H.; Licht, J. D.; Gelb, B. D., Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat. Genet.* **2003,** *34* (2), 148-50.

14. Cohen, P. T., Protein phosphatase 1--targeted in many directions. *J. Cell Sci.* **2002,** *115* (Pt 2), 241-56.

15. Denu, J. M.; Dixon, J. E., Protein tyrosine phosphatases: mechanisms of catalysis and regulation. *Curr. Opin. Chem. Biol.* **1998,** *2* (5), 633-41.

16. Brandao, T. A.; Johnson, S. J.; Hengge, A. C., The molecular details of WPD-loop movement differ in the protein-tyrosine phosphatases YopH and PTP1B. *Arch. Biochem. Biophys.* **2012,** *525* (1), 53-9.

17. Mohebiany, A. N.; Nikolaienko, R. M.; Bouyain, S.; Harroch, S., Receptor-type tyrosine phosphatase ligands: looking for the needle in the haystack. *FEBS J.* **2013,** *280* (2), 388-400.

18. Sacco, F.; Perfetto, L.; Castagnoli, L.; Cesareni, G., The human phosphatase interactome: An intricate family portrait. *FEBS Lett.* **2012,** *586* (17), 2732-9.

19. Hunter, T., Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* **1995,** *80* (2), 225-36.

20. LaRochelle, J. R.; Fodor, M.; Xu, X.; Durzynska, I.; Fan, L.; Stams, T.; Chan, H. M.; LaMarche, M. J.; Chopra, R.; Wang, P.; Fortin, P. D.; Acker, M. G.; Blacklow, S. C., Structural and Functional Consequences of Three Cancer-Associated Mutations of the Oncogenic Phosphatase SHP2. *Biochemistry* **2016,** *55* (15), 2269-77.

21. Ostman, A.; Hellberg, C.; Bohmer, F. D., Protein-tyrosine phosphatases and cancer. *Nat. Rev. Cancer* **2006,** *6* (4), 307-20.

22. Yutzey, K. E.; Colbert, M.; Robbins, J., Ras-related signaling pathways in valve development: ebb and flow. *Physiology (Bethesda)* **2005,** *20*, 390-7.

23. Keilhack, H.; David, F. S.; McGregor, M.; Cantley, L. C.; Neel, B. G., Diverse biochemical properties of Shp2 mutants. Implications for disease phenotypes. *J. Biol. Chem.* **2005,** *280* (35), 30984-93.

24. Yuan, X.; Bu, H.; Zhou, J.; Yang, C. Y.; Zhang, H., Recent Advances of SHP2 Inhibitors in Cancer Therapy: Current Development and Clinical Application. *J. Med. Chem.* **2020,** *63* (20), 11368-11396.

25. Yu, Z. H.; Zhang, Z. Y., Regulatory Mechanisms and Novel Therapeutic Targeting Strategies for Protein Tyrosine Phosphatases. *Chem. Rev.* **2018,** *118* (3), 1069-1091.

26. Roskoski, R., Jr., Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacol. Res.* **2020,** *152*, 104609.

27. Rodriguez, A.; Roy, J.; Martinez-Martinez, S.; Lopez-Maderuelo, M. D.; Nino-Moreno, P.; Orti, L.; Pantoja-Uceda, D.; Pineda-Lucena, A.; Cyert, M. S.; Redondo, J. M., A conserved docking surface on calcineurin mediates interaction with substrates and immunosuppressants. *Mol. Cell* **2009,** *33* (5), 616-26.

28. Grigoriu, S.; Bond, R.; Cossio, P.; Chen, J. A.; Ly, N.; Hummer, G.; Page, R.; Cyert, M. S.; Peti, W., The molecular mechanism of substrate engagement and immunosuppressant inhibition of calcineurin. *PLoS Biol.* **2013,** *11* (2), e1001492.

29. Mullard, A., Phosphatases start shedding their stigma of undruggability. *Nat. Rev. Drug Discov.* **2018,** *17* (12), 847-849.

30. Kohn, M., Turn and Face the Strange: A New View on Phosphatases. *ACS Cent Sci* **2020,** *6* (4), 467-477.

31.Chen, Y. N.; LaMarche, M. J.; Chan, H. M.; Fekkes, P.; Garcia-Fortanet, J.; Acker, M. G.; Antonakos, B.; Chen, C. H.; Chen, Z.; Cooke, V. G.; Dobson, J. R.; Deng, Z.; Fei, F.; Firestone, B.; Fodor, M.; Fridrich, C.; Gao, H.; Grunenfelder, D.; Hao, H. X.; Jacob, J.; Ho, S.; Hsiao, K.; Kang, Z. B.; Karki, R.; Kato, M.; Larrow, J.; La Bonte, L. R.; Lenoir, F.; Liu, G.; Liu, S.; Majumdar, D.; Meyer, M. J.; Palermo, M.; Perez, L.; Pu, M.; Price, E.; Quinn, C.; Shakya, S.; Shultz, M. D.; Slisz, J.; Venkatesan, K.; Wang, P.; Warmuth, M.; Williams, S.; Yang, G.; Yuan, J.; Zhang, J. H.; Zhu, P.; Ramsey, T.; Keen, N. J.; Sellers, W. R.; Stams, T.; Fortin, P. D., Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **2016,** *535* (7610), 148-52.

32. Garcia Fortanet, J.; Chen, C. H.; Chen, Y. N.; Chen, Z.; Deng, Z.; Firestone, B.; Fekkes, P.; Fodor, M.; Fortin, P. D.; Fridrich, C.; Grunenfelder, D.; Ho, S.; Kang, Z. B.; Karki, R.; Kato, M.; Keen, N.; LaBonte, L. R.; Larrow, J.; Lenoir, F.; Liu, G.; Liu, S.; Lombardo, F.; Majumdar, D.; Meyer, M. J.; Palermo, M.; Perez, L.; Pu, M.; Ramsey, T.; Sellers, W. R.; Shultz, M. D.; Stams, T.; Towler, C.; Wang, P.; Williams, S. L.; Zhang, J. H.; LaMarche, M. J., Allosteric Inhibition of SHP2: Identification of a Potent, Selective, and Orally Efficacious Phosphatase Inhibitor. *J. Med. Chem.* **2016,** *59* (17), 7773-82.

33. Shen, D.; Chen, W.; Zhu, J.; Wu, G.; Shen, R.; Xi, M.; Sun, H., Therapeutic potential of targeting SHP2 in human developmental disorders and cancers. *Eur. J. Med. Chem.* **2020,** *190*, 112117.

34. Ulman, A., Formation and Structure of Self-Assembled Monolayers. *Chem. Rev.* **1996,** *96* (4), 1533-1554.

35. Li, J.; Thiara, P. S.; Mrksich, M., Rapid evaluation and screening of interfacial reactions on self-assembled monolayers. *Langmuir* **2007,** *23* (23), 11826-35.

36. Montavon, T. J.; Li, J.; Cabrera-Pardo, J. R.; Mrksich, M.; Kozmin, S. A., Three-component reaction discovery enabled by mass spectrometry of self-assembled monolayers. *Nat. Chem.* **2011,** *4* (1), 45-51.

37. Bayly, A. A.; McDonald, B. R.; Mrksich, M.; Scheidt, K. A., High-throughput photocapture approach for reaction discovery. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** *117* (24), 13261-13266.

38. Hodneland, C. D.; Mrksich, M., Biomolecular Surfaces that Release Ligands under Electrochemical Control. *J. Am. Chem. Soc.* **2000,** *122* (17), 4235-4236.

39. Gurard-Levin, Z. A.; Mrksich, M., Combining self-assembled monolayers and mass spectrometry for applications in biochips. *Annu. Rev. Anal. Chem. (Palo Alto Calif.)* **2008,** *1*, 767-800.

40. Kilian, K. A.; Mrksich, M., Directing stem cell fate by controlling the affinity and density of ligand-receptor interactions at the biomaterials interface. *Angew. Chem. Int. Ed. Engl.* **2012,** *51* (20), 4891-5.

41. Bugga, P.; Mrksich, M., Sequential Photoactivation of Self-Assembled Monolayers to Direct Cell Adhesion and Migration. *Langmuir* **2019,** *35* (17), 5937-5943.

42. Prime, K. L.; Whitesides, G. M., Self-assembled organic monolayers: model systems for studying adsorption of proteins at surfaces. *Science* **1991,** *252* (5009), 1164-7.

43. Devaraj, N. K.; Miller, G. P.; Ebina, W.; Kakaradov, B.; Collman, J. P.; Kool, E. T.; Chidsey, C. E. D., Chemoselective Covalent Coupling of Oligonucleotide Probes to Self-Assembled Monolayers. *J. Am. Chem. Soc.* **2005,** *127* (24), 8600-8601.

44. Houseman, B. T.; Huh, J. H.; Kron, S. J.; Mrksich, M., Peptide chips for the quantitative evaluation of protein kinase activity. *Nat. Biotechnol.* **2002,** *20* (3), 270-4.

45. Haeussling, L.; Ringsdorf, H.; Schmitt, F. J.; Knoll, W., Biotin-functionalized self-assembled monolayers on gold: surface plasmon optical studies of specific recognition reactions. *Langmuir* **1991,** *7* (9), 1837-1840.

46. Helal, K. Y.; Alamgir, A.; Berns, E. J.; Mrksich, M., Traceless Immobilization of Analytes for High-Throughput Experiments with SAMDI Mass Spectrometry. *J. Am. Chem. Soc.* **2018,** *140* (26), 8060-8063.

47. Houseman, B. T.; Gawalt, E. S.; Mrksich, M., Maleimide-Functionalized Self-Assembled Monolayers for the Preparation of Peptide and Carbohydrate Biochips. *Langmuir* **2003,** *19* (5), 1522-1531.

48. Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T., Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* **1991,** *63* (24), 1193A-1203A.

49.Karas, M.; Kruger, R., Ion formation in MALDI: the cluster ionization mechanism. *Chem. Rev.* **2003,** *103* (2), 427-40.

50. Garrett, T. J.; Prieto-Conaway, M. C.; Kovtoun, V.; Bui, H.; Izgarian, N.; Stafford, G.; Yost, R. A., Imaging of small molecules in tissue sections with a new intermediate-pressure MALDI linear ion trap mass spectrometer. *Int. J. Mass spectrom.* **2007,** *260* (2), 166-176.

51. Fujimura, Y.; Miura, D., MALDI Mass Spectrometry Imaging for Visualizing In Situ Metabolism of Endogenous Metabolites and Dietary Phytochemicals. *Metabolites* **2014,** *4* (2), 319-46.

52. Nedelkov, D.; Nelson, R. W., Surface plasmon resonance mass spectrometry: recent progress and outlooks. *Trends Biotechnol.* **2003,** *21* (7), 301-5.

53. Engwegen, J. Y.; Gast, M. C.; Schellens, J. H.; Beijnen, J. H., Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. *Trends Pharmacol. Sci.* **2006,** *27* (5), 251-9.

54. Hu, C. C.; Huang, M. F.; Chang, H. T., Quantitative surface-assisted laser desorption/ionization-MS approaches for bioanalysis. *Bioanalysis* **2013,** *5* (6), 633-5.

55. Wu, C. Y.; Lee, K. C.; Kuo, Y. L.; Chen, Y. C., Revisiting the quantitative features of surface-assisted laser desorption/ionization mass spectrometric analysis. *Philos Trans A Math Phys Eng Sci* **2016,** *374* (2079).

56. Trevor, J. L.; Lykke, K. R.; Pellin, M. J.; Hanley, L., Two-Laser Mass Spectrometry of Thiolate, Disulfide, and Sulfide Self-Assembled Monolayers. *Langmuir* **1998,** *14* (7), 1664-1673.

57. Gong, W.; Elitzin, V. I.; Janardhanam, S.; Wilkins, C. L.; Fritsch, I., Effect of Laser Fluence on Laser Desorption Mass Spectra of Organothiol Self-Assembled Monolayers on Gold. *J. Am. Chem. Soc.* **2001,** *123* (4), 769-770.

58. Mrksich, M., Mass spectrometry of self-assembled monolayers: a new tool for molecular surface science. *ACS Nano* **2008,** *2* (1), 7-18.

59. Su, J.; Mrksich, M., Using mass spectrometry to characterize self-assembled monolayers presenting peptides, proteins, and carbohydrates. *Angew. Chem. Int. Ed. Engl.* **2002,** *41* (24), 4715-8.

60. Berns, E. J.; Cabezas, M. D.; Mrksich, M., Cellular Assays with a Molecular Endpoint Measured by SAMDI Mass Spectrometry. *Small* **2016,** *12* (28), 3811-3818.

61. Guan, W.; Ban, L.; Cai, L.; Li, L.; Chen, W.; Liu, X.; Mrksich, M.; Wang, P. G., Combining carbochips and mass spectrometry to study the donor specificity for the Neisseria meningitidis beta1,3-N-acetylglucosaminyltransferase LgtA. *Bioorg. Med. Chem. Lett.* **2011,** *21* (17), 5025-8.

62. Kornacki, J. R.; Stuparu, A. D.; Mrksich, M., Acetyltransferase p300/CBP associated Factor (PCAF) regulates crosstalk-dependent acetylation of histone H3 by distal site recognition. *ACS Chem. Biol.* **2015,** *10* (1), 157-64.

63. Min, D. H.; Su, J.; Mrksich, M., Profiling kinase activities by using a peptide chip and mass spectrometry. *Angew. Chem. Int. Ed. Engl.* **2004,** *43* (44), 5973-7.

64. Pu, F.; Elsen, N. L.; Williams, J. D., Emerging Chromatography-Free High-Throughput Mass Spectrometry Technologies for Generating Hits and Leads. *ACS Med. Chem. Lett.* **2020,** *11* (11), 2108-2113.

65. Szymczak, L. C.; Kuo, H. Y.; Mrksich, M., Peptide Arrays: Development and Application. *Anal. Chem.* **2018,** *90* (1), 266-282.

66. Chan, W. C.; White, P. D., *Fmoc solid phase peptide synthesis : a practical approach*. Oxford University Press: New York, 2000; p xxiv, 346 p.

67. Frank, R., The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J. Immunol. Methods* **2002,** *267* (1), 13-26.

68. Beyer, M.; Nesterov, A.; Block, I.; Konig, K.; Felgenhauer, T.; Fernandez, S.; Leibe, K.; Torralba, G.; Hausmann, M.; Trunk, U.; Lindenstruth, V.; Bischoff, F. R.; Stadler, V.; Breitling, F., Combinatorial synthesis of peptide arrays onto a microchip. *Science* **2007,** *318* (5858), 1888.

69. Fodor, S. P.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D., Light-directed, spatially addressable parallel chemical synthesis. *Science* **1991,** *251* (4995), 767-73.

70. Min, D. H.; Mrksich, M., Peptide arrays: towards routine implementation. *Curr. Opin. Chem. Biol.* **2004,** *8* (5), 554-8.

71. Gurard-Levin, Z. A.; Kim, J.; Mrksich, M., Combining mass spectrometry and peptide arrays to profile the specificities of histone deacetylases. *ChemBioChem* **2009,** *10* (13), 2159-61.

72. Gurard-Levin, Z. A.; Kilian, K. A.; Kim, J.; Bahr, K.; Mrksich, M., Peptide arrays identify isoform-selective substrates for profiling endogenous lysine deacetylase activity. *ACS Chem. Biol.* **2010,** *5* (9), 863-73.

73. Kuo, H. Y.; DeLuca, T. A.; Miller, W. M.; Mrksich, M., Profiling deacetylase activities in cell lysates with peptide arrays and SAMDI mass spectrometry. *Anal. Chem.* **2013,** *85* (22), 10635-42.

74. Szymczak, L. C.; Mrksich, M., Using Peptide Arrays To Discover the Sequence-Specific Acetylation of the Histidine-Tyrosine Dyad. *Biochemistry* **2019,** *58* (13), 1810-1817.

75. Wood, S. E.; Sinsinbar, G.; Gudlur, S.; Nallani, M.; Huang, C. F.; Liedberg, B.; Mrksich, M., A Bottom-Up Proteomic Approach to Identify Substrate Specificity of Outer-Membrane Protease OmpT. *Angew. Chem. Int. Ed. Engl.* **2017,** *56* (52), 16531-16535.

76. Kightlinger, W.; Lin, L.; Rosztoczy, M.; Li, W.; DeLisa, M. P.; Mrksich, M.; Jewett, M. C., Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol.* **2018,** *14* (6), 627-635.

77. Lin, L.; Kightlinger, W.; Prabhu, S. K.; Hockenberry, A. J.; Li, C.; Wang, L. X.; Jewett, M. C.; Mrksich, M., Sequential Glycosylation of Proteins with Substrate-Specific N-Glycosyltransferases. *ACS Cent Sci* **2020,** *6* (2), 144-154.

78. Techner, J.-M.; Kightlinger, W.; Lin, L.; Hershewe, J.; Ramesh, A.; DeLisa, M. P.; Jewett, M. C.; Mrksich, M., High-Throughput Synthesis and Analysis of Intact Glycoproteins Using SAMDI-MS. *Anal. Chem.* **2020,** *92* (2), 1963-1971.

79. Hunter, T., Signaling--2000 and beyond. *Cell* **2000,** *100* (1), 113-27.

80. Steelman, L. S.; Franklin, R. A.; Abrams, S. L.; Chappell, W.; Kempf, C. R.; Basecke, J.; Stivala, F.; Donia, M.; Fagone, P.; Nicoletti, F.; Libra, M.; Ruvolo, P.; Ruvolo, V.; Evangelisti, C.; Martelli, A. M.; McCubrey, J. A., Roles of the Ras/Raf/MEK/ERK pathway in leukemia therapy. *Leukemia* **2011,** *25* (7), 1080-94.

81. Tartaglia, M.; Mehler, E. L.; Goldberg, R.; Zampino, G.; Brunner, H. G.; Kremer, H.; van der Burgt, I.; Crosby, A. H.; Ion, A.; Jeffery, S.; Kalidas, K.; Patton, M. A.; Kucherlapati, R. S.; Gelb, B. D., Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.* **2001,** *29* (4), 465-8.

82. Niemeyer, C. M., RAS diseases in children. *Haematologica* **2014,** *99* (11), 1653-62.

83. Huang, W.; Saberwal, G.; Horvath, E.; Zhu, C.; Lindsey, S.; Eklund, E. A., Leukemia-associated, constitutively active mutants of SHP2 protein tyrosine phosphatase inhibit NF1 transcriptional activation by the interferon consensus sequence binding protein. *Mol. Cell. Biol.* **2006,** *26* (17), 6311-32.

84. Bei, L.; Shah, C.; Wang, H.; Huang, W.; Platanias, L. C.; Eklund, E. A., Regulation of CDX4 gene transcription by HoxA9, HoxA10, the Mll-Ell oncogene and Shp2 during leukemogenesis. *Oncogenesis* **2014,** *3*, e135.

85. Zhang, J.; Zhang, F.; Niu, R., Functions of Shp2 in cancer. *J. Cell. Mol. Med.* **2015,** *19* (9), 2075-83.

86. Geladopoulos, T. P.; Sotiroudis, T. G.; Evangelopoulos, A. E., A malachite green colorimetric assay for protein phosphatase activity. *Anal. Biochem.* **1991,** *192* (1), 112-116.

87. Takai, A.; Mieskes, G., Inhibitory effect of okadaic acid on the p-nitrophenyl phosphate phosphatase activity of protein phosphatases. *Biochem. J* **1991,** *275 ( Pt 1)*, 233-9.

88. Welte, S.; Baringhaus, K. H.; Schmider, W.; Muller, G.; Petry, S.; Tennagels, N., 6,8-Difluoro-4-methylumbiliferyl phosphate: a fluorogenic substrate for protein tyrosine phosphatases. *Anal. Biochem.* **2005,** *338* (1), 32-8.

89. McAvoy, T.; Nairn, A. C., Serine/Threonine Protein Phosphatase Assays. *Current protocols in molecular biology* **2010,** *92*, 18.18.11-18.18.11.

90.Killilea, S. D.; Cheng, Q.; Wang, Z.-X., Protein Phosphatase Type 1 and Type 2A Assays. In *Protein Phosphatase Protocols*, Ludlow, J. W., Ed. Humana Press: Totowa, NJ, 1998; pp 23-33.

91. Bose, A. K.; Janes, K. A., A High-throughput Assay for Phosphoprotein-specific Phosphatase Activity in Cellular Extracts. *Molecular & Cellular Proteomics* **2013,** *12* (3), 797-806.

92. Su, J.; Mrksich, M., Using Mass Spectrometry to Characterize Self-Assembled Monolayers Presenting Peptides, Proteins, and Carbohydrates. *Angewandte Chemie International Edition* **2002,** *41* (24), 4715-4718.

93. Gurard-Levin, Z. A.; Scholle, M. D.; Eisenberg, A. H.; Mrksich, M., High-Throughput Screening of Small Molecule Libraries using SAMDI Mass Spectrometry. *ACS Combinatorial Science* **2011,** *13* (4), 347-350.

94. Yu, B.; Liu, W.; Yu, W. M.; Loh, M. L.; Alter, S.; Guvench, O.; Mackerell, A. D., Jr.; Tang, L. D.; Qu, C. K., Targeting protein tyrosine phosphatase SHP2 for the treatment of PTPN11-associated malignancies. *Mol. Cancer Ther.* **2013,** *12* (9), 1738-48.

95. Zhang, J. H.; Chung, T. D. Y.; Oldenburg, K. R., A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **1999,** *4* (2), 67-73.

96. Weiss, J. N., The Hill equation revisited: uses and misuses. *FASEB J.* **1997,** *11* (11), 835-41.

97. Poole, A. W.; Jones, M. L., A SHPing tale: perspectives on the regulation of SHP-1 and SHP-2 tyrosine phosphatases by the C-terminal tail. *Cell. Signal.* **2005,** *17* (11), 1323-32.

98. Milanese, A.; Gorincioi, E.; Rajabi, M.; Vistoli, G.; Santaniello, E., New synthesis of 6[3-(1-adamantyl)-4-methoxyphenyl]-2-naphthoic acid and evaluation of the influence of adamantyl group on the DNA binding of a naphthoic retinoid. *Bioorg. Chem.* **2011,** *39* (4), 151-8.

99. Szymczak, L. C.; Huang, C. F.; Berns, E. J.; Mrksich, M., Combining SAMDI Mass Spectrometry and Peptide Arrays to Profile Phosphatase Activities. *Methods Enzymol.* **2018,** *607*, 389-403.

100. Tartaglia, M.; Martinelli, S.; Stella, L.; Bocchinfuso, G.; Flex, E.; Cordeddu, V.; Zampino, G.; Burgt, I.; Palleschi, A.; Petrucci, T. C.; Sorcini, M.; Schoch, C.; Foa, R.; Emanuel, P. D.; Gelb, B. D., Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. *Am. J. Hum. Genet.* **2006,** *78* (2), 279-90.

101. Grossmann, K. S.; Rosário, M.; Birchmeier, C.; Birchmeier, W., The Tyrosine Phosphatase Shp2 in Development and Cancer. In *Adv. Cancer Res.*, George, F. V. W.; George, K., Eds. Academic Press: 2010; Vol. Volume 106, pp 53-89.

102. Hof, P.; Pluskey, S.; Dhe-Paganon, S.; Eck, M. J.; Shoelson, S. E., Crystal structure of the tyrosine phosphatase SHP-2. *Cell* **1998,** *92* (4), 441-50.

103. Tsai, C. F.; Wang, Y. T.; Yen, H. Y.; Tsou, C. C.; Ku, W. C.; Lin, P. Y.; Chen, H. Y.; Nesvizhskii, A. I.; Ishihama, Y.; Chen, Y. J., Large-scale determination of absolute phosphorylation stoichiometries in human cells by motif-targeting quantitative proteomics. *Nat Commun* **2015,** *6*, 6622.

104. Xue, L.; Wang, W. H.; Iliuk, A.; Hu, L.; Galan, J. A.; Yu, S.; Hans, M.; Geahlen, R. L.; Tao, W. A., Sensitive kinase assay linked with phosphoproteomics for identifying direct kinase substrates. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109* (15), 5615-20.

105. Zhu, H.; Klemic, J. F.; Chang, S.; Bertone, P.; Casamayor, A.; Klemic, K. G.; Smith, D.; Gerstein, M.; Reed, M. A.; Snyder, M., Analysis of yeast protein kinases using protein chips. *Nat. Genet.* **2000,** *26*, 283.

106. Fasolo, J.; Sboner, A.; Sun, M. G.; Yu, H.; Chen, R.; Sharon, D.; Kim, P. M.; Gerstein, M.; Snyder, M., Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. *Genes Dev.* **2011,** *25* (7), 767-78.

107. Zhang, Z. Y.; Thieme-Sefler, A. M.; Maclean, D.; McNamara, D. J.; Dobrusin, E. M.; Sawyer, T. K.; Dixon, J. E., Substrate specificity of the protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90* (10), 4446-50.

108. Zhang, Z. Y.; Dixon, J. E., Protein tyrosine phosphatases: mechanism of catalysis and substrate specificity. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1994,** *68*, 1-36.

109. Alchab, F.; Sibille, E.; Ettouati, L.; Bana, E.; Bouaziz, Z.; Mularoni, A.; Monniot, E.; Bagrel, D.; Jose, J.; Le Borgne, M.; Chaimbault, P., Screening of indeno[1,2-b]indoloquinones by MALDI-MS: a new set of potential CDC25 phosphatase inhibitors brought to light. *J. Enzyme Inhib. Med. Chem.* **2016,** *31* (sup3), 25-32.

110. Winter, M.; Bretschneider, T.; Kleiner, C.; Ries, R.; Hehn, J. P.; Redemann, N.; Luippold, A. H.; Bischoff, D.; Buttner, F. H., Establishing MALDI-TOF as Versatile Drug Discovery Readout to Dissect the PTP1B Enzymatic Reaction. *SLAS Discov* **2018,** *23* (6), 561-573.

111. Winter, M.; Ries, R.; Kleiner, C.; Bischoff, D.; Luippold, A. H.; Bretschneider, T.; Buttner, F. H., Automated MALDI Target Preparation Concept: Providing Ultra-High-Throughput Mass Spectrometry-Based Screening for Drug Discovery. *SLAS Technol* **2018**, 2472630318791981.

112. Meng, X.; Wei, J.; Wang, Y.; Zhang, H.; Wang, Z., The role of peptide microarrays in biomedical research. *Analytical Methods* **2018,** *10* (38), 4614-4624.

113. Garaud, M.; Pei, D., Substrate profiling of protein tyrosine phosphatase PTP1B by screening a combinatorial peptide library. *J. Am. Chem. Soc.* **2007,** *129* (17), 5366-7.

114. Ren, L.; Chen, X.; Luechapanichkul, R.; Selner, N. G.; Meyer, T. M.; Wavreille, A. S.; Chan, R.; Iorio, C.; Zhou, X.; Neel, B. G.; Pei, D., Substrate specificity of protein tyrosine phosphatases 1B, RPTPalpha, SHP-1, and SHP-2. *Biochemistry* **2011,** *50* (12), 2339-56.

115. Selner, N. G.; Luechapanichkul, R.; Chen, X.; Neel, B. G.; Zhang, Z. Y.; Knapp, S.; Bell, C. E.; Pei, D., Diverse levels of sequence selectivity and catalytic efficiency of protein-tyrosine phosphatases. *Biochemistry* **2014,** *53* (2), 397-412.

116. Palma, A.; Tinti, M.; Paoluzi, S.; Santonico, E.; Brandt, B. W.; Hooft van Huijsduijnen, R.; Masch, A.; Heringa, J.; Schutkowski, M.; Castagnoli, L.; Cesareni, G., Both Intrinsic Substrate Preference and Network Context Contribute to Substrate Selection of Classical Tyrosine Phosphatases. *J. Biol. Chem.* **2017,** *292* (12), 4942-4952.

117. Sun, H.; Tan, L. P.; Gao, L.; Yao, S. Q., High-throughput screening of catalytically inactive mutants of protein tyrosine phosphatases (PTPs) in a phosphopeptide microarray. *Chem. Commun. (Camb.)* **2009,** (6), 677-9.

118. Weeks, A. M.; Wells, J. A., Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **2018,** *14* (1), 50-57.

119. Kohn, M.; Gutierrez-Rodriguez, M.; Jonkheijm, P.; Wetzel, S.; Wacker, R.; Schroeder, H.; Prinz, H.; Niemeyer, C. M.; Breinbauer, R.; Szedlacsek, S. E.; Waldmann, H., A microarray strategy for mapping the substrate specificity of protein tyrosine phosphatase. *Angew. Chem. Int. Ed. Engl.* **2007,** *46* (40), 7700-3.

120. Xue, A. Y.; Szymczak, L. C.; Mrksich, M.; Bagheri, N., Machine Learning on Signal-to-Noise Ratios Improves Peptide Array Design in SAMDI Mass Spectrometry. *Anal. Chem.* **2017,** *89* (17), 9039-9047.

121. Becker, C. F.; Wacker, R.; Bouschen, W.; Seidel, R.; Kolaric, B.; Lang, P.; Schroeder, H.; Muller, O.; Niemeyer, C. M.; Spengler, B.; Goody, R. S.; Engelhard, M., Direct readout of protein-protein interactions by mass spectrometry from protein-DNA microarrays. *Angew. Chem. Int. Ed. Engl.* **2005,** *44* (46), 7635-9.

122. Gogolin, L.; Schroeder, H.; Itzen, A.; Goody, R. S.; Niemeyer, C. M.; Becker, C. F., Protein-DNA arrays as tools for detection of protein-protein interactions by mass spectrometry. *ChemBioChem* **2013,** *14* (1), 92-9.

123. Pundir, S.; Martin, M. J.; O'Donovan, C.; UniProt, C., UniProt Tools. *Curr Protoc Bioinformatics* **2016,** *53*, 1 29 1-15.

124. Zhao, B. M.; Keasey, S. L.; Tropea, J. E.; Lountos, G. T.; Dyas, B. K.; Cherry, S.; Raran-Kurussi, S.; Waugh, D. S.; Ulrich, R. G., Phosphotyrosine Substrate Sequence Motifs for Dual Specificity Phosphatases. *PLoS One* **2015,** *10* (8), e0134984.

125. Turner, A. H.; Lebhar, M. S.; Proctor, A.; Wang, Q.; Lawrence, D. S.; Allbritton, N. L., Rational Design of a Dephosphorylation-Resistant Reporter Enables Single-Cell Measurement of Tyrosine Kinase Activity. *ACS Chem. Biol.* **2016,** *11* (2), 355-62.

126. Yan, B. X.; Sun, Y. Q., Glycine residues provide flexibility for enzyme active sites. *J. Biol. Chem.* **1997,** *272* (6), 3190-4.

127. Lu, K. P.; Liou, Y. C.; Zhou, X. Z., Pinning down proline-directed phosphorylation signaling. *Trends Cell Biol.* **2002,** *12* (4), 164-72.

128. Andrews, L. D.; Zalatan, J. G.; Herschlag, D., Probing the origins of catalytic discrimination between phosphate and sulfate monoester hydrolysis: comparative analysis of alkaline phosphatase and protein tyrosine phosphatases. *Biochemistry* **2014,** *53* (43), 6811-9.

129. Neel, B. G.; Chan, G.; Dhanji, S., Chapter 98 - SH2 Domain-Containing Protein-Tyrosine Phosphatases. In *Handbook of Cell Signaling (Second Edition)*, Bradshaw, R. A.; Dennis, E. A., Eds. Academic Press: San Diego, 2010; pp 771-809.

130. Formisano, P.; Sohn, K. J.; Miele, C.; Di Finizio, B.; Petruzziello, A.; Riccardi, G.; Beguinot, L.; Beguinot, F., Mutation in a conserved motif next to the insulin receptor key autophosphorylation sites de-regulates kinase activity and impairs insulin action. *J. Biol. Chem.* **1993,** *268* (7), 5241-8.

131. Hubbard, S. R., The insulin receptor: both a prototypical and atypical receptor tyrosine kinase. *Cold Spring Harb. Perspect. Biol.* **2013,** *5* (3), a008946.

132. White, M. F., Insulin signaling in health and disease. *Science* **2003,** *302* (5651), 1710-1.

133. Salmeen, A.; Andersen, J. N.; Myers, M. P.; Tonks, N. K.; Barford, D., Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1B. *Mol. Cell* **2000,** *6* (6), 1401-12.

134. Feldhammer, M.; Uetani, N.; Miranda-Saavedra, D.; Tremblay, M. L., PTP1B: a simple enzyme for a complex world. *Crit. Rev. Biochem. Mol. Biol.* **2013,** *48* (5), 430-45.

135. Casey, G. R.; Stains, C. I., Interrogating Protein Phosphatases with Chemical Activity Probes. *Chemistry (Easton)* **2018,** *24* (31), 7810-7824.

136. Casey, G. R.; Stains, C. I., A fluorescent probe for monitoring PTP-PEST enzymatic activity. *Analyst* **2020**.

137. Huang, C. F.; Mrksich, M., Profiling Protein Tyrosine Phosphatase Specificity with Self-Assembled Monolayers for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry and Peptide Arrays. *ACS Comb Sci* **2019,** *21* (11), 760-769.

138. Szymczak, L. C.; Sykora, D. J.; Mrksich, M., Using Peptide Arrays to Profile Phosphatase Activity in Cell Lysates. *Chemistry (Easton)* **2020,** *26* (1), 165-170.

139. McEwen, A. E.; Escobar, D. E.; Gottardi, C. J., Signaling from the adherens junction. *Subcell. Biochem.* **2012,** *60*, 171-96.

140. Daugherty, R. L.; Gottardi, C. J., Phospho-regulation of Beta-catenin adhesion and signaling functions. *Physiology (Bethesda)* **2007,** *22*, 303-9.

141. Roura, S.; Miravet, S.; Piedra, J.; Garcia de Herreros, A.; Dunach, M., Regulation of E-cadherin/Catenin association by tyrosine phosphorylation. *J. Biol. Chem.* **1999,** *274* (51), 36734-40.

142. Roper, J. C.; Mitrossilis, D.; Stirnemann, G.; Waharte, F.; Brito, I.; Fernandez-Sanchez, M. E.; Baaden, M.; Salamero, J.; Farge, E., The major beta-catenin/E-cadherin junctional binding site is a primary molecular mechano-transductor of differentiation in vivo. *Elife* **2018,** *7*.

143. Yan, H. X.; Yang, W.; Zhang, R.; Chen, L.; Tang, L.; Zhai, B.; Liu, S. Q.; Cao, H. F.; Man, X. B.; Wu, H. P.; Wu, M. C.; Wang, H. Y., Protein-tyrosine phosphatase PCP-2 inhibits beta-catenin signaling and increases E-cadherin-dependent cell adhesion. *J. Biol. Chem.* **2006,** *281* (22), 15423-33.

144. Simoneau, M.; Coulombe, G.; Vandal, G.; Vezina, A.; Rivard, N., SHP-1 inhibits beta-catenin function by inducing its degradation and interfering with its association with TATA-binding protein. *Cell. Signal.* **2011,** *23* (1), 269-79.

145. van Veelen, W.; Le, N. H.; Helvensteijn, W.; Blonden, L.; Theeuwes, M.; Bakker, E. R.; Franken, P. F.; van Gurp, L.; Meijlink, F.; van der Valk, M. A.; Kuipers, E. J.; Fodde, R.;

Smits, R., beta-catenin tyrosine 654 phosphorylation increases Wnt signalling and intestinal tumorigenesis. *Gut* **2011,** *60* (9), 1204-12.

146. Tucci, V.; Kleefstra, T.; Hardy, A.; Heise, I.; Maggi, S.; Willemsen, M. H.; Hilton, H.; Esapa, C.; Simon, M.; Buenavista, M. T.; McGuffin, L. J.; Vizor, L.; Dodero, L.; Tsaftaris, S.; Romero, R.; Nillesen, W. N.; Vissers, L. E.; Kempers, M. J.; Vulto-van Silfhout, A. T.; Iqbal, Z.; Orlando, M.; Maccione, A.; Lassi, G.; Farisello, P.; Contestabile, A.; Tinarelli, F.; Nieus, T.; Raimondi, A.; Greco, B.; Cantatore, D.; Gasparini, L.; Berdondini, L.; Bifone, A.; Gozzi, A.; Wells, S.; Nolan, P. M., Dominant beta-catenin mutations cause intellectual disability with recognizable syndromic features. *J. Clin. Invest.* **2014,** *124* (4), 1468-82.

147. Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E., PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **2015,** *43* (Database issue), D512-20.

148. Dingerdissen, H. M.; Torcivia-Rodriguez, J.; Hu, Y.; Chang, T. C.; Mazumder, R.; Kahsay, R., BioMuta and BioXpress: mutation and expression knowledgebases for cancer biomarker discovery. *Nucleic Acids Res.* **2018,** *46* (D1), D1128-D1136.

149. Kan, Z.; Jaiswal, B. S.; Stinson, J.; Janakiraman, V.; Bhatt, D.; Stern, H. M.; Yue, P.; Haverty, P. M.; Bourgon, R.; Zheng, J.; Moorhead, M.; Chaudhuri, S.; Tomsho, L. P.; Peters, B. A.; Pujara, K.; Cordes, S.; Davis, D. P.; Carlton, V. E.; Yuan, W.; Li, L.; Wang, W.; Eigenbrot, C.; Kaminker, J. S.; Eberhard, D. A.; Waring, P.; Schuster, S. C.; Modrusan, Z.; Zhang, Z.; Stokoe, D.; de Sauvage, F. J.; Faham, M.; Seshagiri, S., Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **2010,** *466* (7308), 869-73.

150. Tinti, M.; Nardozza, A. P.; Ferrari, E.; Sacco, F.; Corallino, S.; Castagnoli, L.; Cesareni, G., The 4G10, pY20 and p-TYR-100 antibody specificity: profiling by peptide microarrays. *N. Biotechnol.* **2012,** *29* (5), 571-7.

151. Jho, E. H.; Zhang, T.; Domon, C.; Joo, C. K.; Freund, J. N.; Costantini, F., Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Mol. Cell. Biol.* **2002,** *22* (4), 1172-83.

152. Varone, A.; Spano, D.; Corda, D., Shp1 in Solid Cancers and Their Therapy. *Front. Oncol.* **2020,** *10*, 935.

153. Labun, K.; Montague, T. G.; Krause, M.; Torres Cleuren, Y. N.; Tjeldnes, H.; Valen, E., CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **2019,** *47* (W1), W171-W174.

154. Guan, L.; Zhu, S.; Han, Y.; Yang, C.; Liu, Y.; Qiao, L.; Li, X.; Li, H.; Lin, J., Knockout of CTNNB1 by CRISPR-Cas9 technology inhibits cell proliferation through the Wnt/beta-catenin signaling pathway. *Biotechnol. Lett* **2018,** *40* (3), 501-508.

155. Kobayashi, W.; Ozawa, M., The epithelial-mesenchymal transition induced by transcription factor LEF-1 is independent of beta-catenin. *Biochem Biophys Rep* **2018,** *15*, 13-18.

156. Moully, E. H.; Berns, E. J.; Mrksich, M., Label-Free Assay of Protein Tyrosine Phosphatase Activity in Single Cells. *Anal. Chem.* **2019,** *91* (20), 13206-13212.

157. Pu, F.; Elsen, N. L.; Williams, J. D., Emerging Chromatography-Free High-Throughput Mass Spectrometry Technologies for Generating Hits and Leads. *ACS Med. Chem. Lett.* **2020**.

158. Yuzhalin, A. E., Citrullination in Cancer. *Cancer Res.* **2019,** *79* (7), 1274-1284.

159. Harauz, G.; Ishiyama, N.; Hill, C. M.; Bates, I. R.; Libich, D. S.; Fares, C., Myelin basic protein-diverse conformational states of an intrinsically unstructured protein and its roles in myelin assembly and multiple sclerosis. *Micron* **2004,** *35* (7), 503-42.

160. Boggs, J. M.; Rangaraj, G.; Heng, Y. M.; Liu, Y.; Harauz, G., Myelin basic protein binds microtubules to a membrane surface and to actin filaments in vitro: effect of phosphorylation and deimination. *Biochim. Biophys. Acta* **2011,** *1808* (3), 761-73.

161. Kim, J. K.; Mastronardi, F. G.; Wood, D. D.; Lubman, D. M.; Zand, R.; Moscarello, M. A., Multiple sclerosis: an important role for post-translational modifications of myelin basic protein in pathogenesis. *Mol. Cell. Proteomics* **2003,** *2* (7), 453-62.

162. Boggs, J. M.; Rangaraj, G.; Gao, W.; Heng, Y. M., Effect of phosphorylation of myelin basic protein by MAPK on its interactions with actin and actin binding to a lipid membrane in vitro. *Biochemistry* **2006,** *45* (2), 391-401.

163. Grant, J.; Goudarzi, S. H.; Mrksich, M., High-Throughput Enzyme Kinetics with 3D Microfluidics and Imaging SAMDI Mass Spectrometry. *Anal. Chem.* **2018,** *90* (21), 13096-13103.

164. Grant, J.; Modica, J. A.; Roll, J.; Perkovich, P.; Mrksich, M., An Immobilized Enzyme Reactor for Spatiotemporal Control over Reaction Products. *Small* **2018**, e1800923.

165. Grant, J.; O'Kane, P. T.; Kimmel, B. R.; Mrksich, M., Using Microfluidics and Imaging SAMDI-MS To Characterize Reaction Kinetics. *ACS Cent Sci* **2019,** *5* (3), 486-493.

166. O'Kane, P. T.; Mrksich, M., An Assay Based on SAMDI Mass Spectrometry for Profiling Protein Interaction Domains. *J. Am. Chem. Soc.* **2017,** *139* (30), 10320-10327.