

**Bicyclic Peptides as Inhibitors of Protein-Protein Interactions:  
The development of conformational phosphotyrosine mimetics**

A dissertation submitted by

Justin S. Quartararo

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Chemistry

Tufts University

February, 2015

Adviser: Professor Joshua A. Kritzer



The interactions that occur between proteins in a living system dictate many vital cellular events, and the proper control of the onset and duration of these intermolecular interactions is critical for healthy cellular activity. The aberrant regulation of protein-protein interactions has been correlated with a variety of human pathologies, and the development of chemical modulators of these interactions has therefore become the focus of a variety of medicinal chemistry campaigns. However, harvesting therapeutic lead compounds capable of antagonizing protein-protein interactions from libraries of traditional small molecules has often been challenging. Traditional small molecules represent an area of chemical space that is structurally ill-equipped for disrupting such interactions, as these compounds frequently lack the sheer size necessary to achieve a high target affinity and to discriminate among the binding pockets of similar proteins. Larger species, such as whole proteins and other biologic agents, often overcome this obstacle at the expense of poor bioavailability and metabolic stability. As a result, many protein-protein interactions have been historically intractable targets.

The work discussed herein represents an effort to target two such problematic protein-protein interactions that have been implicated in a diversity of human disease states. This work is unique in its approach, as we sought to discover and characterize inhibitors of these protein-protein interactions from a region of chemical space with physicochemical properties intermediate between those of traditional small molecules and large biologics. The work discussed in this dissertation concerns the discovery, development and structural characterization of constrained peptide scaffolds that uniquely mimic phosphotyrosine so as to antagonize the interactions of the Grb2-SH2 domain and PTP1B. This

dissertation also details the cell-penetrating capabilities of these novel peptides, as well as the atomic-level details that dictate phosphotyrosine mimicry and cellular internalization. The strategy presented herein represents an application of a potentially broad paradigm for targeting phosphotyrosine-binding proteins, and efforts remain underway to continue to use this technology for the discovery of chemical modulators of a wide variety of other phosphotyrosine-mediated protein-protein interactions.



## Acknowledgements

The work that will be presented in this dissertation would not have been possible without the help and support of my advisor, Professor Joshua Kritzer. Joshua has not only helped me to hone my skills in the lab, he has taught me what it means to be an independent scientist. Thank you Joshua for helping me to overcome any problems I have faced throughout the course of my research, and thank you for your tremendous support of my work. When the opportunities to further my research by pursuing new techniques and collaborations arose, you were supportive at every turn. You have been incredibly helpful in my professional development, as a chemist, teacher and thinker, and for this I am extremely grateful.

This work is also the product of the support and assistance of many inspiring scientists outside of my lab that I have had the fortune of interacting with over the last five years. Thank you to Professors James Baleja and Andrew Bohm for all of your help in learning and understanding techniques in structural biology that were quite new to me. I cannot overstate how much I appreciate your welcoming me into your labs and the time and attention you have given me. Thank you to Dr. Celia Harrison for your tremendous help and support as I pursued X-ray crystallography; I consider myself very fortunate for the knowledge you have shared with me. Thank you to Professor Yu-Shan Lin and Dr. Hongtao Yu, without whom some very interesting structural studies discussed in this dissertation would not have been possible. Finally, I would like to thank the members of my thesis committee, Professors Clay Bennett, Elena Rybak-Akimova and Jianmin Gao. I am inspired by your work and very grateful to have each of you as a part of my committee.

This work was also made possible with the help and support of my dear friends and family. I want to thank my mother and father for their constant support of my aspirations. The successes I have been able to achieve throughout my life are a product of their love and devotion. I am grateful for all of the trips that I have taken with my father to parks, museums, and planetariums, for I am convinced that it was the totality of these excursions at an early age that helped me to cultivate an interest in science that would ultimately help shape the path of my career. I am forever thankful for the support of my mother; she worked to instill in me the values of hard work and education since as far back as I can remember, and I am eternally grateful.

To my brother and best friend, I am thankful that regardless of where life has taken us I have always had the blessing of your support. I am exceedingly proud to call you my colleague in science and exceptionally fortunate to have you in my life. Thank you for all of your help and support, and I am now truly excited to see how your limitless scientific career will blossom.

To Daniela, I am eternally grateful for your unconditional love and support. You are the love of my life and the source of my inspiration. Thank you for always pushing me to be a better student, teacher, scientist, and person. You have been with me through all of my exams and all the late nights in the lab, and I am blessed beyond words to have been able to come home to you at the end of it all.

## Table of Contents

### **Chapter 1 – Developing Cyclic and Bicyclic Peptides as Phosphotyrosine Surrogates that Antagonize the Grb2-SH2 Domain**

1.1 – An introduction to the challenge of protein-protein interactions.....	1
1.2 – Constrained peptides: the intersection of small molecules and large biologics.....	3
1.3 – The Src homology 2 (SH2) domain of growth factor receptor-bound protein 2.....	6
1.4 – The design and synthesis of cyclic peptide inhibitors of the Grb2-SH2 domain.....	7
1.5 – Monitoring Grb2-SH2 inhibition by fluorescence polarization.....	19
1.6 – Identifying cyclic peptide inhibitors of the Grb2-SH2 domain.....	28
1.7 – Identifying bicyclic peptide inhibitors of the Grb2-SH2 domain.....	31
1.8 – Assessing the selectivity of cyclic and bicyclic peptide inhibitors for the SH2 domain of Grb2.....	46
1.9 – Investigating the serum stability of cyclic and bicyclic peptides.....	53
1.10– Peptide pre-organization enables recognition by anti-phosphotyrosine antibodies	55

### **Chapter 2 – Investigating the Cell-Penetrating Capabilities of Cyclic and Bicyclic Phosphotyrosine Mimetics**

2.1 – General rules for the cellular uptake of cationic and amphipathic peptides.....	58
2.2 – Investigating the cellular phenotypes of cyclic and bicyclic Grb2-SH2 domain inhibitors.....	63
2.3 – Characterizing the cellular uptake of peptide analogs of G1HT1 and BC1.....	64

**Chapter 3 – Analyzing the Structural Features of Cyclic and Bicyclic  
Phosphotyrosine Mimetics that Promote Target Affinity and Cell  
Penetration by X-ray Crystallography and Nuclear Magnetic Resonance  
Spectroscopy**

3.1 – Structural investigations of previous phosphotyrosine isosteres that target the Grb2-SH2 domain.....	81
3.2 – Crystallization of the Grb2-SH2 domain in complex with cyclic and bicyclic phosphotyrosine mimics.....	86
3.3 – Probing the structures of BC1 and G1HT1 by multidimensional NMR.....	110
3.4 – The interaction of BC1 with aqueous solvent: A molecular dynamics simulation..	130

**Chapter 4 – Expanding Conformational Phosphotyrosine Mimicry to Target  
Protein Tyrosine Phosphatase 1B**

4.1 – Protein tyrosine phosphatase 1B as a drug target for obesity and type-II diabetes.....	136
4.2 – Engineering cyclic peptide PTP1B inhibitors based on G1HT1.....	150
4.3 – Applying cysteine bis-alkylation to the development of bicyclic peptide inhibitors of PTP1B.....	157
4.4 – Future directions: Structural investigations and cellular activities of cyclic and bicyclic PTP1B inhibitors.....	162

**Appendices**

Appendix 1: X-ray diffraction data collection parameters.....	166
---	-----

<b>References.....</b>	<b>167</b>
------------------------	------------

## **Chapter One – Constrained peptides as phosphotyrosine surrogates to antagonize the Grb2-SH2 domain**

### **1.1 – An introduction to the challenge of protein-protein interactions**

Protein-protein interactions represent a challenging class of targets for combating human pathology.<sup>1-4</sup> These interactions are vital for the proper functioning of signal transduction processes that regulate healthy cell physiology. One very common motif employed by the nature in signal cascades is the phosphorylation of specific tyrosine residues. Tyrosine phosphorylation is a metabolically reversible process, as dedicated kinases catalyze phosphoryl transfer to tyrosine residues in key signaling proteins, while phosphatases execute the hydrolysis of these moieties. Meanwhile, tyrosine phosphorylation often serves to effect conformational changes in a protein that effect its activation or deactivation, and adaptor proteins provide a means of relaying the phosphotyrosine signal between upstream and downstream proteins in a signaling pathway. Healthy cellular physiology is tantamount to the delicate homeostasis of these processes. The aberrant activity of signal transduction pathways due to the overexpression or misregulation of key participants in these cascades has been correlated to a wide array of human diseases.<sup>5-9</sup>

Targeting protein-protein interactions has proven challenging by way of conventional medicinal chemistry approaches. The development of small molecules that bind Grb2 and PTP1B with high affinity and selectivity has slowed due to the large surface areas over which molecular recognition occurs. Such targets often lack the ability to bind ligands of high size-efficiency, a property exploited by conventional small molecule therapeutics. For this reason, such

targets have gradually come to be labeled as “undruggable.” However, today, the perception of the druggable portion of the human proteome is changing.<sup>1</sup> Large protein-protein interactions are being targeted successfully with higher molecular weight ligands, often macrocyclic and peptidic in nature.<sup>10-14</sup> Peptides are a promising class of lead compounds to target protein-protein interactions.<sup>15</sup> Their high diversity and larger size qualifies these compounds as useful molecular recognition agents over extended contact surfaces. However, unmodified peptides suffer from poor target affinity and selectivity due to conformational flexibility. Additionally, these compounds are prone to proteolytic degradation and have high de-solvation energies, rendering them generally unable to penetrate the cell membrane.<sup>16,17</sup> Peptides that have been modified so as to achieve structural pre-organization have frequently overcome these challenges.<sup>18,19</sup> Such conformational restriction can be installed via intramolecular cross-links, a strategy that has been applied extensively to rigidify alpha helices. So-called “stapled” helices exhibit improved helicity, target affinity, selectivity and cell-penetrating characteristics.<sup>10-13,20-22</sup> We seek to apply this conformational constraint to a head-to-tail cyclic peptide. Alpha helices and small cyclic peptides share several structural characteristics, including the ability of side chains to form well-defined binding epitopes and conformational flexibility in aqueous solution. We therefore hypothesized that intramolecular cross-linking can effect structural pre-organization in head-to-tail cyclic peptides, rendering these compounds more potent, selective, protease resistant and cell-penetrant. Examples of highly potent bicyclic peptides are abundant in nature, yet these compounds are largely underexplored in medicinal chemistry.<sup>23-31</sup>

## 1.2 – Constrained peptides: The intersection of small molecules and large biologics

Achieving a compromise between the promising features of large biologic agents and small molecule therapeutics has posed a long-standing challenge to the medicinal chemistry community, and particularly to the development of modulators of protein-protein interactions, as both classes of compounds have distinct advantages and disadvantages.<sup>3,4,32</sup> Peptides and proteins, for instance, offer the promise of high target affinity and selectivity by virtue of the ability to form intermolecular contacts over a wide surface area, but often suffer from poor bioavailability and metabolic stability. Small molecule lead compounds offer the promise of high oral bioavailability, but can often be marked by poor affinity, low selectivity and toxicity due to off-target effects.

Conformationally-constrained bicyclic peptides, or cyclic peptides containing an intramolecular cross-link, represent a highly underexplored, unique region of chemical space that is well-suited to tackle the challenges of difficult-to-target protein surfaces. Many examples of these types of constrained peptides exist in nature, yet these species have received little attention in the development of novel therapeutics.<sup>25,29–31</sup> The central premise that underlies our hypothesis in the projects discussed herein is as follows: that modifying peptides in such a way so as to structurally pre-organize a defined binding epitope can increase binding affinity and specificity for a specific target. In addition, we theorized that conformational restriction by intramolecular cross-linking can lead to enhanced cell penetration and resistance to proteolysis. This hypothesis is predicated on the success that stapled alpha-helices have enjoyed.<sup>10–13,20,22,33</sup> It has been shown that, by installing intramolecular cross-links that structurally rigidify the

helix, so-called “stapled” peptides have been observed to exhibit enhanced helicity, improved cell penetration and are less prone to degradation by proteases. Short alpha helical peptides and cyclic peptides share many similar features, including amino acid side chains that form well-defined binding epitopes, and both classes of peptide suffer from conformational flexibility in aqueous solution. While prominent examples of stapled alpha helical peptides have been reported, installing rationally-designed intramolecular cross-links within a head-to-tail cyclic scaffold is only recently emerging as a platform for peptide drug discovery.<sup>14,34–36</sup> We believe that, by stabilizing the bioactive conformation of these molecules, we can simultaneously achieve improved target affinity, selectivity, cell-penetrating capabilities and resistance to proteolysis. It was proposed that bicyclic peptides offer a promising new therapeutic modality by targeting protein surfaces that have proven difficult to inhibit by conventional means, and we sought to apply this technology to the inhibition of growth factor receptor-bound protein 2 (Grb2).

Bicyclic peptides are well-represented in a sampling of potent inhibitors of protein-protein interactions in nature. Sunflower trypsin inhibitor 1 (SFTI1) is an established example of a robust antagonist of a protease-substrate interaction.<sup>37–41</sup> SFTI1 is a 14-residue head-to-tail cyclic peptide containing two anti-parallel strands that are joined by a disulfide bond, rendering the overall structure bicyclic. This peptide exhibits high inhibitory potency ( $K_i = 0.1$  nM) and a rigidified structure in solution. Another example of this structural paradigm in nature can be found in  $\alpha$ -amanitin, an 8-residue bicyclic peptide containing an unusual cross-link between 6-hydroxytryptophan and cysteine residues.<sup>29</sup> This potent RNA polymerase II and III inhibitor is endogenous to several species of *Amanita* mushrooms, including the death cap mushroom and the destroying angel.

Despite the presence of an unusual cross-link and uncommon amino acids (e.g. hydroxyproline and isoleucine-(4*R*,5)-diol, this peptide is ribosomally synthesized and has an oral LD<sub>50</sub> of 0.1 mg/kg.<sup>23</sup> Despite the presence of this structural element in native biology, synthetic efforts to design and access similar structures for the inhibition of protein-protein interactions have focused chiefly on the discovery of monocyclic peptides. These include screening paradigms for head-to-tail peptide macrocycles<sup>42,43</sup> and the structure-based design of stapled helical peptides.<sup>13</sup> While recent efforts have demonstrated the synthetic accessibility of bicyclic, doubly-stapled helices and platforms have been developed for the high-throughput screening of bicyclic miniature proteins,<sup>34–36</sup> to date, there remain strikingly few examples of rationally designed bicyclic peptide inhibitors of protein-protein interactions.

The point of entry into this project is an 11-residue cyclic peptide that transiently mimics phosphotyrosine in a conformation-dependant manner. This dissertation will present extensive evidence that demonstrates the utility of this strategy, as we have used the careful application of intramolecular cross-links in order to develop this compound into a bicyclic peptide with a structurally pre-organized phosphotyrosine-like epitope. This compound exhibits improved target affinity, selectivity, protease resistivity and cell-penetrating properties.<sup>44</sup> We sought to use structure-based design to further develop inhibitors of the Grb2-SH2 domain and probe their phenotypes in live cancer cells. We also aimed to expand this strategy to the generation of peptide bicycles that target PTP1B so as to further demonstrate the utility of this platform for turning peptides into promising therapeutic leads.

### **1.3 – The Src homology 2 domain of growth factor receptor-bound protein 2**

Growth factor receptor-bound protein 2 (Grb2) plays a critical role in the activation of the Ras-MAPK signal cascade, providing the physical linkage between Her2/EGFR phosphorylation and downstream effector proteins.<sup>45,46</sup> Ras-MAPK signaling results in cell growth and proliferation, and the aberrant regulation of this pathway is a hallmark of certain cancers, including breast cancer.<sup>5,9</sup> As such, Grb2 is a validated breast cancer target, and has been the subject of a variety of studies that seek the inhibition of this adaptor protein.<sup>9,47-54</sup> Grb2 is comprised of one N-terminal SH2 domain that recognizes and binds phosphorylated tyrosine residues (pTyr 1068) on the intracellular tails of the EGFR, as well as two C-terminal SH3 domains that bind proline-rich motifs in downstream proteins.<sup>46</sup> Inhibiting the SH2 domain of Grb2 is an established paradigm for Ras-MAPK signal inhibition *in vitro*, but the molecular-level details of this interaction have rendered difficult the development of anti-Grb2 therapeutics.<sup>9,51</sup> First, the Grb2-EGFR interaction is not restricted to the phosphotyrosine binding pocket of Grb2, but rather occurs over an extended interface of approximately 1500 square angstroms, rendering this interaction difficult to antagonize with a small molecule. Furthermore, there are over 100 known human proteins that contain a distinct SH2 domain, 10 of which are known to contain multiple SH2 domains.<sup>55</sup> SH2 domains represent a class of highly conserved sequences that bind phosphotyrosine, and selectively targeting one among the host of others presents a challenge that medicinal chemistry is still working to solve. These problems are compounded by the intracellular location of the Grb2-EGFR interaction. The SH2 domain of Grb2 is highly cationic (pI > 9) and binds negatively-charged phosphotyrosine, rendering the development of a membrane-permeable phosphotyrosine surrogate a difficult task.

To solve these problems, we sought to steer away from conventional approaches that employ small molecule phosphotyrosine isosteres for Grb2 inhibition (table 1.1), and instead identify constrained peptide scaffolds capable of disrupting the Grb2-EGFR interaction by presenting a conformation-dependent phosphotyrosine mimic. The sections that follow will detail our approach to targeting the SH2 domain of Grb2.

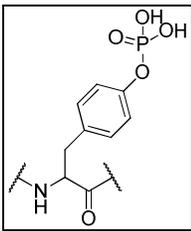
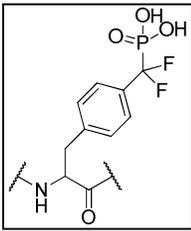
#### **1.4 – The design and synthesis of cyclic peptide inhibitors of the Grb2-SH2 domain**

We sought to discover constrained bicyclic peptides capable of antagonizing the SH2 domain of Grb2, and in doing so gain insight into a potentially general strategy of modulating cyclic peptide conformation by the installation of intramolecular covalent cross-links. Many previous efforts to discover antagonists of the Grb2-SH2 domain have focused on the development of small molecules containing chemical functionality that is isosteric or isoelectronic to phosphotyrosine (table 1.1).<sup>52–54</sup> The common chemical functionality that these compounds share is a polar anionic, aromatic acid. Examples of families of molecules that share these traits include non-hydrolyzable phosphotyrosine isosteres, such as phosphonomethyl phenylalanine and the increasingly potent difluorophosphonomethyl phenylalanine, which features an electron-withdrawing CF<sub>2</sub> moiety in place of the aryl oxygen of phosphotyrosine in order to simulate the electrostatic surface potential of phosphotyrosine while also being resistant to hydrolysis under physiological conditions. Other examples include structures featuring the highly anionic malonic acid-substituted phenylalanine, tyrosine and cyclohexene residues. With pK<sub>1</sub> on the order of 2.5 and pK<sub>2</sub> approximately 5.6, small molecules that incorporate these pharmacophores are likely to be multiply

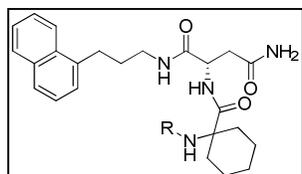
negatively charged at physiological pH, which may explain the notoriously low cellular uptake that has been observed for such molecules.<sup>52,56</sup> These compounds all employ a similar pharmacophore that likely engages the SH2 domain at the partially negatively charged oxygen atoms of the malonyl or phosphonomethyl moieties by establishing hydrogen bonds with hydroxyl protons of Ser88 and Ser90 and/or by forming electrostatic contacts with cationic Arg67 and Arg86 residues (figure 1.2).<sup>57</sup> The strong ionic character of these structures coupled with their relatively small molecular volume may explain the inability of many of these compounds to inhibit the Grb2-SH2 domain with high selectivity and the low cellular uptake associate with these molecules.

We began this work by searching for an ideal starting point molecule that would provide a forum for installing conformational constraints into a peptide that weakly inhibits the Grb2-SH2 domain. We therefore used as a point of entry

**Table 1. Representative Grb2-SH2 Domain Inhibitors**

Entry	Inhibitor	IC <sub>50</sub>	K <sub>d</sub>	IC <sub>50</sub> Assay	K <sub>d</sub> Assay	Ref
1	Ac-NXVNIE-NH <sub>2</sub> X = 	4.7 μM		ELISA		49
2		0.9 μM		ELISA		49

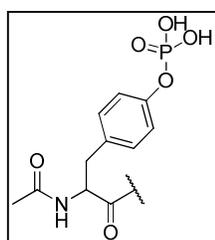
X =



ELISA

3

R =



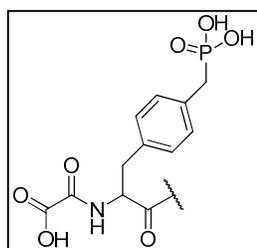
6.4 nM

ELISA

45

4

R =



10 nM

128 nM

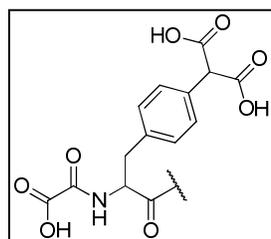
ELISA

Surface Plasmon Resonance (SPR)

45

5

R =

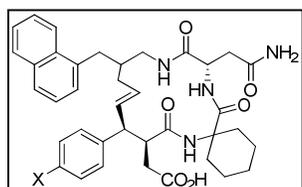


22 nM

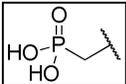
ELISA

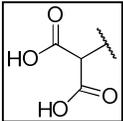
SPR

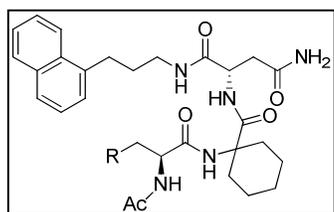
46



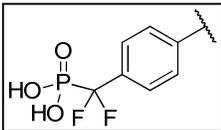
45

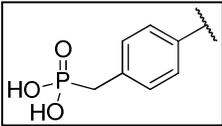
6 X =  1.4 nM 0.9 nM ELISA SPR 46

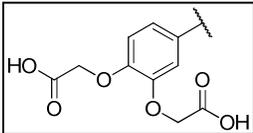
7 X =  4.3 nM 12 nM ELISA SPR 47

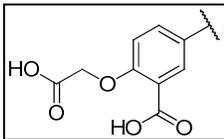


45

8 R =  80 nM SPR 45

9 R =  85 nM SPR 45

10 R =  > 100 μM SPR 49

11 R =  6.7 μM SPR 49

peptide **G1**, an 11-residue peptide rendered monocyclic by virtue of an intramolecular disulfide bridge.<sup>47</sup> This peptide was discovered by phage display as part of an effort to identify peptide modulators of the Grb2-SH2 domain. **G1** represents a novel inhibitory modality among Grb2 antagonists, as this peptide is comprised exclusively of naturally-occurring amino acids. There is no discrete phosphorylated tyrosine residue, and the molecule lacks a moiety that is explicitly isosteric or isoelectronic with phosphotyrosine. Nonetheless, the IC<sub>50</sub> of **G1** has been measured to be 20 μM in surface plasmon resonance assays.<sup>47</sup> It has been proposed that this compound, as well as the thioether derivative **G1TE** (figure 1.1) binds the SH2 domain of Grb2 in a unique fashion. Previous structure-activity relationships suggest that the Glu1, Tyr3 and Asn5 residues of this cyclic peptide are essential for preserving binding activity, while substitutions elsewhere in the molecule are tolerated to varying extents. Still other SAR investigations have revealed that the dihedral angles within Glu1-Leu2-Tyr3 motif, as well as the affinity of the molecule for the Grb2-SH2 domain, are sensitive to the

substitution pattern at the AA2 position. Additionally, it has been shown by solution-phase NMR experiments that **G1** accesses a multitude of conformational ensembles at room temperature.<sup>58</sup> Taken together with the results of *in silico* molecular docking experiments, this data provides evidence for the existence of a distinct conformation of **G1** in which the Glu1 and Tyr3 residues are oriented with respect to one another such that they comprise a discontinuous phosphotyrosine mimetic. We therefore posit that **G1** has access to a unique conformation that presents a phosphotyrosine-like epitope to the surface of Grb2, and that this conformation can be modulated by introducing structural changes elsewhere in the molecule. These structural qualities make **G1** a well-suited candidate for applying conformational modulation by intramolecular cross-linking,

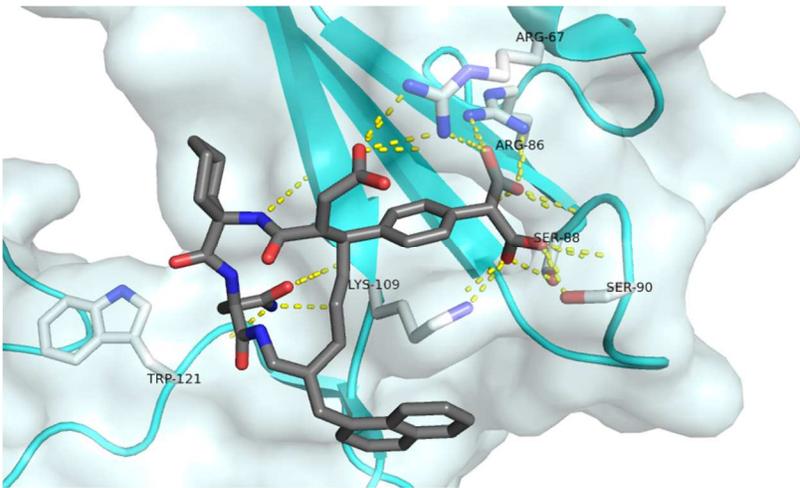
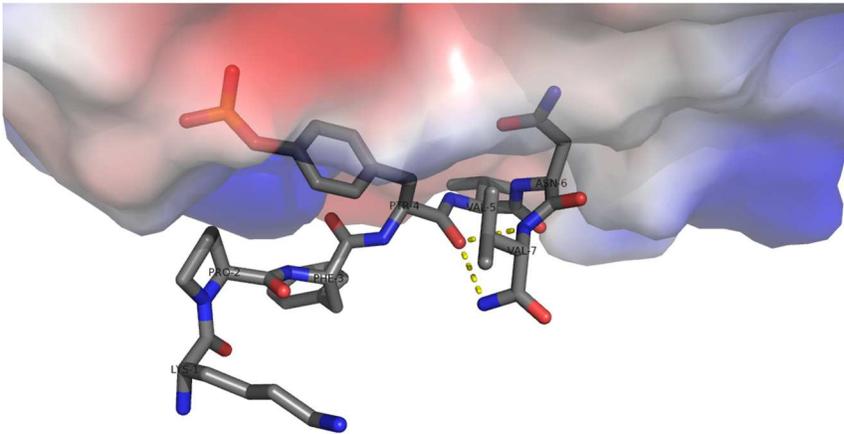
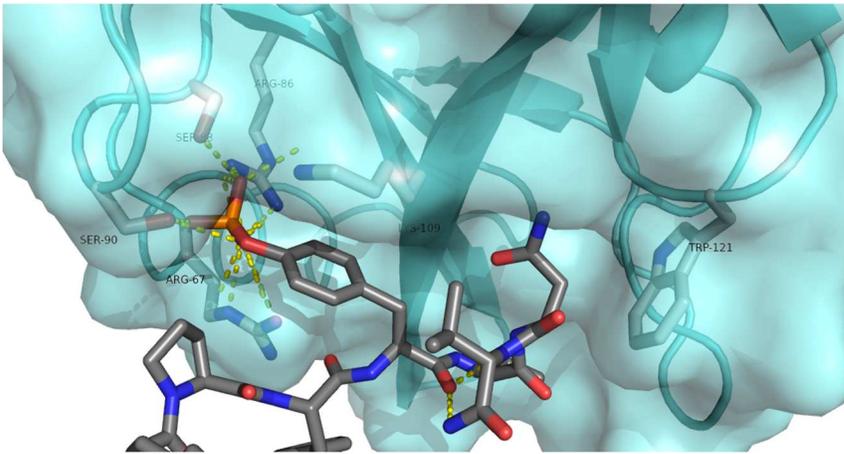


Figure 1.1. *Top, middle*: X-ray crystal structures of the Grb2-SH2 domain in complex with pYVNV peptide ligand. PDB ID: 1TZE *Bottom*: X-ray crystal structures of the Grb2-SH2 domain in complex with macrocyclic peptide inhibitor developed by Burke and colleagues. PDB ID: 1A0A

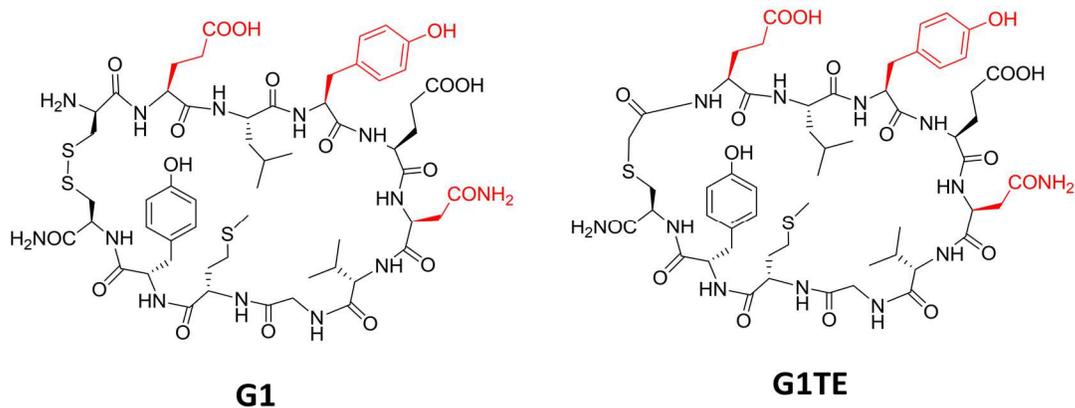


Figure 1.2. Structures of **G1** and **G1TE** peptides.

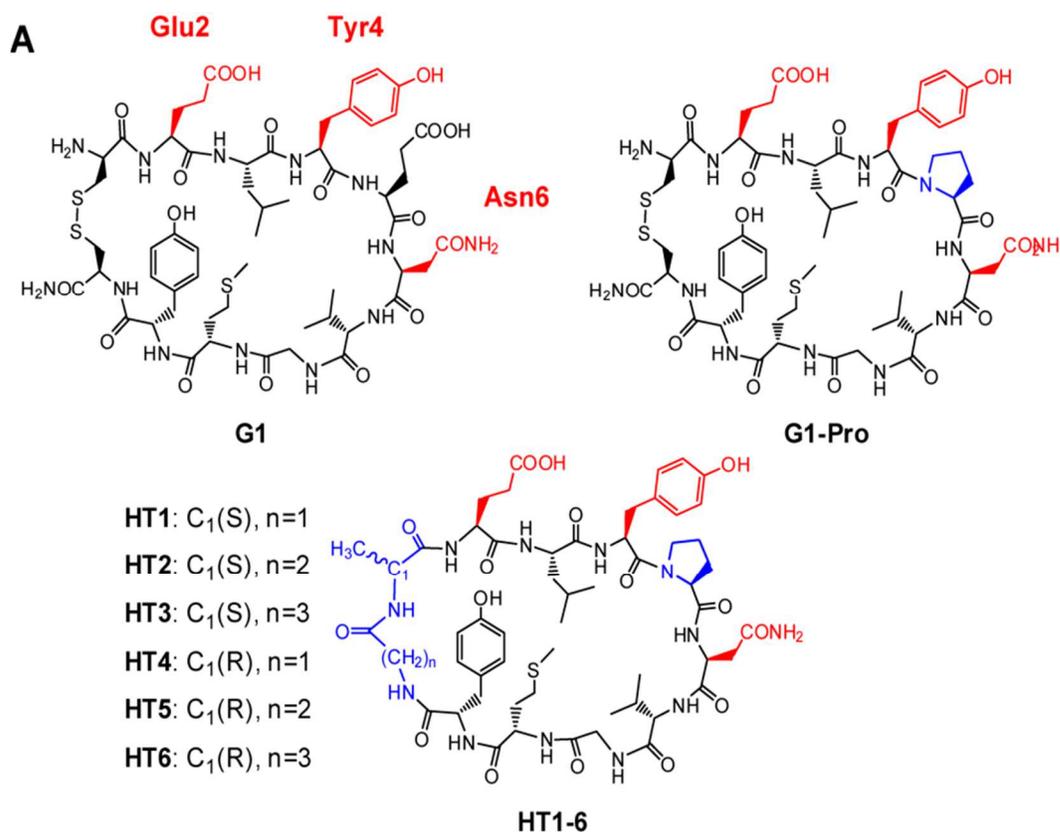


Figure 1.3. Design of the **G1HT** series. **G1**-inspired head-to-tail peptide macrocycles were synthesized so as to vary the overall macrocycle length and stereochemical configuration of the head-to-tail linker.

as we seek to selectively stabilize the unique bioactive conformation of this compound, and in turn structurally pre-organize the phosphotyrosine-like surface.

Prior to the installation of intramolecular cross-links, we first sought to replace the disulfide bridge with an explicit amide bond so as to furnish a head-to-tail cyclic peptide. This substitution was motivated by the hypothesis that, in its current state, **G1** would suffer from poor serum stability and proteolytic degradation due to the presence of the sulfur-sulfur bond. The electrophilicity and redox activity of disulfide bridges has been well documented, and this structural feature of **G1** was perceived as detrimental to the development of a serum-stable Grb2 antagonist capable of being internalized by mammalian cells. In order to replace this bond with an amide bond and retain the original structure of **G1**, we investigated a series of dipeptide linkers that scan a variety of linker lengths and stereochemical configurations. It was essential to explicitly install and test a variety of linkers because of the conformational differences between a disulfide bond and an amide bond: the diversity of bond lengths, atomic hybridizations and dihedral geometries rendered non-obvious to us which linker would maximally preserve the bioactivity of **G1**. We therefore designed a series of 6 head-to-tail cyclic peptides, **G1HT1-6**, each of which incorporates a linker with a unique length and stereochemical orientation (figure 1.3). Specifically, we varied the identity of amino acids AA10 and AA11 in the peptide sequence. To introduce different stereocenters into the linker, we introduced either Ala or d-Ala at position AA11, and to vary the linker length, we inserted either Gly,  $\beta$ -Ala, or  $\gamma$ -aminobutyric acid at position AA10. In addition, we introduced yet another structural change in this series of peptides at position AA4. In the parent peptide **G1**, the amino acid at this position is glutamic acid. Previous SARs demonstrate

that the acid functionality at this position is not necessary for SH2 domain binding.<sup>47,50,51</sup> Moreover, it has also been shown that replacement of the straight chain carboxylic acid at this position with a conformationally rigid cyclic side chain, such as cyclohexylalanine, improves the target affinity.<sup>50</sup> Molecular modeling of **G1** suggests that the bioactive conformation of the peptide includes a turn nucleated by the amino acid in position AA4. Therefore, we made the substitution of glutamic acid to proline at the onset of this investigation in order to facilitate the formation of this turn. Proline was expected to favorably nucleate a turn at this position as a result of its structurally rigidified cyclic side chain. We hypothesized this would have the effect of engendering a more favorable entropy change for the folding of the peptide into its bioactive structure, which would manifest itself in a more favorable free energy of binding and hence a more potent inhibitor. In addition to structurally pre-organizing the peptide scaffold, we further proposed that this substitution would improve the cell penetrating capacity of this iteration of peptides, as the removal of a negative charge from the system would render the peptide increasingly hydrophobic and enhance the lipid-binding capabilities of these compounds.

## **Materials and Methods**

### *Synthesis of disulfide-cyclized **G1***

Linear **G1** was synthesized by solid phase peptide synthesis. Starting with a Rink amide resin, the peptide chain was extended with repeated cycles of Fmoc-deprotection (20% piperidine in DMF) and coupling (5 eqv. amino acid, 5 eqv. PyBOP, 5 eqv. HOBt, and 13 eqv. DIPEA). Following global deprotection and cleavage with trifluoroacetic acid:water:ethanedithiol:triisopropylsilane (94:2.5:2.5:1), peptides were purified by RP-HPLC using water/acetonitrile gradients with 0.1% trifluoroacetic acid. Pure ( $\geq 95\%$ ) peptides were then dissolved to roughly 500  $\mu\text{M}$  in phosphate buffer (pH = 8.7), and these solutions were gently stirred open to the atmosphere overnight (12 hr) to oxidize the intramolecular disulfide bridge. To quench the reaction, the pH of the mixture was brought to below 3.0. The mixtures were then rotary-evaporated and lyophilized. The crude was then re-dissolved in DMSO and purified by RP-HPLC using water/acetonitrile gradients with 0.1% trifluoroacetic acid.

#### *Head-to-tail cyclic peptide synthesis*

All solid-phase resins, Fmoc-amino acids, and coupling reagents were purchased from EMD Biosciences unless otherwise noted. Head-to-tail cyclic peptides were synthesized using established solid-phase peptide synthesis procedures. Resins were swelled in dichloromethane 2 x 10 min at the start of each synthesis.

One round of peptide synthesis proceeded as follows: deprotection of N-terminal Fmoc in 20% piperidine/DMF (2 x 8 min); washing of resin in DMF (3 x 3 min); coupling of proceeding Fmoc-amino acid (5 eq. Fmoc-AA, 5 eq. PyBOP, 5 eq. HOBt, 13 eq. DIEA, 45 min); washing of resin in DMF (3 x 3 min). Following assembly of the linear peptide chain on solid support, removal of C-terminal or side-chain allyl protecting groups was achieved by the addition of 3 eq.

$\text{Pd}^0(\text{PPh}_3)_4$  in 37:2:1 chloroform:acetic acid:N-methylmorpholine to the peptidyl-

resin. The mixture was gently agitated under an Ar atmosphere for 2 hr. The resin was then washed in sequence with 0.5% sodium diethyldithiocarbamate trihydrate in DMF (2 x 5 min); 0.5% DIEA in DMF (2 x 5 min); and DMF (3 x 5 min). The resin-bound peptide was then cyclized by treatment with 5 equivalents of PyBOP, 5 equivalents of HOBt and 13 equivalents of DIEA in DMF for 24-48 hours. Coupling reagents were replenished every 12 hours as needed. Extent of cyclization was assessed by ninhydrin test or by acidic cleavage of 50-100 beads of resin and analysis of products by matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry.

Cleavage and global deprotection was carried out by first drying the solid resin after successive rinses in dichloromethane and methanol. The dried peptidyl-resin was then treated with a solution of 94:2.5:2.5:1 trifluoroacetic acid (TFA):water:ethanedithiol (EDT):triisopropylsilane (TIS) for 2 hr at room temperature under gentle agitation. The cleavage cocktail was then rotary-evaporated to approximately 2 mL and was subsequently triturated in ice-cold ether. The suspension was then centrifuged, and the crude pellet was dissolved in DMSO.

#### *Peptide purification*

Disulfide-cyclized **G1** and all head-to-tail cyclic peptides were purified by reverse phase HPLC using a preparative scale C8 column (Vydac, 0.46  $\mu$ m particle size) and a water/acetonitrile/TFA solvent system. Peptides were purified using a 5-55% acetonitrile gradient over 30 minutes. Final purity of all cyclic peptide products was at least 95% as determined by analytical re-injections of purified products.

## 1.5 – Monitoring Grb2-SH2 inhibition by fluorescence polarization

In order to screen the G1HT series of peptides for inhibitory activity against the Grb2-SH2 domain, we optimized a fluorescence polarization-based assay. This paradigm was informed by the well-documented history of other investigations that have successfully implemented fluorescence polarization techniques to measure the affinity of peptides and small molecules alike for various SH2 domains. Typically, this experiment can be run in two different ways. A fluorescence polarization direct binding experiment detects the formation of a complex between a large molecule and a smaller fluorescent species by comparing the intensity of plane-polarized fluorescent radiation that remains parallel to the excitation plane to the intensity of fluorescence in the plane perpendicular to the excitation plane. This type of experiment enables the calculation of the equilibrium constant,  $K_d$ , for the dissociation of this binary complex. This experiment can also be designed in a competitive binding format in order to monitor the disruption of this complex by an exogenous inhibitor. Therefore, a convenient means of screening a series of potential inhibitor compounds in this format can be achieved by first optimizing a direct binding FP experiment between the SH2 domain and a known, fluorescently-labeled molecule that binds the SH2 domain with a high affinity. The dose-dependent effect of an inhibitor on the formation of this complex can subsequently be monitored in a separate experiment.

Fluorescence polarization represents an ideal method for measuring the affinity of the peptides proposed herein for the Grb2-SH2 domain for several reasons. The analytical readout of a fluorescence polarization experiment is fluorescence intensity that is confined to the planes parallel and perpendicular to that of the

excitation radiation. The difference between these values relative to the total fluorescence intensity emitted is formally the fluorescence polarization of the system:

$$FP = [F(\text{Parallel}) - F(\text{Perpendicular})]/[F(\text{Parallel}) + F(\text{Perpendicular})]$$

The fraction of fluorescent radiation that remains confined to the parallel plane, and hence the magnitude of the FP measurement, is a manifestation of the physical phenomenon of a decrease in the rotational diffusion coefficient of a fluorophore upon binding a larger molecule – often canonically referred to as a “receptor.” This hindrance of rotational diffusion is intensified as the difference in molecular volume between the fluorophore and receptor increases. Previous investigations in screening phosphorylated peptides for consensus sequences in SH2 domain ligands provide a logical starting point for the design of a fluorescent tracer that can be used in direct and competitive binding experiments. The pTyr-Val-Asn-Val sequence has been identified and validated as a high-nanomolar ligand of the Grb2-SH2 domain. In order to design a fluorescent probe based on this sequence, we decided to introduce two continuous  $\beta$ -Ala residues immediately N-terminal to the phosphotyrosine. We proposed that this dipeptide would serve as a semi-flexible linker between the phosphotyrosine pharmacophore and a fluorescent dye that could be appended at the N-terminus of the molecule. The molar masses of this proposed peptide, 1024 g/mol, is approximately ten-fold less than the molar mass of the Grb2-SH2 domain. Moreover, the Grb2-SH2 domain used in these experiments was prepared as a maltose-binding protein (MBP) fusion in order to facilitate rapid affinity chromatography (see Materials and Methods). It has been shown that SH2 domains are modular protein sub-units that retain functionality even as portions

of chimeric fusion constructs, and so for the purposes of furnishing a more massive protein, we proposed leaving the MBP affinity tag on the final protein product. This would engender an even larger molecular weight differential between the SH2 domain and the fluorescent tracer, which would in turn manifest itself in a larger fluorescence polarization signal and signal-to-noise ratio. Moreover, our hypothesis that the presence of the MBP tag would not attenuate SH2-domain binding can be explicitly tested, as previous work has characterized the thermodynamic parameters of the Grb2-SH2:pYVNV complex formation. Analysis of S/N in our system and comparison of our calculated  $K_d$  to published values would collectively provide unambiguous indicators of the utility of this experimental design.

To begin the process of assay optimization, we synthesized and purified the aforementioned fluorescently-labeled peptide tracer,  $\beta$ Ala- $\beta$ Ala-pTyr-Val-Asn-Val, and subsequently functionalized the N-terminus of this peptide by conjugation with fluorescein isothiocyanate (FITC). The  $K_d$  of the Grb2-SH2:pYVNV interaction has been previously reported as 291 nM.<sup>59</sup> We therefore titrated this probe with a C-terminal MBP conjugate of the Grb2-SH2 domain at concentrations between approximately 100-fold below and 100-fold above this value. Additionally, the concentration of fluorescently-labeled probe was varied in pilot experiments from 0.5 nM up to 50 nM, one order of magnitude below the expected  $K_d$ . A variety of buffer compositions were screened for optimal S/N, including Tris-buffered saline (TBS) and phosphate-buffered saline (PBS) in the presence and absence of the non-ionic detergent (0.1%) Tween 20 or bovine-serum albumin (BSA). These latter agents were included in assay optimization screens in order to prevent non-specific binding. Ultimately, the conditions that

yielded the lowest observed error within each experiment and from trial to trial were 1x PBS (137 mM NaCl, 10 mM KHPO<sub>4</sub> and 0.18 mM K<sub>2</sub>HPO<sub>4</sub>, pH = 7.4). No significant difference in binding activity was observed at each of the different probe concentrations. We settled on a final probe concentration of 5 nM, in line with previous SH2 domain inhibitor screens. This condition was also optimal for the sake of comparison, as previous investigations have determined the K<sub>d</sub> of the Grb2-SH2:pYVNV interaction in 1x PBS as well.

After performing the direct binding experiment as described above, the K<sub>d</sub> of the interaction between the MBP-conjugated Grb2-SH2 domain with fluorescein-labeled pYVNV was calculated as 500 nM in 1x PBS (figure 1.5). This differs

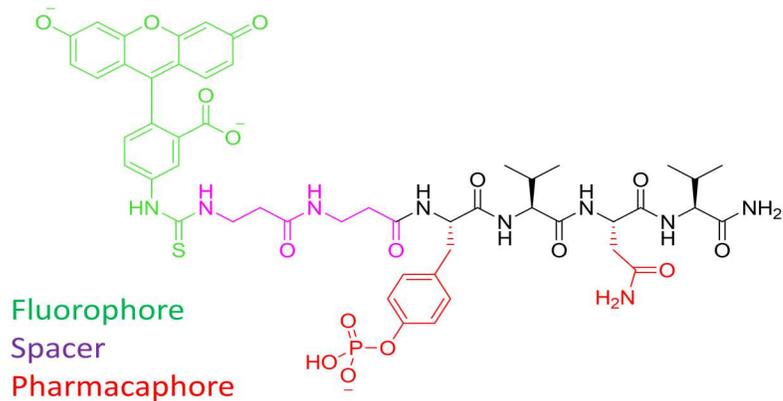


Figure 1.4. Design of fluorescent probe of Grb2-SH2 domain for fluorescence polarization assays. The probe contains a Grb2-specific phosphotyrosine motif (pYVNV), a fluorescein moiety, and a dipeptide spacer ( $\beta$ Ala- $\beta$ Ala) between the two to prevent the fluorophore from sterically hindering Grb2-binding.

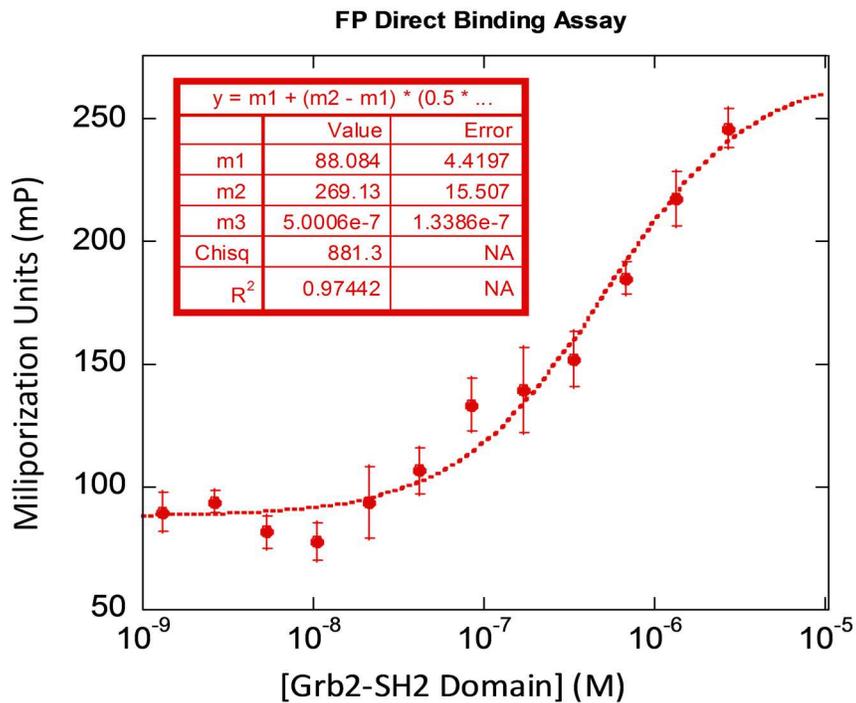


Figure 1.5. Results of FP direct binding experiment between pYVNV-probe and Grb2-SH2 domain.

from the published value of 291 nM by approximately a factor of 3. This provides evidence that the collective introduction of the fluorescein tag at the N-terminus of the fluorescent tracer and/or the inclusion of the MBP tag at the C-terminus of the SH2 domain does moderately attenuate ligand binding. However, this value is acceptable from the perspective of optimizing an inhibitor screen, as the binding affinity is sufficiently high to allow the use of a moderate concentration of probe (5 nM) and protein (0.91  $\mu$ M) in subsequent competitive binding experiments (see Materials and Methods, below, for a detailed explanation of the factors that dictate selection of these final concentrations in competition assays).

## **Materials and Methods**

### *Expression and purification of recombinant Grb2-SH2 domain*

All bacterial cells used for protein expression were purchased from New England Biolabs. The SH2 domains of Grb2 and Tensin1 were expressed and purified as N-terminal MBP fusion proteins. pMAL<sup>TM</sup>-C2x plasmids (New England Biolabs) encoding these fusion constructs under the control of T7 RNA polymerase were gifts from Professor Dehua Pei (Department of Chemistry, Ohio State University). Briefly, BL21(DE3) competent *E. coli* were transformed with plasmid DNA encoding Grb2-SH2 or Tensin1-SH2. The next day, a single colony was picked and grown in 10- mL LB-ampicillin media overnight. This culture was then diluted into 1-L 2XYT + 20% glucose and grown at 37° C for several hours until the sample reached an OD<sub>600</sub> of 0.4 - 0.6. Induction was achieved by the addition of IPTG to a final concentration of 350  $\mu$ M, after which point the culture was incubated at 30° C for 2.5 hours.

The bacteria were then harvested by centrifugation at 2,500 g for 10 min at 4° C. Pellets were re-suspended in Tris-buffered saline solution (20 mM Tris, pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol) containing protease inhibitor cocktail (Sigma-Aldrich). Cells were lysed by rod sonication (4 x 30 second pulses with 1 minute of rest in between each pulse) and centrifuged at 9,000 g for 30 min. Supernatant containing crude lysate was diluted 1:2 in Tris-buffered saline containing 10 mM  $\beta$ -mercaptoethanol and 1x protease inhibitor cocktail (Sigma-Aldrich).

MBP-containing protein was subsequently purified by amylose column chromatography (New England Biolabs). For the purification of recombinant SH2 domain from a 1 L culture of BL21 *E. coli*, 5 mL of pre-swollen, aqueous amylose resin was prepared in a fritted polystyrene column and maintained at 4° C. All buffers and clarified cell lysates were stored on ice throughout the purification process. Care was taken throughout the procedure to ensure the column was submerged in cold buffer or lysates at all times and that at no point was the column allowed to run dry. The resin was gently conditioned with 10 column volumes of amylose column buffer (250 mM NaCl, 50 mM Tris-HCl, pH = 7.4, 1 mM EDTA) under gravity flow at an approximate flow rate of 1 mL/minute. Diluted, clarified cell lysates was then passed over the column twice at an approximate flow rate of 1 mL/minute. Column flow-through was collected and stored for SDS-PAGE analysis. The resin was then washed with 10 column volumes of amylose column buffer. MBP fusion protein was eluted by treating the resin with amylose column buffer containing 2 mM  $\beta$ -mercaptoethanol and 25 mM maltose monohydrate (Sigma-Aldrich). Eluate was collected in 1 mL fractions on ice. Pre-induction BL21, clarified lysates, column flow-through and

eluate fractions were analyzed by SDS-PAGE. Pure fractions were pooled and concentrated using centrifugal 9K MWCO protein concentrators (Thermo Scientific).

#### *Fluorescence polarization direct binding experiments*

Fluorescence polarization assays were conducted in Corning flat-bottom, black 96-well plates. Direct-binding experiments were performed by incubating 5 nM fluorescein-labeled phosphotyrosine-containing ligand (FI- $\beta$ Ala- $\beta$ Ala-pTyr-Val-Asn-Val) with the Grb2-SH2 domain at concentrations ranging from 1 nM to 1.5  $\mu$ M in phosphate-buffered saline (pH 7.2). This allowed the selection of optimum SH2 domain concentrations for competitive binding assays. This direct binding experiment also enabled the calculation of the dissociation equilibrium constant,  $K_d$ , of the Grb2-SH2:FI- $\beta$ Ala- $\beta$ Ala-pTyr-Val-Asn-Val complex. Non-linear regression analysis was performed using the Kaleidagraph software suite.

#### *Fluorescence polarization competitive binding experiments*

Based on the results of direct binding experiments, competition assays were carried out by incubating fluoresceinated pYVNV peptide at a constant concentration of 5 nM with 0.91  $\mu$ M Grb2-SH2 domain and varying inhibitor concentrations. The concentration of Grb2-SH2 domain was selected so as to maximize the change in fluorescence polarization that would be observed upon

inhibitor-induced dissociation of the Grb2-SH2:pYVNV complex. According to the results shown in figure 1.5, at a Grb2-SH2 domain concentration of 0.91  $\mu$ M and a fluorescein-labeled probe concentration of 5 nM, the fluorescence polarization measured is approximately 200 millipolarization (mP) units. At low concentrations of Grb2-SH2 domain (< 10 nM), the observed fluorescence polarization is approximately 95 mP units. Therefore, at this concentration of protein and fluorophore-labeled probe, a change in fluorescence polarization on the order of 100 mP units is expected. This concentration of protein is therefore optimal for competitive binding assays for a variety of reasons: (i) the fluorescence polarization delta of 100 mP provides a strong signal/noise ratio, ideal for the rapid screening of a series of candidate peptide inhibitors; (ii) the polarization signal is not yet saturated at this concentration of Grb2-SH2 domain, which is necessary to ensure that the fluorescence polarization readout is maximally sensitive to effective binding events between Grb2 and inhibitor peptides; and (iii) the sub-micromolar concentration of Grb2-SH2 domain is well within the solubility limit of this protein, as Grb2-SH2 has been shown to aggregate at higher concentrations. Competition assays were carried out by incubating 5 nM pYVNV ligand, 0.91  $\mu$ M Grb2-SH2 with inhibitor concentrations ranging from sub-nM to 125  $\mu$ M in a flat-bottom, black, polystyrene 96-well plate (Corning). Half-maximal inhibitory concentrations ( $IC_{50}$ ) were computed using the Kaleidagraph software.

### **1.6 – Identifying cyclic peptide inhibitors of the Grb2-SH2 domain**

Initial screens of the G1HT series of cyclic peptides to identify a head-to-tail cyclic peptide that lacked a labile disulfide bridge but retained the biological activity of the parent peptide, **G1**, yielded interesting results. Of the 6 designed cyclic peptides, **G1HT2-6** were inactive in FP competition assays up to final

concentrations of 62.5  $\mu\text{M}$ . **G1HT2** and **4** only began to demonstrate inhibitory activity at concentrations exceeding 125  $\mu\text{M}$ , and even at these concentrations the inhibition cannot be unambiguously distinguished from noise. **G1HT1**, however, functioned as a competent inhibitor in this assay, disrupting the Grb2-SH2 domain:pYVNV interaction with an  $\text{IC}_{50}$  of 3.6  $\mu\text{M}$  (figure 1.6). Not only does **G1HT1** retain the activity of the precursor peptide **G1**, this inhibitory activity actually represents a greater than 3-fold increase in potency.

We theorize that this increase in target affinity can be explained at least in part by an increase in conformational restriction. While **G1** was rendered macrocyclic by virtue of a sulfur-sulfur bond (bond dissociation energy approximately 60 kcal/mol), **G1HT1** is tethered by a much stronger amide bond (bond dissociation energy typically greater than 100 kcal/mol).<sup>60</sup> These differences in bond strength are manifested in very different geometric orientations: the sulfur-sulfur bond length of a disulfide bridge is expected to be greater than that of the carbon-

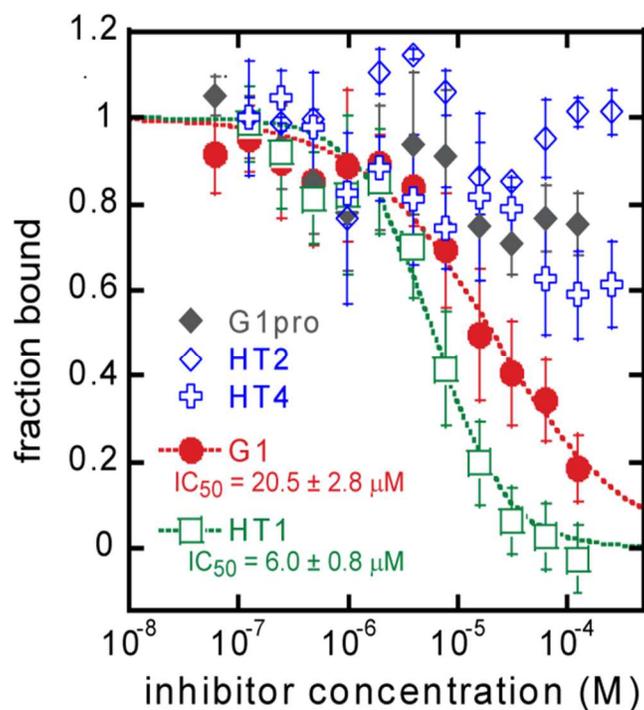


Figure 1.7. Results of FP competitive binding assays in **G1HT** series. No binding was observed for **G1HT3**, **5**, and **6**.

Peptide	Calculated MW	Observed m/z	Species
G1	1321.5	1320.4	[M-H]
G1-pro	1289.5	1288.5	[M-H]
G1HT1	1195.3	1194.2	[M-H]
G1HT2	1209.3	1208.3	[M-H]
G1HT3	1223.3	1222.3	[M-H]
G1HT4	1195.3	1194.2	[M-H]
G1HT5	1209.3	1208.3	[M-H]
G1HT6	1223.3	1222.3	[M-H]
G1HT7	1118.2	1117.3	[M-H]
G1HT8	1174.3	1173.3	[M-H]

Table 1.2. Characterization of Grb2-SH2 domain-targeting cyclic peptides.

nitrogen bond length of a peptide bond. The carbon-nitrogen bond in an amide moiety is between two sp<sup>2</sup>-hybridized trigonal planar centers, and the short bond length, high bond energy and resistance to rotation at room temperature have been rationalized on the basis of partial pi bonding (MO theory) and double-bond character (resonance theory) between the carbon and nitrogen atoms.<sup>61</sup>

Taken together, the relaxed length and enhanced rotational capacity of the sulfur-sulfur bond as compared to the amide bond provide an explanation for the increased conformational restriction in the **G1HT1** system, which appears to have the effect of structurally pre-organizing the phosphotyrosine-like epitope of the molecule as judged by the increase in inhibitory capability.

The different geometric properties of a disulfide-bridged macrocycle and a lactam cyclic peptide also suggest an explanation for the surprising differences in the inhibitory activities among the peptides of the **G1HT** series. At the onset of the inhibitor design process, we theorized that replacement of the disulfide bridge of **G1** with an amide bond to furnish a functionally equivalent peptide would not be a trivial substitution. The results of these initial inhibitor screens confirmed this intuition. Our attempts to mimic the overall structure of **G1** by replacing a disulfide with a motif that possesses a distinct set of physicochemical properties and conformational tendencies were mostly unsuccessful. This result underscores what would eventually become a common observation throughout this project: that seemingly subtle changes in the design of **G1**-inspired macrocycles (addition/deletion of a methylene group, inversion of a stereocenter) can actually lead to large structural changes, which is often a source of attrition in inhibitor discovery in this system. This speaks to the high conformational sensitivity of this cyclic peptide scaffold, as we conclude that the application of conformational constraint by the introduction of a head-to-tail lactam bond into

this macrocycle must be done carefully and precisely. Excessive conformational freedom (e.g. **G1HT3**, **6**) and structural pre-organization of the molecule in an orientation that is stereochemically not suitable for target binding (e.g. **G1HT4**) lead to large, deleterious losses in target affinity. These conclusions are not only drawn from a class of six cyclic peptides, but rather from a larger pool of cyclic peptides as well as bicyclic peptides inspired by **G1** and **G1HT1**.

### **1.7 – Identifying bicyclic peptide inhibitors of the Grb2-SH2 domain**

Following the discovery of **G1HT1**, we sought to begin the process of further structurally rigidifying this macrocycle by installing covalent side-chain-to-side-chain cross-links. We hypothesized that the precise placement of an intramolecular conformational constraint in this system would further enhance the pharmacological profile of **G1HT1**, with the expectation that, as was the case for the **G1HT** peptides, many of the introduced variations could drastically attenuate inhibitory activity. With this in mind, we applied analysis from previous SARs that have been reported for **G1** and used this information in molecular modeling experiments in order to develop a structural model of **G1HT1** bound to the phosphotyrosine-binding site of the Grb2-SH2 domain.

In order to perform this structural investigation, we used the Molecular Operating Environment (MOE) software package (Chemical Computing Group). We began by constructing a model of **G1HT1** in which the side chains of the glutamic acid and tyrosine residues – the putative components of the phosphotyrosine-like pharmacophore – were oriented towards one another such that a hydrogen bond was established between the phenol of tyrosine and the carboxylic acid of glutamic acid. The asparagine side chain was subsequently oriented so as to project outward from the peptide surface, away from the phosphotyrosine-like

surface. This decision was informed by SARs and structural investigations that have demonstrated the presence of an asparagine-binding pocket peripheral to the primary phosphotyrosine-binding site at the Grb2-SH2 domain surface. The dual binding elements of **G1HT1** were then manually inserted into their respective binding pockets on the protein surface using the published crystal structure of the Grb2-SH2 domain (PDB ID: 1TZE).<sup>62</sup> The peptide was subsequently energy minimized in the context of the target protein, with several notable constraints: (i) a Hooke's law force constant of 10 kcal/mol was applied to the side chains of the tyrosine and glutamic acid residues, and (ii) only atoms of the peptide were allowed to move during the simulation, while atoms of the target protein were rendered completely rigid.

The result of the simulation was a structural model of the **G1HT1**:Grb2-SH2 interaction (figure 1.8). This model provided a starting point for the introduction of covalent cross-links, as several potential sites for side-chain substitutions were identified on the basis of spatial proximity. The two most promising pairs of potentially proximal residues included the combinations of Leu2-Met8 and Val6-Tyr9. In this preliminary structural model, the terminal methyl protons of Leu2 are within 2.3 angstroms of the Met8 methyl protons, while the phenolic hydroxyl proton of Tyr9 is within 2.5 angstroms of the terminal methyl protons of Val6. Based on the results of this molecular modeling investigation, we designed a series of bicyclic peptides that introduce aspartic acid and lysine substitutions at these positions so as to facilitate the formation of lactam cross-links between these two positions. With a variety of linker chemistries at our disposal, we initially chose lactam staples as a starting point for bicyclic peptide screens. Lactam bicyclic peptides were expected to be readily accessible synthetic targets that could be achieved using common, commercially available Fmoc-protected

amino acids. In addition, several bicyclic variants containing all-carbon intramolecular tethers were proposed that could be accessed via ring-closing olefin metathesis. Despite the availability of amino acid precursors, we were still faced with the task of devising a synthetic strategy to assemble bicyclic peptides entirely on resin. **To date, there had been no reported examples of such a synthesis.** The advantages to a synthesis in which the entire functionalized macrocycle is assembled on solid support prior are numerous. On-resin cyclization of peptides and peptidomimetics has been shown to yield cyclic products in higher crude purity than cyclization in solution phase. Peptide cyclization on solid support has often resulted in suppressed racemization and decreased quantities of polymeric peptides.<sup>63</sup> These advantages emerge in part from the kinetics of cyclization on resin as opposed to in solution phase, as an elevated relative rate of intramolecular versus intermolecular cyclization is common to on-resin cyclization procedures. High dilution conditions are frequently employed during solution-phase cyclization processes in order to prevent or suppress peptide polymerization due to intermolecular couplings or cyclization events.

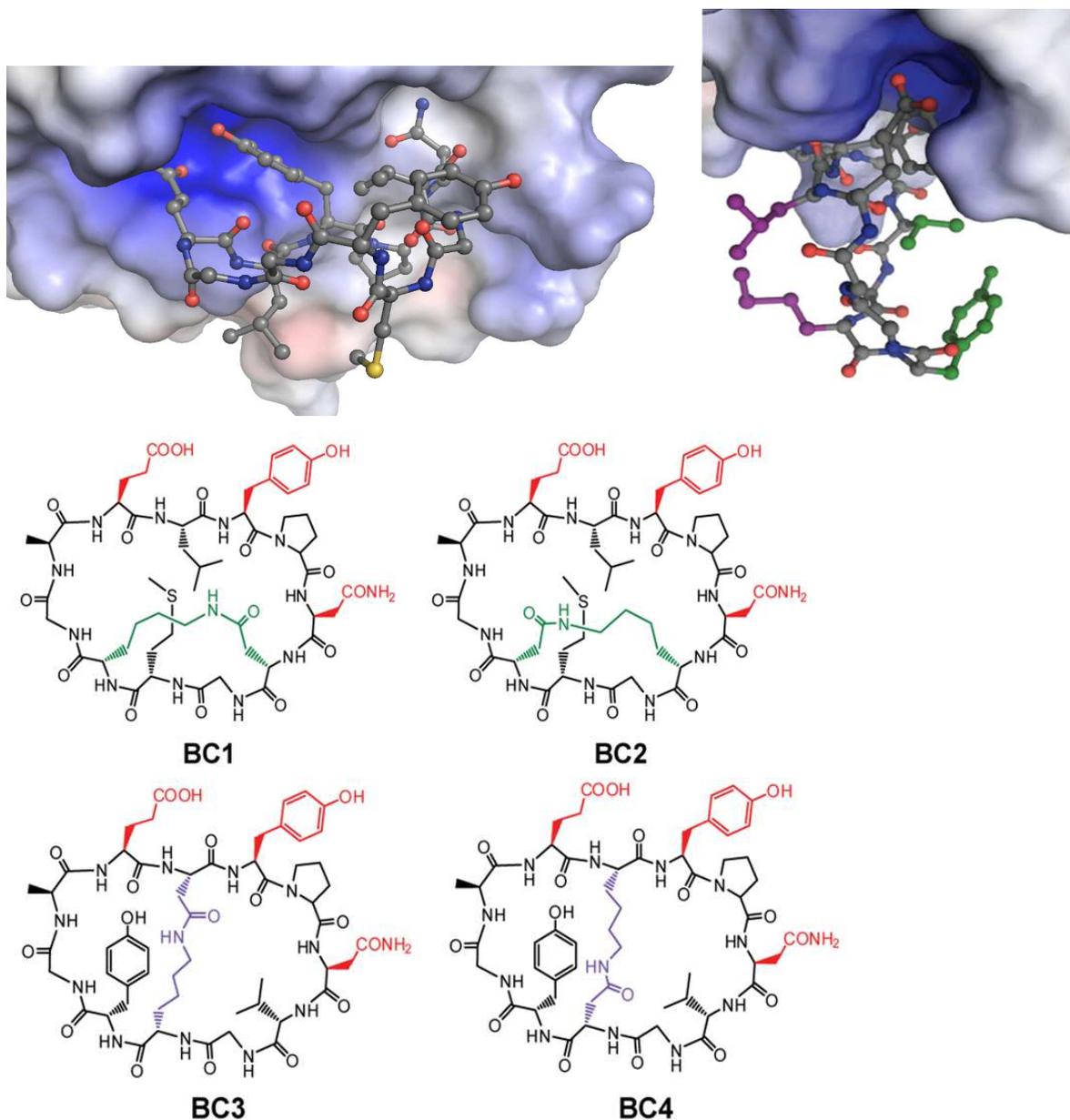


Figure 1.8. *Top*: Molecular modeling experiments suggested potential cross-linking sites within the **G1HT1** scaffold. *Bottom*: Initial lactam-bridged bicyclic peptide series.

These conditions are inconvenient for a number of reasons. Cyclization reactions or peptides have been shown to proceed most robustly in nonvolatile, polar, aprotic solvents that are poorly capable of solvating electronegative nucleophilic

atoms, such as the N-terminus of a peptide chain. Typical solvents include dimethylformamide (DMF) and N-methylpyrrolidone (NMP), which exhibit normal boiling points of 153° C and 202° C, respectively. These elevated boiling points render the removal of large amounts of these solvents *in vacuo* a time-consuming task. More importantly, high dilution conditions often require long cyclization times, which can result in the occurrence of base-promoted side reactions such as racemization, dehydration and aspartimide formation.<sup>64,65</sup> During the on-resin cyclization of a peptide, a side-chain functionality is typically bound to the solid support via a covalent linkage that is labile under conditions that are orthogonal to those employed during the assembly of the linear peptide chain.<sup>63</sup> As the molar substitution of the resin is tantamount to the effective concentration of the protruding peptide chain, a high dilution environment can be simulated by controlling the loading capacity of the resin. Typical resin substitutions employed for the synthesis of cyclic peptides on resin range from 0.1 to 0.4 mmol of linker per gram of resin. These low-loading capacity resins can be purchased commercially or accessed from higher loading resins by loading the first amino acid of a peptide chain using sub-stoichiometric quantities of amino acid and subsequently capping the resin via treatment with reagents capable of irreversibly attenuating the reactivity of the resin handle. This is commonly achieved by the acylation of residual free amines of an amino resin with acetic anhydride or the etherification of unreacted 2-chlorotrityl chloride handles with methanol.<sup>63</sup>

Though synthetic strategies for the assembly of cyclic peptide macrocycles on solid support have been established, the challenge of accessing peptide bicycles entirely on resin necessitated the optimization of an orthogonal protection strategy. We sought to devise a synthesis that would enable the site-selective

removal of protecting groups from side chain or C-terminal functionality without cleaving protecting groups or other chemical moieties elsewhere in the peptide. To this end, two synthetic strategies were ultimately utilized (see Materials and Methods) that could enable independent cyclization of the side chains and N- and C-termini. As was the case for the synthesis of the **G1HT** series of peptides, the glutamic acid side chain provided a convenient functionality for tethering of a side chain to 4-(hydroxymethyl)phenoxymethyl polystyrene (Wang resin). These synthetic routes differ in the order of cyclization events: route A proceeds by cross-linking the Lys and Asp side chains first, and subsequently cyclizing the N- and C-termini immediately prior to global deprotection and cleavage from the resin. Route B proceeds through assembly of a head-to-tail macrocyclic intermediate on resin, followed by selective Lys/Asp deprotection and inter-side chain coupling. It was ultimately determined that the lactam-bridged bicyclic peptides proposed herein could be synthesized by one of these two methods, and that the highest-yielding synthetic route could be correlated to the location of the lactam tether within the peptide.

For the synthesis of olefin-stapled bicyclic peptides, a third synthetic strategy was employed. Here, the decision of the ordering of cyclization events was informed by several factors: (i) the olefinic amino acids that will ultimately comprise the intramolecular tether are stable to the basic conditions used to remove the N-terminal Fmoc protection, and (ii) the ring-closing metathesis reaction can be performed under more forcing conditions than typical amide-bond forming reactions. Therefore, it was ultimately determined that the optimal synthetic strategy was one in which side-chain tethering was deferred to the end of the synthesis, regardless of the location of the cross-link.

Based on the results of molecular modeling investigations of **G1HT1**, an initial series of lactam-bridged bicyclic peptides was designed as a best-odds starting point for the design of improved Grb2 inhibitors (figure 1.8). Upon testing these peptides in our optimized competitive binding FP assay, it became apparent that many of the proposed cross-links were deleterious to inhibitory activity, similar to the results of the G1HT series. One of these peptides, **BC1**, not only retained the inhibitory capacity of the precursor peptide, but it exhibited a ten-fold increase in potency relative to **G1HT1** and a nearly 60-fold increase in potency relative to disulfide-bridged **G1**. Interestingly, even the closely-related analog **BC2** exhibited a dramatic reduction in biological activity, as inhibition was not observed *in vitro*

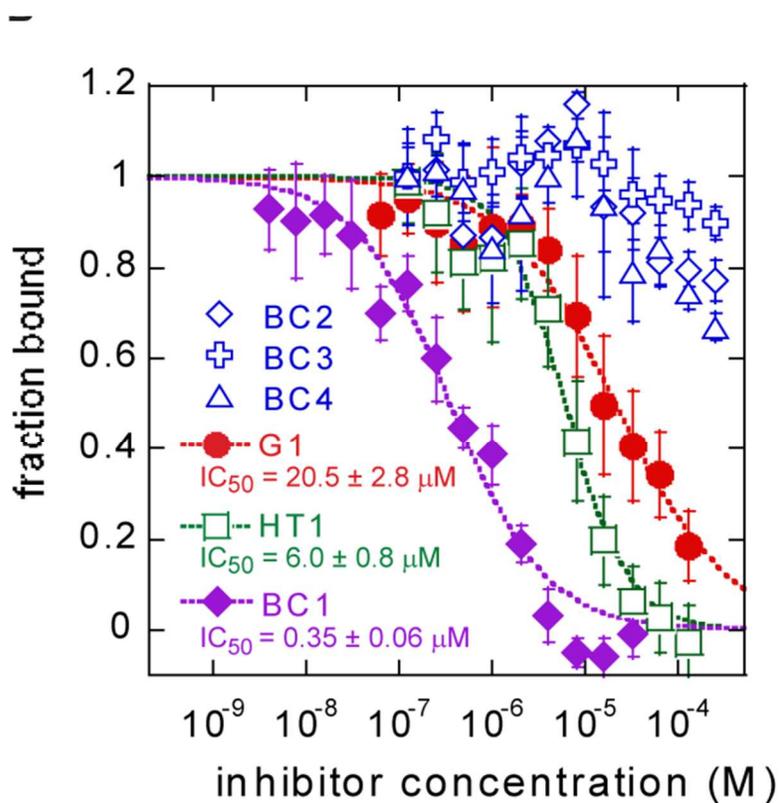


Figure 1.9. Results of initial bicyclic peptide series in FP competition assay. Most of the initially tested bicyclic peptides did not preserve the inhibitory activity of the parent compounds, **G1** and **G1HT1**. For clarity, many of these variants have been excluded here (see figure 1.11 for results of remaining bicyclic peptide variants). **BC1**, however, exhibited an improved affinity for the Grb2-SH2 domain based on the calculated  $IC_{50}$  value of  $0.35 \mu\text{M}$ .

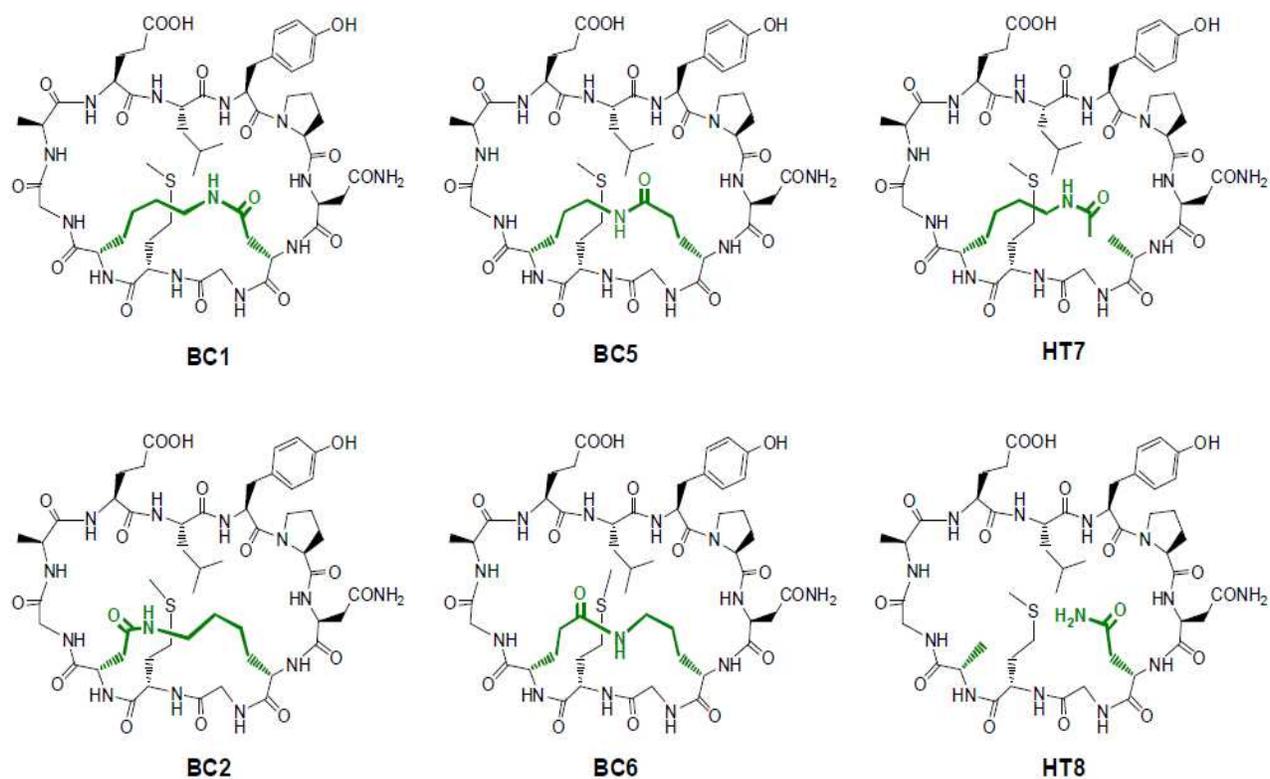


Figure 1.10. Additional cyclic and bicyclic peptides designed based on the structures of **BC1** and **G1HT1**.

up to a concentration of 125  $\mu\text{M}$ . **BC5**, the olefinic analog of **BC1**, also did not exhibit significant inhibition of Grb2. Collectively, these observations provide evidence that the precise locations of the C=O and N-H moieties of the side-chain cross-link are crucial to the inhibitory activity of the peptide bicycle.

To further probe the positional tolerance of the amide bond in the side-chain cross-link, we designed and synthesized variants of **BC1** that alter the location of the amide functional group and that are void of this moiety entirely (figure 1.10-1.12). To understand which aspects of the staple contributed to the higher potency of **BC1**, we tested analogs lacking a covalently linked staple. **HT7** differs from **HT1** in that an Asn is substituted for Val7, and **HT8** differs from **HT1** in that an acetylated Lys is substituted for Tyr10. Each of these possesses part of the staple of **BC1**, including the amide group, but is still monocyclic. **HT7** and **HT8** show little inhibitory activity (figure 1.11). These results show that the conformational constraint of the staple, and not the simple addition of a specific amide group, was responsible for **BC1**'s increased potency. They also provide further evidence of the exquisite conformational sensitivity of the macrocycle scaffold, since **HT7** and **HT8** have effectively lost the inhibitory potency of **HT1**. We then tested two additional variants of **BC1** which maintained the staple length but altered the position and orientation of the amide bond within the staple (figure 1.10). All of these were also inactive in the competition assay. Small changes in the staple thus led to large losses in inhibitory potency, implying that the conformation of **BC1** is precisely controlled by the staple's composition and geometry.

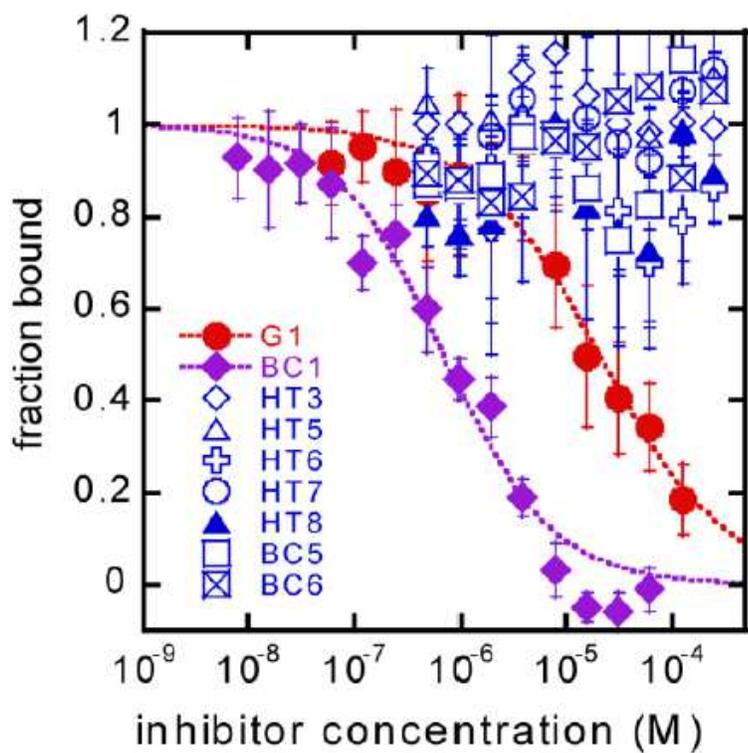


Figure 1.11. FP competitive binding assay results of all negative peptides. **BC1** and **G1** are plotted for comparison.

Peptide	Calculated MW	Observed m/z	Species
BC1	1158.3	1157.3	[M-H]
BC2	1158.3	1157.3	[M-H]
BC3	1176.2	1175.3	[M-H]
BC4	1176.2	1175.3	[M-H]
BC5	1158.3	1157.3	[M-H]
BC6	1158.3	1157.3	[M-H]

Table 1.3. Characterization of Grb2-SH2 domain-targeting, lactam-bridged bicyclic peptides.

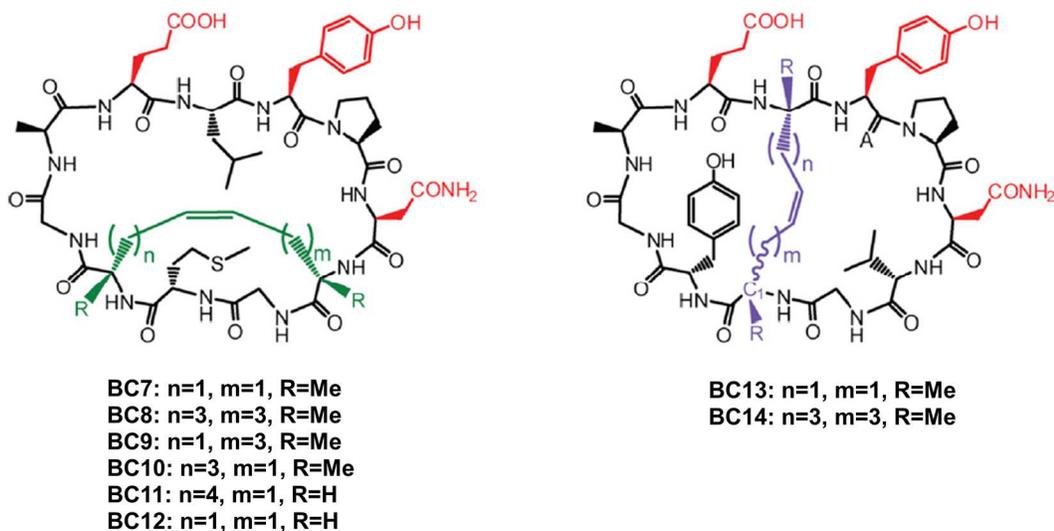


Figure 1.12. Design of olefin-bridged bicyclic peptides designed to antagonize the Grb2-SH2 domain.

Peptide	Calculated MW	Observed m/z	Species
BC7	1127.3	1126.1	[M-H]
BC8	1183.4	1182.6	[M-H]
BC9	1155.3	1154.1	[M-H]
BC10	1155.3	1154.5	[M-H]
BC11	1141.3	1140.1	[M-H]
BC12	1099.2	1098.4	[M-H]
BC13	1145.2	1144	[M-H]
BC14	1201.3	1200.1	[M-H]

Table 1.4. Characterization of Grb2-SH2 domain-targeting, olefin-bridged bicyclic peptides.

## Materials and Methods

### *Bicyclic peptide synthesis: Route A (Used for synthesis of **BC1**-like peptides)*

Starting with low-loaded (0.2 – 0.3 mmol/g) Wang resin, Fmoc-Glu-ODmab was loaded using a DMAP/HBTU coupling procedure as described.<sup>44</sup> The peptide chain was extended with repeated cycles of Fmoc-deprotection (20% piperidine in DMF) and coupling (5 eqv. amino acid, 5 eqv. PyBOP, 5 eqv. HOBt, and 13 eqv. DIPEA). Following assembly of the linear chain up until the second amino acid of the intramolecular cross-link, the side-chain protecting groups of Asp and Lys were removed by treatment with tetrakis(triphenylphosphine) palladium reagent (3.0 equivalents) in a 37:2:1 solution of chloroform:acetic acid:N-methylmorpholine. The reaction mixture was shielded from light and allowed to proceed for 3 hours at room temperature under gentle agitation on an orbital shaker. The resin was subsequently washed thoroughly with DMF (3 x 3 minutes), 0.1% (w/v) diethyldithiocarbamic acid in DMF (2 x 5 minutes), and 0.1% DIPEA in DMF (2 x 4 minutes). The resin was washed twice more in DMF before proceeding to the removal of the N-terminal Fmoc protecting group in the usual way. The peptide chain was then completed by repeated Fmoc

deprotections and amino acid couplings. At the conclusion of the linear sequence assembly, the C-terminal Dmb protecting group was removed by treatment of the resin with a 2% hydrazine in DMF solution (4 x 3 minutes). The resin was then washed thoroughly in DMF (2 x 4 min) before treating with 20% piperidine in DMF to ensure complete removal of the N-terminal Fmoc protecting group (2 x 8 minutes). The resin was then washed with DMF (2 x 4 minutes) and treated with PyBOP, HOBT and DIPEA for 24-48 hours at room temperature under Ar atmosphere. Coupling reagents were replenished every 12 hours, and reaction efficiency was periodically determined by small-scale cleavage and MALDI MS analysis and/or ninhydrin tests. Large-scale cleavage and global deprotection was achieved by treatment of the resin with 94:2.5:2.5:1 TFA:EDT:H<sub>2</sub>O:TIS. The reaction was allowed to proceed at room temperature for 2 hours under mild agitation. The resin was filtered and the filtrate concentrated by rotary evaporation to a crude oil. This was subsequently triturated in ice-cold diethyl ether, yielding white precipitate immediately. Precipitates were centrifuged in 50-mL conical vials for 10 minutes at 3,000 x g and were dissolved in DMSO:acetonitrile:water mixtures for purification by HPLC.

*Bicyclic peptide synthesis: Route B (Used for synthesis of **BC4**-like peptides)*

Amino acid loading and peptide synthesis was carried out as detailed above, except the entire linear peptide chain was assembled by repeated Fmoc deprotections and amino acid couplings before removing any side-chain protecting groups. Upon completion of the linear chain assembly, the C-terminal allyl ester was cleaved by treatment with palladium as described above. The resin was washed as described and treated with piperidine to remove the N-terminal Fmoc. Head-to-tail cyclization was achieved by treating the resin with

PyBOP, HOBt and DIEA for 24-48 hours, and cyclization efficiency was monitored by small-scale cleavages and ninhydrin tests. The resin was then washed in DMF (3 x 5 minutes) and the side-chain iv-Dde protecting group (Lys) and Dmab protecting group (Asp) were removed by treatment of the resin with 2% hydrazine in DMF as described above. The resin was washed as described and treated with PyBOP, HOBt and DIEA for 48-72 hours. Coupling reagents were replenished every 12 hours.

Cleavage and global deprotection was carried out by first drying the solid resin after successive rinses in dichloromethane and methanol. The dried peptidyl-resin was then treated with a solution of 94:2.5:2.5:1 trifluoroacetic acid (TFA):water:ethanedithiol (EDT):triisopropylsilane (TIS) for 2 hr at room temperature under gentle agitation. The cleavage cocktail was then rotary-evaporated to approximately 2 mL and was subsequently triturated in ice-cold ether. The suspension was then centrifuged, and the crude pellet was dissolved in DMSO.

#### *Peptide purification*

Peptides were purified by reverse phase HPLC using a preparative scale C8 column (Vydac, 0.46  $\mu\text{m}$  particle size) and a water/acetonitrile/TFA solvent system. Peptides were purified using a 5-55% acetonitrile gradient over 30 minutes. Final purity of all cyclic peptide products was at least 95% as determined by analytical re-injections of purified products.

## 1.8 – Assessing the selectivity of cyclic and bicyclic peptides for the SH2 domain of Grb2

A great deal of structure-activity data on **G1**, **HT1**, and **BC1** has demonstrated that the ability to bind Grb2-SH2 is highly conformation-dependent.<sup>49–51</sup> The high sensitivity of **BC1** to any conformational perturbation led us to ask whether it might be highly selective for Grb2-SH2 over other SH2 domains. Since binding requires not only the pTyr-mimicking epitope formed by Glu2 and Tyr4, but also the correct conformation of the entire molecule, we reasoned that **BC1** might discriminate between variant loops outside the pTyr-binding pocket. To test the specificity of our compounds, we set up a second competition fluorescence polarization assay using the SH2 domain from tensin-1 (tensin-SH2, figure 1.13). Previous screening efforts to determine consensus peptide sequences that bind the tensin-SH2 domain reveal that this protein recognizes phosphorylated tyrosine in the context of a Asp-Asn-Val sequence.<sup>66–68</sup> Given the similarity of this consensus amino acid sequence to the pTyr-Val-Asn-Val consensus for the Grb2-SH2 domain, we reasoned that the tensin-SH2 domain constitutes a stringent test for binding specificity. We found that tensin-SH2 binds the same pTyr-Val-Asn-Val probe that binds Grb2-SH2, with a similar dissociation constant (figure 1.13). This finding is consistent with values from other assays, and also with prior work establishing similar consensus binding sequences for these SH2 domains.<sup>59</sup> Surprisingly, **G1** inhibits tensin-SH2 more potently than it does Grb2-SH2 (IC<sub>50</sub> for tensin-SH2 of 6.4 ± 0.3 μM), while **HT1** and **BC1** inhibit tensin-SH2 with much less potency (IC<sub>50</sub>'s of 59 ± 4 and 23 ± 1 μM, respectively). Thus, **BC1** exhibits 66-fold selectivity for Grb2-SH2 over tensin-SH2, compared to 0.31-fold for **G1**. This represents a 200-fold increase in selectivity for Grb2-SH2. It should be noted, however, that at least part of the promiscuity of **G1** can be rationalized

on the electrostatic character of the peptide immediately C-terminal to the ELY motif, as the presence of a second glutamic acid likely contributes favorably to binding a peripheral site on the tensin surface. This is in line with the positional preference for an acidic residue in the +1 position in consensus sequence screens.

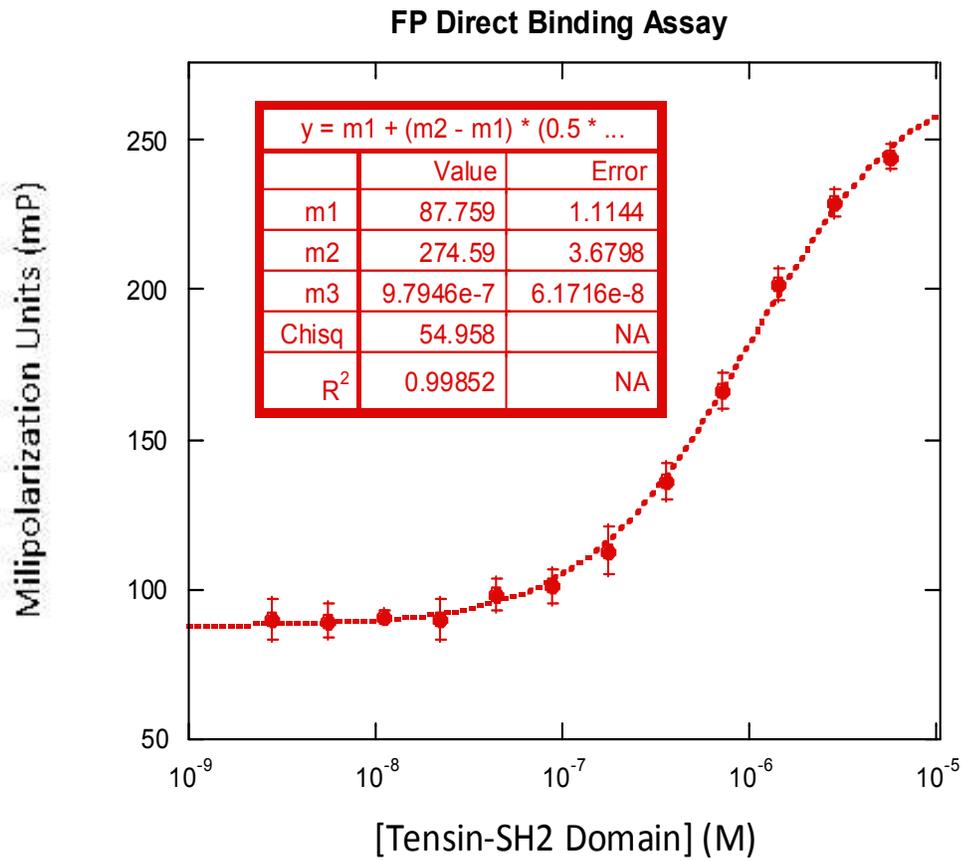


Figure 1.13. Results of FP direct binding assay between pYVNV-containing fluorescent probe and recombinant Tensin-SH2 domain.

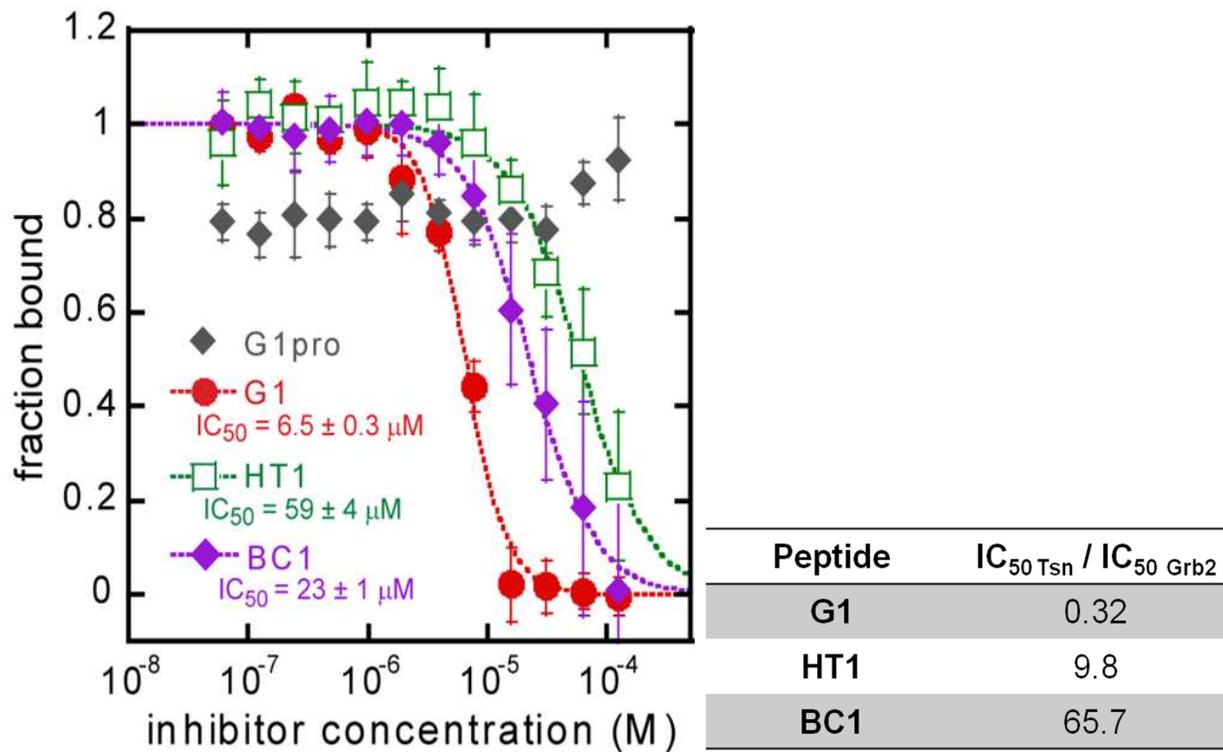


Figure 1.14. Results of FP competitive binding experiments to determine the inhibitory activity of Grb2-SH2 domain-targeting peptides against the Tensin-SH2 domain.

### Materials and Methods

#### *Expression and purification of recombinant Tensin-SH2 domain*

All bacterial cells used for protein expression were purchased from New England Biolabs. The SH2 domains of Tensin1 were expressed and purified as N-terminal MBP fusion proteins. pMAL™-C2x plasmids (New England Biolabs) encoding

these fusion constructs under the control of T7 RNA polymerase were gifts from Professor Dehua Pei (Department of Chemistry, Ohio State University). Briefly, BL21(DE3) competent *E. coli* were transformed with plasmid DNA encoding Tensin1-SH2. The next day, a single colony was picked and grown in 10- mL LB-ampicillin media overnight. This culture was then diluted into 1-L 2XYT + 20% glucose and grown at 37° C for several hours until the sample reached an OD<sub>600</sub> of 0.4 - 0.6. Induction was achieved by the addition of IPTG to a final concentration of 350 μM, after which point the culture was incubated at 30° C for 2.5 hours.

The bacteria were then harvested by centrifugation at 2,500 g for 10 min at 4° C. Pellets were re-suspended in Tris-buffered saline solution (20 mM Tris, pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol) containing protease inhibitor cocktail (Sigma-Aldrich). Cells were lysed by rod sonication (4 x 30 second pulses with 1 minute of rest in between each pulse) and centrifuged at 9,000 g for 30 min. Supernatant containing crude lysate was diluted 1:2 in Tris-buffered saline containing 10 mM β-mercaptoethanol and 1x protease inhibitor cocktail (Sigma-Aldrich).

MBP-containing protein was subsequently purified by amylose column chromatography (New England Biolabs). For the purification of recombinant SH2 domain from a 1 L culture of BL21 *E. coli*, 5 mL of pre-swollen, aqueous amylose resin was prepared in a fritted polystyrene column and maintained at 4° C. All buffers and clarified cell lysates were stored on ice throughout the purification process. Care was taken throughout the procedure to ensure the column was submerged in cold buffer or lysates at all times and that at no point was the column allowed to run dry. The resin was gently conditioned with 10 column

volumes of amylose column buffer (250 mM NaCl, 50 mM Tris-HCl, pH = 7.4, 1 mM EDTA) under gravity flow at an approximate flow rate of 1 mL/minute. Diluted, clarified cell lysates was then passed over the column twice at an approximate flow rate of 1 mL/minute. Column flow-through was collected and stored for SDS-PAGE analysis. The resin was then washed with 10 column volumes of amylose column buffer. MBP fusion protein was eluted by treating the resin with amylose column buffer containing 2 mM  $\beta$ -mercaptoethanol and 25 mM maltose monohydrate (Sigma-Aldrich). Eluate was collected in 1 mL fractions on ice. Pre-induction BL21, clarified lysates, column flow-through and eluate fractions were analyzed by SDS-PAGE. Pure fractions were pooled and concentrated using centrifugal 9K MWCO protein concentrators (Thermo Scientific).

#### *Fluorescence polarization direct binding experiments*

Fluorescence polarization assays were conducted in Corning flat-bottom, black 96-well plates. Direct-binding experiments were performed by incubating 5 nM fluorescein-labeled phosphotyrosine-containing ligand (FI- $\beta$ Ala- $\beta$ Ala-pTyr-Val-Asn-Val) with the Grb2-SH2 domain at concentrations ranging from 1 nM to 1.5  $\mu$ M in phosphate-buffered saline (pH 7.2). This allowed the selection of optimum SH2 domain concentrations for competitive binding assays. This direct binding experiment also enabled the calculation of the dissociation equilibrium constant,  $K_d$ , of the Grb2-SH2:FI- $\beta$ Ala- $\beta$ Ala-pTyr-Val-Asn-Val complex using the Kaleidagraph software package.

#### *Fluorescence polarization competitive binding experiments*

Based on the results of direct binding experiments, competition assays were carried out by incubating fluoresceinated pYVNV peptide at a constant concentration of 5 nM with 1.2  $\mu$ M Tsn-SH2 domain and varying inhibitor concentrations. The concentration of Tsn-SH2 domain was selected so as to maximize the change in fluorescence polarization that would be observed upon inhibitor-induced dissociation of the Tsn-SH2:pYVNV complex. According to the results shown in figure 1.13, at a Tsn-SH2 domain concentration of 1.2  $\mu$ M and a fluorescein-labeled probe concentration of 5 nM, the fluorescence polarization measured is approximately 200 millipolarization (mP) units. At low concentrations of Grb2-SH2 domain (< 10 nM), the observed fluorescence polarization is approximately 90 mP units. Therefore, at this concentration of protein and fluorophore-labeled probe, a change in fluorescence polarization on the order of 100 mP units is expected. Competition assays were carried out by incubating 5 nM pYVNV ligand, 1.2  $\mu$ M Tsn-SH2 with inhibitor concentrations ranging from sub-nM to 125  $\mu$ M in a flat-bottom, black, polystyrene 96-well plate (Corning). Half-maximal inhibitory concentrations ( $IC_{50}$ ) were computed as described previously for Grb2-SH2 screens (*vide supra*).

### **1.9 – Investigating the serum stability of cyclic and bicyclic peptides**

Natural peptide bicycles such as  $\alpha$ -amanitin and sunflower trypsin inhibitor I are uniquely resistant to proteolytic degradation.<sup>69</sup> To test whether our designed peptide bicycles were similarly stable, we optimized an HPLC-based degradation assay to monitor the lifetime of constrained peptides in buffered human serum (see Materials and Methods). Briefly, we incubated 0.9 mg/mL of **G1**, **G1-Pro**, **HT1**, and **BC1** in human serum and monitored their degradation by HPLC (figure 1.15). **G1** and **G1-Pro** showed steady degradation, with less than 30% of the

peptide remaining after seven hours and no detectable amounts of peptide remaining after 24 hours. **HT1** was only 50% degraded after 24 hours. **BC1**, by contrast, remained completely intact after 24 hours and was less than 15% degraded after 48 hours in human serum.

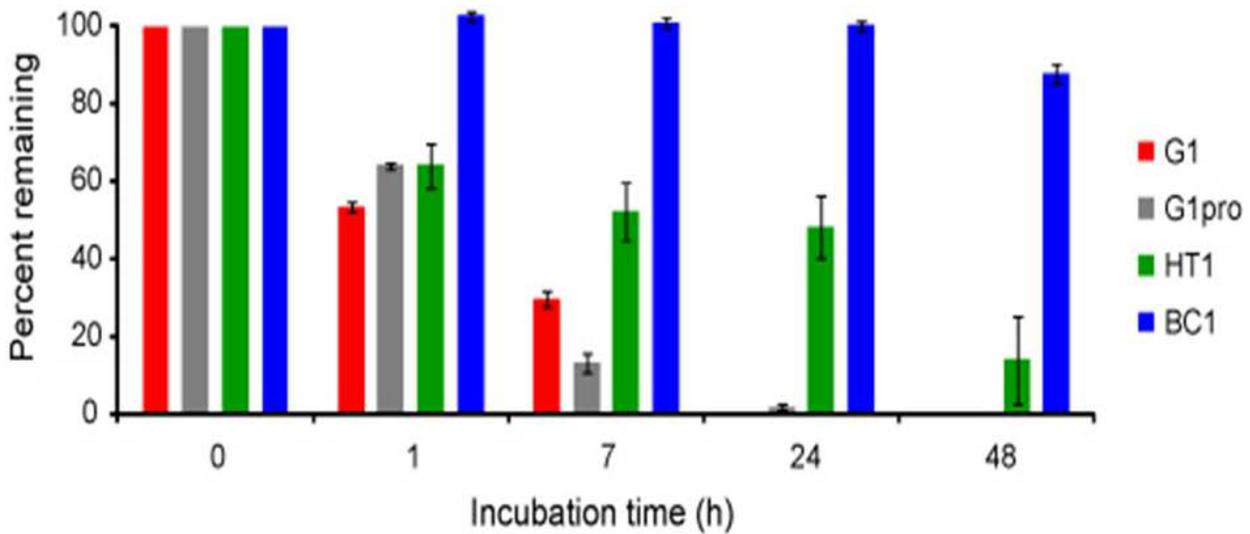


Figure 1.15. Human serum stability assay. Error bars represent the standard deviation of three independent trials.

## Materials and Methods

### *HPLC-based serum stability assay*

The human serum stability assay was performed essentially as described . Human male type-AB serum (Sigma-Aldrich) was diluted 9:1 in phosphate-buffered saline. Peptides were incubated in buffered serum at 37° C at a final concentration of 0.9 mg/mL. At the indicated times, 50 µL aliquots were removed and quenched by adding 200 µL ice-cold methanol. Serum proteins were pelleted, and the supernatant was analyzed by HPLC. Values were normalized to the zero time point. Error bars represent three independent trials.

### **1.10 – Peptide pre-organization enables recognition by anti-phosphotyrosine antibodies**

The peptide bicycle strategy is based on the idea that structural stabilization of the binding epitope will promote greater inhibitory potency and selectivity. For **BC1**, this would mean that the bicyclic structure rigidly pre-organizes the pTyr-mimicking epitope. To test this hypothesis, we spotted serial dilutions of **G1**, **G1-Pro**, **HT1**, and **BC1** onto nitrocellulose membranes and probed them with the commonly used anti-pTyr antibodies 4G10 and PY20 (figure 1.16). EGF-stimulated cell lysate and our own pTyr-containing probe peptide were used as controls to validate the peptide spotting and blotting procedures. **BC1** consistently captured both anti-pTyr antibodies, while **G1**, **G1-Pro**, and **HT1** were unable to capture anti-pTyr antibodies. This demonstrated that **BC1** can maintain a conformation that mimics pTyr even when adsorbed to nitrocellulose. This provides direct evidence that the bicyclic structure of **BC1** successfully pre-organizes the molecule into its protein-binding structure.



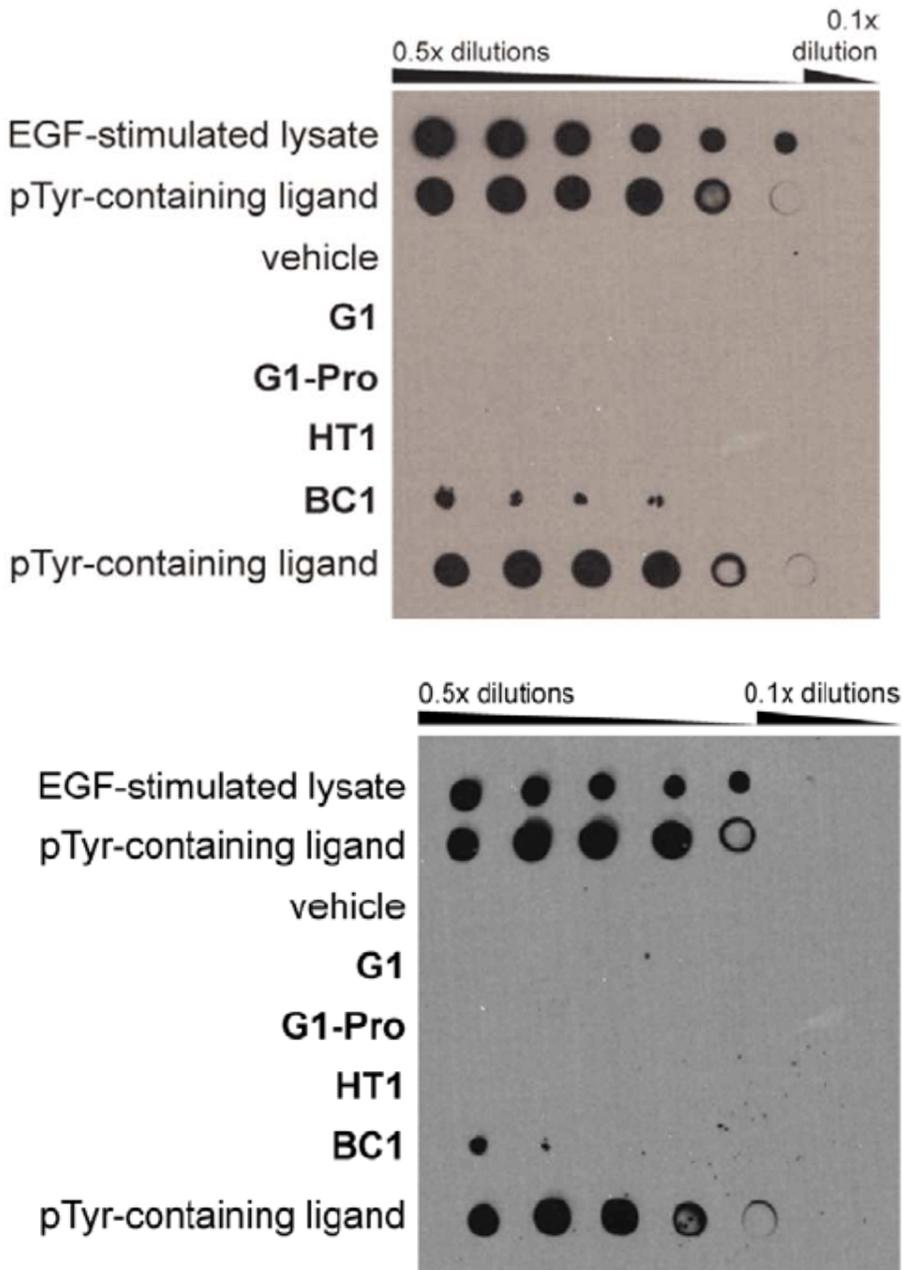


Figure 1.16. Peptide dot blot reveals that bicyclic **BC1** is capable of capturing anti-phosphotyrosine antibodies, even when adsorbed to nitrocellulose membrane. The peptides were spotted directly onto nitrocellulose and probed with either anti-phosphotyrosine antibody 4g10 (EMD Millipore, top) or pY20 (bottom). Initial peptide concentration is 750  $\mu$ M.

## **Materials and Methods**

### *Anti-phosphotyrosine dot blot*

Peptides were diluted to 500  $\mu$ M in Tris-buffered saline, pH 7.4, with 0.1% Tween-20. These stocks were serially diluted and 1  $\mu$ L drops were spotted directly onto a nitrocellulose membrane. Spots were allowed to completely dry. The membrane was blocked with 0.2% bovine-serum albumin for 1 h. The membrane was then incubated with anti-pTyr antibody (4G10 or PY20, Millipore), followed by HRP-conjugated secondary antibody (goat-anti-Mouse IgG:HRP conjugate, Sigma-Aldrich). The membrane was then washed and incubated with SuperSignal Chemiluminescent Substrate (Pierce) and imaged on film.

## **Chapter Two – Investigating the cell-penetrating capabilities of cyclic and bicyclic phosphotyrosine mimetics**

### **2.1 – General rules for the cellular uptake of cationic and amphipathic peptides**

There is currently a vast body of work dedicated to understanding the types of molecular architectures that are capable of being internalized by mammalian cells. Cultivating an understanding of what types of molecular architectures can penetrate the cell and how they do so is paramount to the development of biological probes and therapeutic lead compounds that can effectively

antagonize intracellular targets. While the last two decades have seen a resurgence in the use of constrained peptides for the development of inhibitors of protein-protein interactions,<sup>13,20</sup> there remains a paucity of knowledge of the structural and electronic requirements that inform the internalization of these compounds into mammalian cells. Seminal work in the field has shown that certain classes of peptides are capable of being taken up by mammalian cells.<sup>70</sup> However, most of these are polycationic sequences or amphipathic alpha-helical peptides.<sup>71-73</sup> There are strikingly few reports of negatively charged peptides that are internalized by mammalian cells, and quite often a single negative charge is enough to prevent internalization in small molecules and peptides alike.<sup>22,74-78</sup> Moreover, very little is known about the mechanisms by which constrained, non-cationic peptides are taken up by cells.<sup>16,17,74,79</sup> This knowledge is critical to the development of this class of compounds as a potential therapeutic modality for intramolecular targets.

In the present section of this dissertation, investigations will be presented that seek to characterize the cellular uptake of a series of neutral and anionic peptides that are conformationally constrained by virtue of head-to-tail cyclization (**G1HT** series) and side-chain-to-side-chain cross links (**BC** series). This section will begin with a brief review of the general rules that have been established for the internalization of polycationic peptides and constrained, amphipathic  $\alpha$ -helices. Work will then be presented in this section that demonstrates that the neutral and anionic constrained peptide derivatives of the **BC** and **G1HT** series are capable of being internalized by a mammalian breast epithelial cancer line, despite the frequent presence of a net negative charge. We begin to characterize

the mechanism of uptake of these peptides by probing the effects of temperature and modulators of endocytosis in order to elucidate the process by which these molecules are internalized. Collectively, the data presented herein provides evidence for distinct structural features that promote cellular internalization and suggests a starting point for the development of a set of general rules that dictate the internalization of neutral and anionic constrained peptides.

A great deal of work has been dedicated to understanding the precise mechanisms by which cationic peptides are internalized by mammalian cells. Common examples of cell-penetrating peptides with cationic character include the HIV-derived transactivator of transcription (TAT) peptide, the antennapedia peptide, and polyarginine, among others.<sup>72,80</sup> It has been shown that TAT and TAT-like peptides enter the mammalian cell by multiple mechanisms, including lipid raft macropinocytosis<sup>81</sup> as well as via membrane destabilization that occurs past a certain threshold concentration (typically 10  $\mu$ M) of peptide that varies from one cell line to another.<sup>70</sup> Many of these cationic peptides have been shown to undergo internalization by energy-dependant processes, including endocytosis and macropinocytosis. TAT has been shown to be capable of binding anionic proteoglycans on the extracellular face of the plasma membrane and subsequently induce actin remodeling events that facilitate a GTP-dependant pinching off of the membrane and concomitant internalization of the peptide. Seminal efforts have also revealed the ability of other polycationic sequences to induce lipid raft-mediated endocytosis.<sup>71,82,83,81</sup>

In addition to the variety of energy-dependant mechanisms by which these positively charged peptides are internalized, still other investigations have demonstrated the ability of these peptides to be internalized under conditions that slow or inhibit molecular events crucial to endocytosis. Analysis of the cellular

uptake of cationic peptides such as TAT and antennapedia by live-cell imaging of fluorescently-labeled peptides and flow cytometry in the presence of endocytosis inhibitors and at reduced temperatures have demonstrated that these peptides are indeed capable of being internalized in an energy-independent manner as well.<sup>82</sup> It has been shown that this route of internalization is operative above a certain threshold concentration of the cationic peptide that appears to be specific to a given mammalian cell line. Suggested mechanistic features that could account for this mode of internalization include local destabilization of the plasma membrane through electrostatic interactions between the peptide and sulfate-rich carbohydrate moieties on the extracellular surface of the membrane. Alternative mechanistic proposals suggest that these cationic peptides are capable of inducing the formation of pores within the membrane, which subsequently enables these molecules to diffuse passively into the cell according to its chemical gradient.<sup>71</sup>

It is important to emphasize that under conditions in which endocytic modes of internalization are operative, the peptides that are internalized remain confined to endosomes and are hence topologically exterior to the cell. Peptides that are endocytosed are not capable of accessing the cytosol until these molecules escape the endosomes, a process that has been shown to occur at unique phases of the endosomal maturation process contingent upon the ionic character and spatial charge distribution of the internalized peptide. Recent investigations aimed at elucidating the precise points in the progression of the endosome during which cationic peptides achieve access to the cytosol have revealed that certain polyarginine-containing structures are capable of escape from early (Rab5+) endosomes, while still others, such as Tat and Arg<sub>8</sub>, are not capable of endosomal escape until maturation to a later (Rab7+) state.<sup>84</sup> In each case, the

key determinant of the temporality of cytosolic access appears to be the distribution and relative geometric orientations of cationic residues upon the peptide surface.

Previous investigations on the cell-penetrating capabilities of constrained peptides have focused primarily on the uptake of stapled  $\alpha$ -helical peptides. This class of conformationally-restricted peptides represents a highly sought-after synthetic target that has been exploited in the inhibition of many medically-relevant protein-protein interactions. The structural pre-organization of an  $\alpha$ -helix has been achieved by the application of a multitude of chemical cross-linking strategies, many of which feature the tethering of two solvent-exposed side-chains on the same face of the helix with a linker of the appropriate length and stereochemical configuration to structurally stabilize one or more helical turns.<sup>13,22,33</sup> This paradigm has been executed using a broad variety of linker strategies, including the formation of lactam bridges, disulfide bonds, coordination of side-chain carboxylate functions with divalent metal cations, ring-closing olefin metathesis, and recently, the tethering of cysteine residues with bifunctional electrophilic alkyl or aryl linkers.<sup>20,76,85–87</sup> To date, a series of general rules for the development of stapled helical peptides that target helix-binding sites at PPI interfaces have been established.<sup>13</sup> Not surprisingly, much of the research that has been performed to investigate the cell-penetrating capabilities of these compounds has focused on characterizing the internalization of olefin-stapled helical peptides, as these compounds comprise some of the most influential examples of tethered peptides as inhibitors of intracellular protein targets.

Olefin-stapled peptides have been shown to be internalized by mammalian cells to varying extents depending upon the physicochemical properties of the peptide

sequence. One of the most important factors that dictates the cellular uptake of these peptides is electrostatic charge: most cell-penetrant stapled helical peptides are amphipathic or cationic.<sup>72,88–91</sup> Negative charges within the peptide sequence have historically been associated with poor cell-penetrating propensity. In certain cases, the presence of even a single negative charge has been sufficient to attenuate cellular uptake in this class of compounds (as well as in small molecules).<sup>13,75</sup> Net neutral or positive charge in stapled helices has been correlated efficient cellular uptake as judged by multiple lines of evidence, including live cell confocal microscopy, FACS analysis and emerging label-free techniques. Recent prominent examples of the inappropriate incorporation of negative charge into these scaffolds have been shown to be deleterious to the promising pharmacological profile of an olefin-stapled BimBH3 helical inhibitor of the pro-survival protein Bcl-2,<sup>13,75</sup> a finding which has reinforced the correlation between charge and cellular uptake in constrained helical scaffolds and has further reinvigorated the need for a set of general rules that govern the uptake of anionic peptides that adopt other three-dimensional conformations.

## **2.2 – Investigating the cellular phenotypes of cyclic and bicyclic Grb2-SH2 domain inhibitors**

**BC1** and its analogs are inhibitors of the SH2 domain of Grb2, an intracellular adaptor protein that potentiates cell growth and proliferation via the Ras/Raf/MAPK pathway. Grb2 is over-expressed in human epithelial breast cancer cells, and the MDA-MB-453 breast cancer line is a model system for probing the effects of Grb2-SH2 inhibitors *in vitro*.<sup>92</sup> The anti-proliferative effects of **BC1** and its monocyclic precursors, as well as its methyl ester, were tested over the course of a 4-day experiment, and only very weak anti-proliferative

effects were observed at concentrations between 10 and 30  $\mu\text{M}$ . The methyl ester of **BC1** was synthesized (BC1-OMe) and tested in this assay in case the negative charge of **BC1** limited the cell penetration. However, the methyl ester exhibited similarly weak anti-proliferative activity. Competitive co-immunoprecipitation experiments in whole MDA-MB-453 cells aimed at determining the anti-Grb2 activity of our bicyclic peptide in living cells revealed

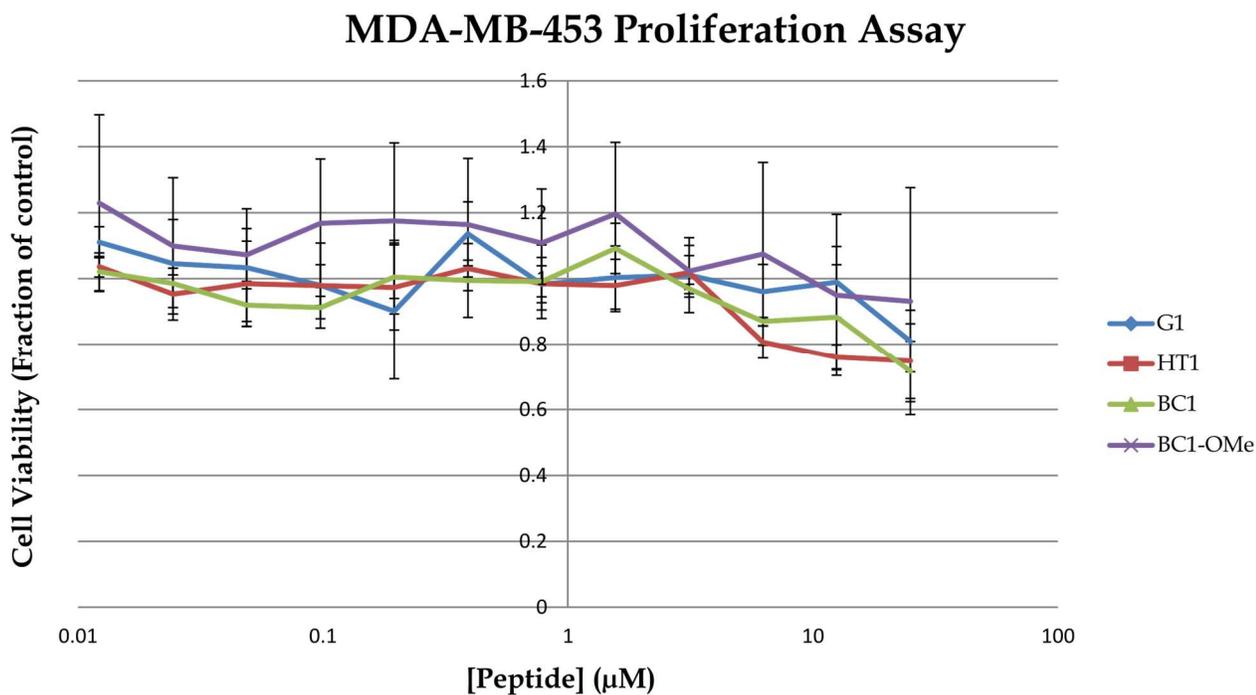


Figure 2.1. Analysis of the anti-proliferative effects of cyclic and bicyclic Grb2-SH2 domain inhibitors. only very weak reduction in Grb2-associated phosphor-EGFR (figure 2.2).

### 2.3 – Characterizing the cellular uptake of peptide analogs of G1HT1 and BC1

This section will describe the work that has been done to understand the cell-penetrating capacities of our designed peptide inhibitors of the Grb2-SH2

domain. We used confocal fluorescence microscopy to investigate the cellular uptake of our cyclic and bicyclic scaffolds using MDA-MB-453 cells. Variants of **HT1** and **BC1** labeled with BODIPY-FL were prepared (figure 2.3) with the dye conjugated to Glu1 to eliminate the negative charge (**HT1-Neut** and **BC1-Neut**), or to a Cys residue substituted for Met8 to maintain a net charge of -1 (**BC1-Neg**). To further probe the role of conformational constraint in promoting cell penetration, we also prepared analogous dye-labeled linear analogs, as well as a dye-labeled bicyclic scaffold with a different lactam cross-link (**BC3-Neut**). Confocal microscopy revealed that both **BC1-Neg** and **BC1-Neut** are internalized by MDA-MB-453 cells (figure 2.4) Cellular uptake was quantified using two independent measurements: ImageJ analysis of confocal microscopy images, and direct fluorometry of cell lysates following incubation and washing (figure 2.6). Both techniques showed similar trends. The neutral **BC1-Neut** was internalized as effectively as a BODIPY-labeled Tat peptide (**Tat-Bdy**), and the anionic **BC1-Neg** was internalized almost as effectively. Interestingly, **BC3-Neut** was internalized to a similar extent, implying that multiple bicyclic structures can be efficiently taken up by cells.

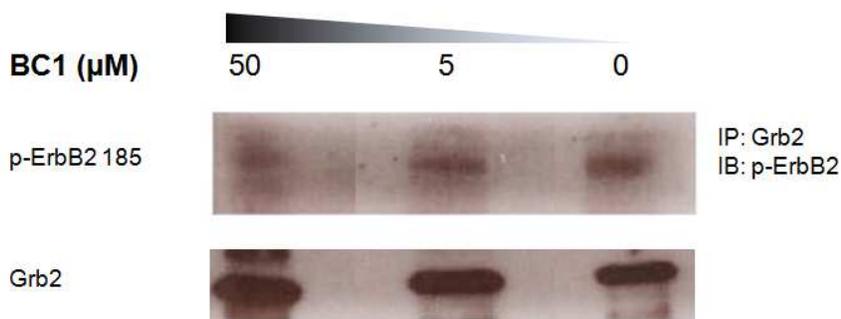


Figure 2.2. Results of competitive co-immunoprecipitation experiment in whole MDA-MB-453 cells. Cultured cells were incubated with various concentrations of **BC1** (5 or 50  $\mu\text{M}$ ) or 0.1% DMSO vehicle (0  $\mu\text{M}$  **BC1**). After lysis, Grb2 was immunoprecipitated and associated phosphor-EGFR was detected (top). A western blot to detect total Grb2 content was performed as a control (bottom).

to directly investigate the role of conformational restriction on cellular uptake, we repeated these experiments with the less rigid, monocyclic peptide **HT1-Neut** and found that it is taken up by cells to a similar extent as **BC1-Neut**. Linear peptides labeled at Glu1 (**Lin-Neut**) or at the N-terminus (**Lin-Neg**) were not effectively internalized – while some fluorescence was associated with cell lysates at high concentrations, confocal microscopy revealed only dim fluorescence localized to the exterior of the cell (figure 2.4). Thus, for these scaffolds we observed a general correlation between structural rigidity and cell penetration. Similar results were observed for fluorescein-labeled peptides as well as peptides with the dye attached at a third location (figure 2.5). We therefore conclude that the internalization observed is an intrinsic property of the constrained scaffolds.

Some previous work has correlated degree of structure to cell penetration for other classes of constrained peptides. However, the precise mechanisms by which constrained peptides are taken up by the cell are still poorly understood.<sup>13,74,79</sup> We sought to begin to characterize the mechanism of uptake for our cyclic and bicyclic scaffolds. The patterns of fluorescence associated with **BC1-Neut** and **BC1-Neg** contain a few bright punctate signals as well as more diffuse fluorescence that permeates the cytoplasm, with less fluorescence in the nucleus. **HT1-Neut** and **BC3-Neut** have more of their fluorescence distributed throughout the cytoplasm, with fewer visible punctae. These results are consistent with endocytotic entry for these peptides, with some extent of endosomal escape.<sup>93</sup> At 4° C, overall fluorescence intensity was dramatically reduced for all internalized cyclic and bicyclic scaffolds, but much less reduced for **Tat-Bdy**, as expected due to the multiple modes by which Tat is internalized.<sup>71,82,83,81</sup> The role of clathrin-mediated endocytosis in facilitating cellular uptake was studied using the small-molecule dynamin inhibitor Dynasore.<sup>94</sup> Pre-treatment with Dynasore resulted in increased formation of punctate fluorescence for **Tat-Bdy**, but resulted in increased cell-surface localization for internalized cyclic and bicyclic scaffolds (figure 2.4, quantified in figure 2.7). Taken together, these data implicate an energy-dependent, endocytotic pathway for uptake of these constrained peptides that differs substantially from the uptake mechanism of Tat. **BC1** and its methyl ester showed only mild anti-proliferative effects in MDA-MB-43 cells (figure 2.1), consistent with the mid-range IC<sub>50</sub> (0.35 μM) observed *in vitro*.<sup>44</sup> Future work will apply emerging label-free methods, pathway-specific assays, and assays for quantitating endosomal escape to further refine our understanding of these peptides' surprising cellular internalization.<sup>78,84,95</sup> Long-term, we believe this work

represents a promising new route towards intrinsically cell-penetrant scaffolds for inhibiting intracellular protein-protein interactions.

In order to quantitate the fluorescence intensity patterns observed by confocal microscopy, the resulting images were quantitated with respect to total, background-corrected cell fluorescence (figure 2.7) and also plot profile analysis (figures 2.8-2.10). Background subtraction of individual cells within images from each condition tested (37° C; 37° C, 80 μM dynasore; 4° C) resulted in a clear correlation between structural rigidification and cellular uptake at 37° C in the presence and absence of dynasore, as judged by corrected total cellular fluorescence (figure 2.7). For profile analysis, individual representative cells were selected from each micrograph of MDA-MB-453 cells incubated with peptide at 37° C in the presence and absence of dynasore. Using the ImageJ software package, linear rays were constructed that extended through the center of each cell, and gray value distributions were calculated along these lines. All bicyclic and cyclic peptides exhibit fluorescence intensity maxima in between the endpoints of the linear paths, with overall intensities well above background, indicating true cellular internalization. By contrast, the linear variants display little fluorescence in these regions (figure 2.8). These data further support the conclusion that unstructured, linear peptides may be able to bind the exterior of the cell, but are not internalized, while **BC1** and its cyclic and bicyclic variants are capable of being internalized. Treatment of cells with dynasore resulted in the

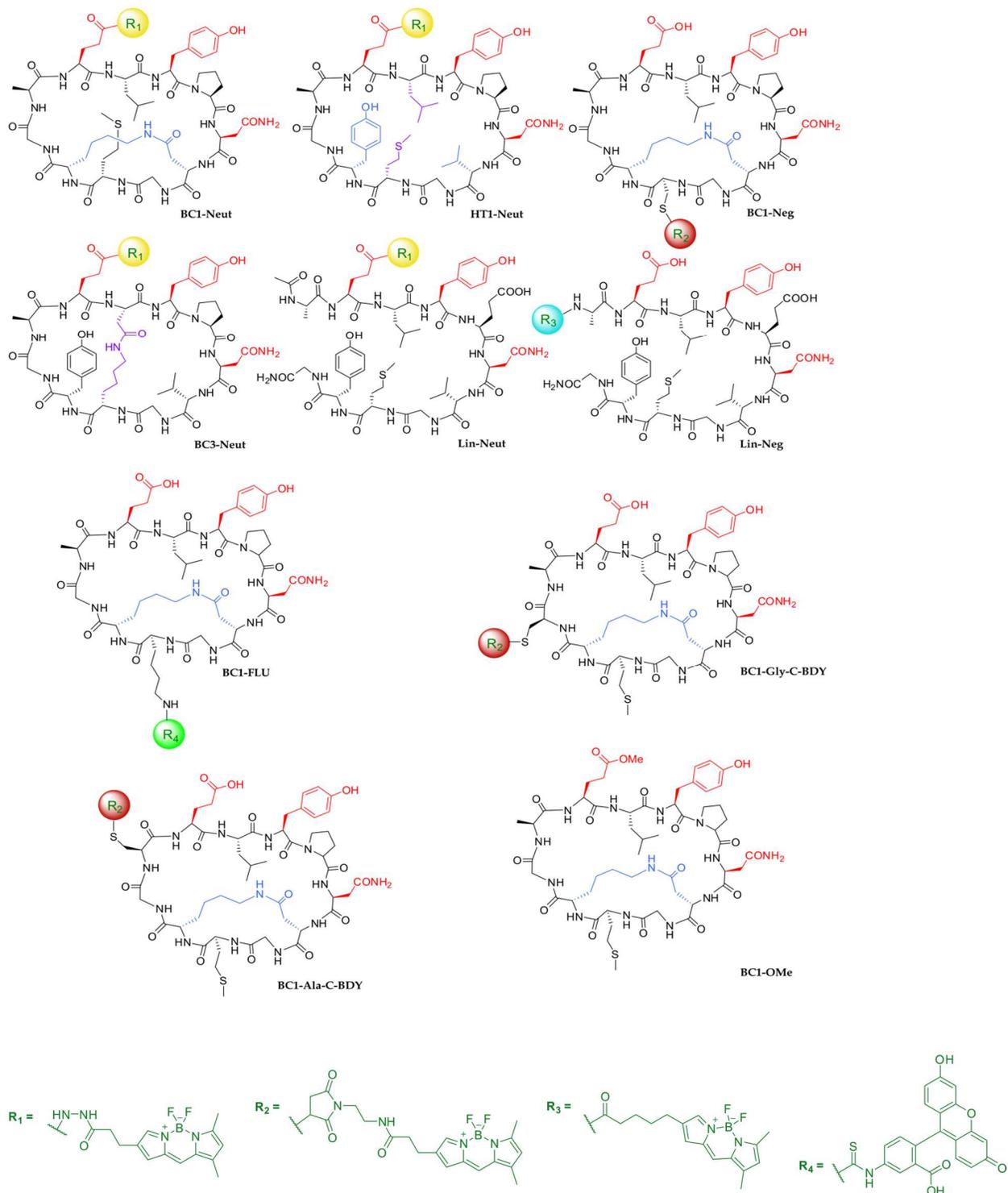


Figure 2.3 Fluorescently-labeled linear, cyclic and bicyclic peptides.

observation of increased punctate fluorescence for **Tat-Bdy**, and the observation of increased fluorescence localized to the cell surface for most of the bicyclic and cyclic scaffolds. Direct comparison of the fluorescence distribution profiles of cells treated with dye-labeled peptide in the presence and absence of dynasore helps to highlight these differences (figure 2.10). The FI profiles of **BC1-Neut** and **HT1-Neut** exhibit local maxima near the endpoints of the cell upon treatment with dynasore, while the highest fluorescence intensity values of cells not treated with dynasore occur in the interior of a linear path across the cell. This same trend is observed for **BC3-Neut**, while comparison of the FI profiles of **Tat-Bdy** reveals a largely similar distribution in the presence and absence of the dynamin inhibitor. Together, these data provide evidence for an energy-dependent, endocytic mechanism of internalization for the cyclic and bicyclic scaffolds. The results of the background-subtracted total cell fluorescence analysis (figure 2.7) reflect the trends observed from direct fluorometric analysis of cell lysates (figure 2.6), and along with the fluorescence distribution profiles below, collectively provide several lines of evidence that support the conclusion that the cyclic and bicyclic scaffolds are internalized by mammalian cells.

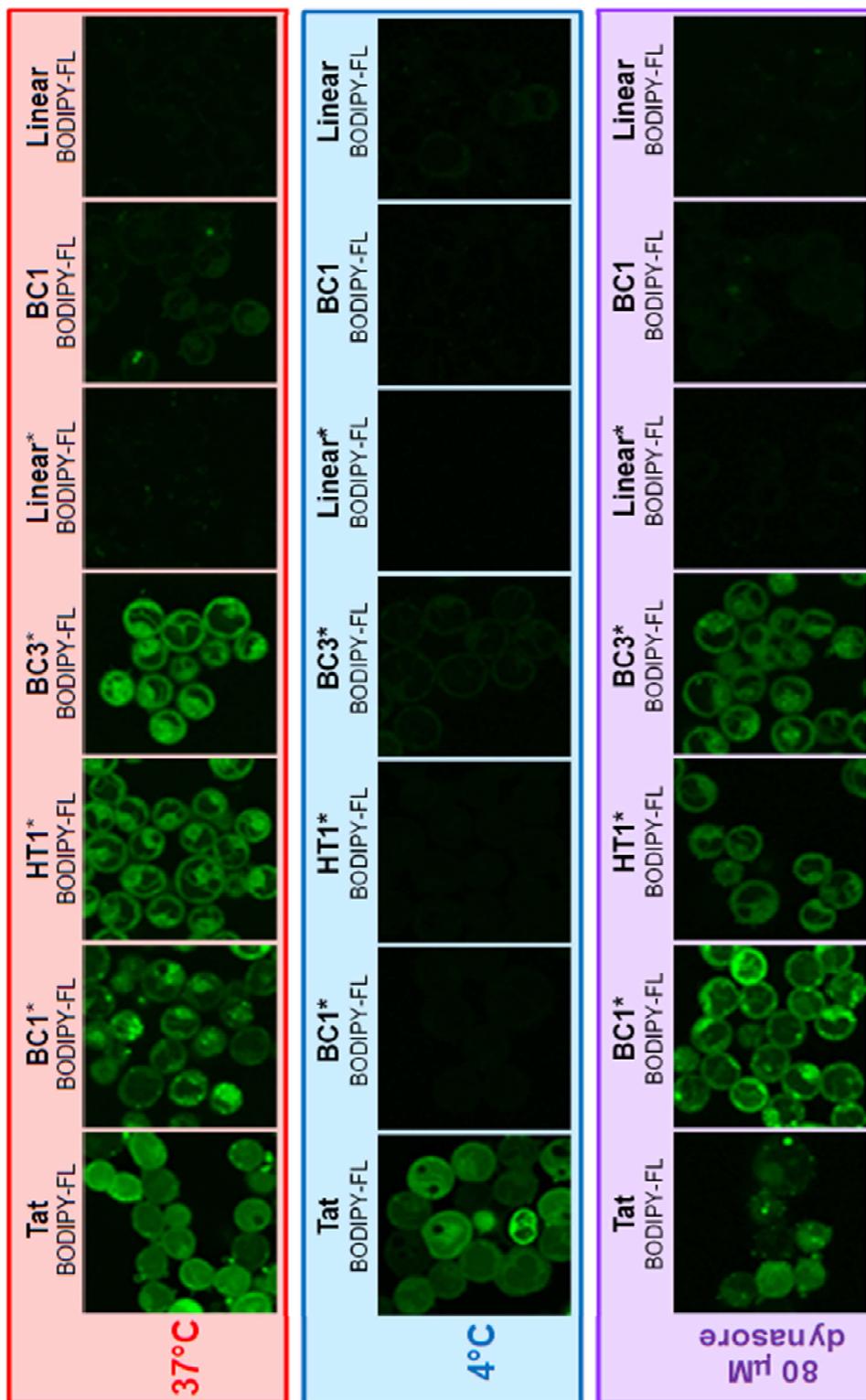


Figure 2.4 Summary of confocal microscopy-based cellular internalization results.

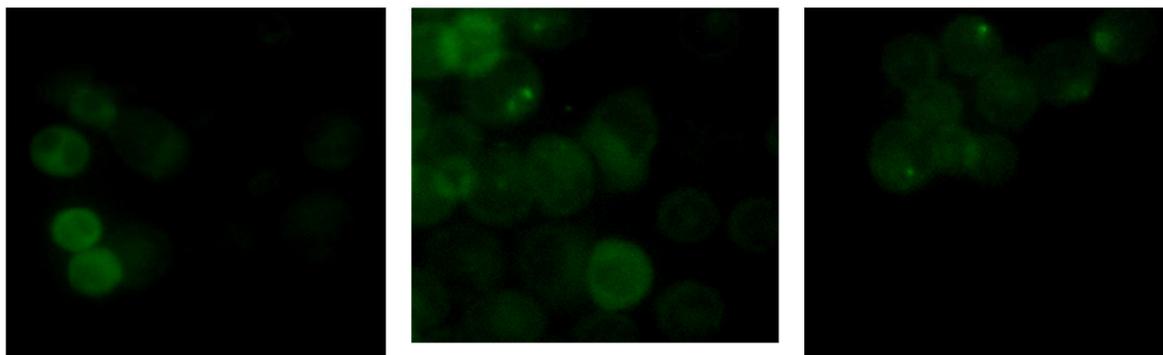
**BC1-Lys-FLU****BC1-Ala-C-Bdy****BC1-Gly-C-Bdy**

Figure 2.5 Summary of fluorescence microscopy-based cellular internalization results of alternatively-labeled bicyclic analogs.

Peptide	Calculated MW	Observed m/z	Species
BC1	1158.8	1157.3	[M-H]
BC1-OMe	1172.8	1173.2	[M+H]
HT1	1195.4	1194.7	[M-H]
BC1-Neg	1544.4	1524.1	[M-F]
BC1-Neut	1446.4	1427.6	[M-F]
HT1-Neut	1483.5	1463.8	[M-F]
Lin-Neut	1557.4	1555.9	[M-H]
Lin-Neg	1529.5	1527.8	[M-H]
BC3-Neut	1464.3	1444.7	[M-F]
BC1-FLU	1544.6	1546.2	[M+H]
BC1-Ala-C-Bdy	1604.6	1602.0	[M-H]
BC1-Gly-C-Bdy	1618.6	1616.5	[M-H]
TAT-Bdy	2078.1	2077.3	[M-H]

Table 2.1. Characterization of fluorescently-labeled linear, cyclic and bicyclic peptide analogs.

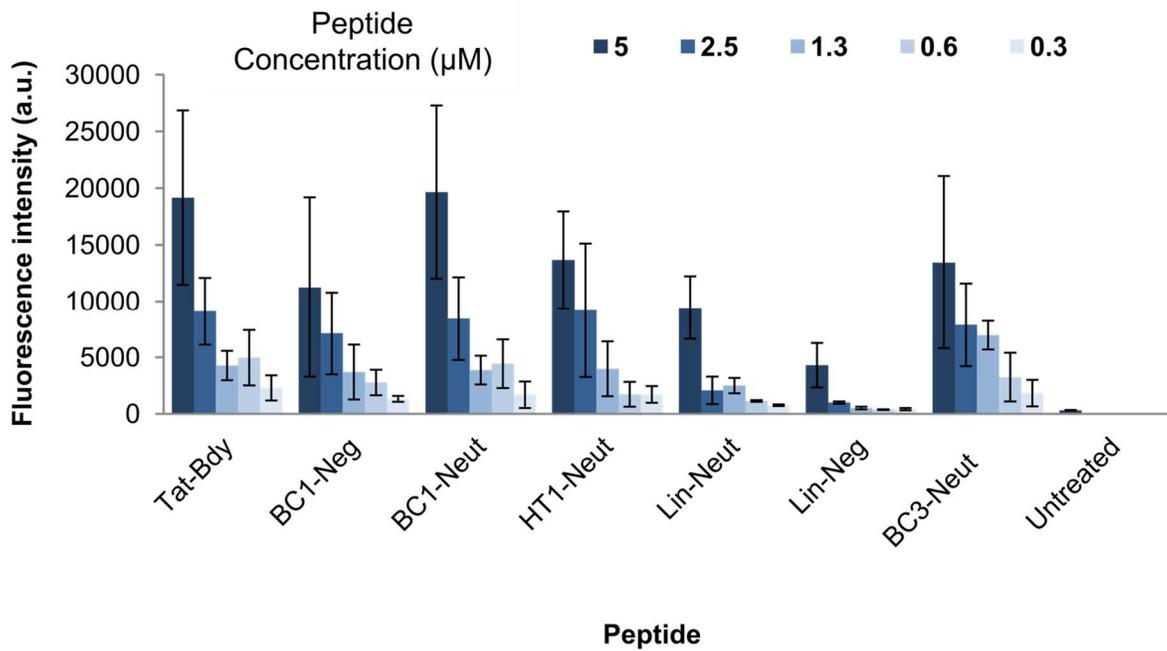


Figure 2.6. Fluorescence of MDA-MB-453 cell lysates following incubation of live cells with fluorescently-labeled linear, cyclic and bicyclic peptide variants.

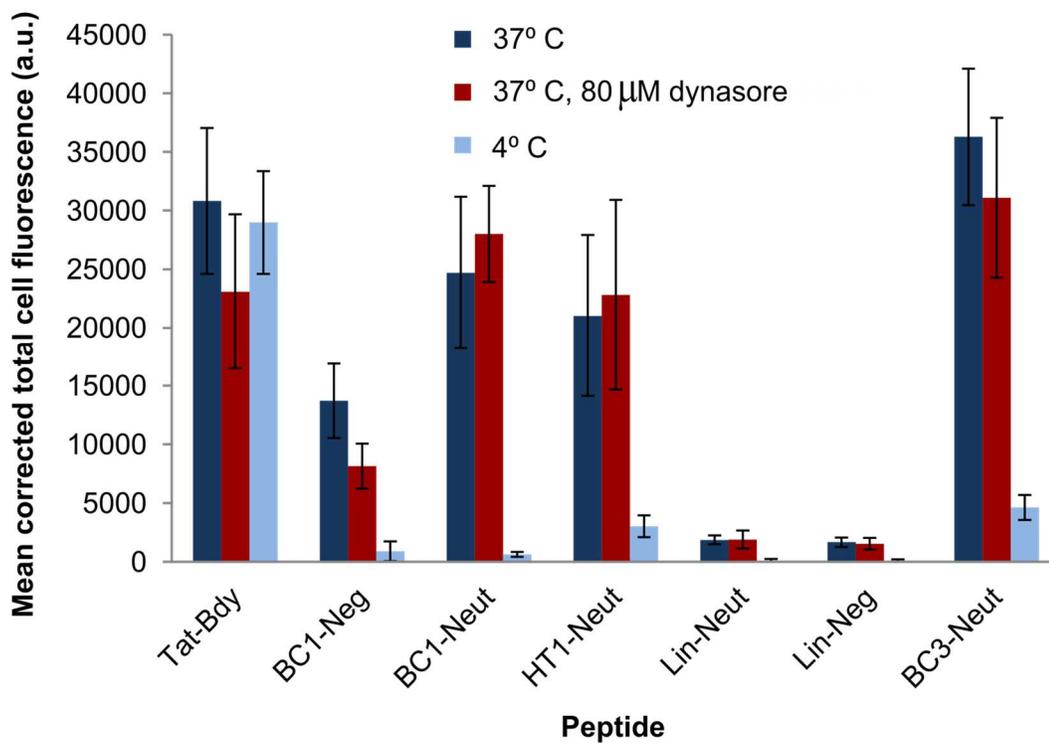


Figure 2.7. ImageJ quantitation of confocal microscopy images represented in figure 2.4.

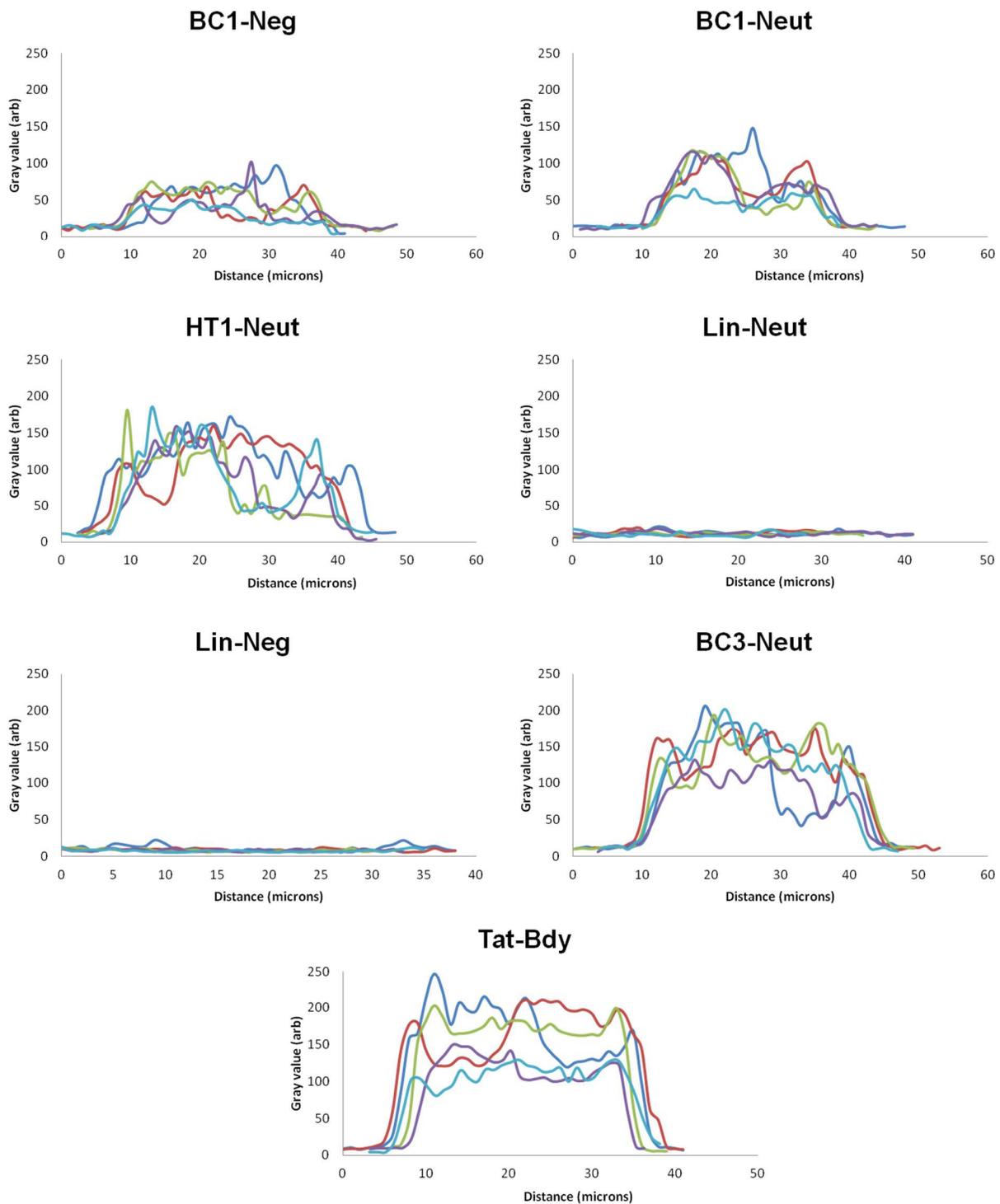


Figure 2.8. Plot profile analysis of confocal fluorescence micrographs of MDA-MB-453 cells incubated with fluorescently-labeled peptide variants in the *absence* of dynasore. The fluorescence profiles of 5 individually selected cells are overlaid for each peptide.

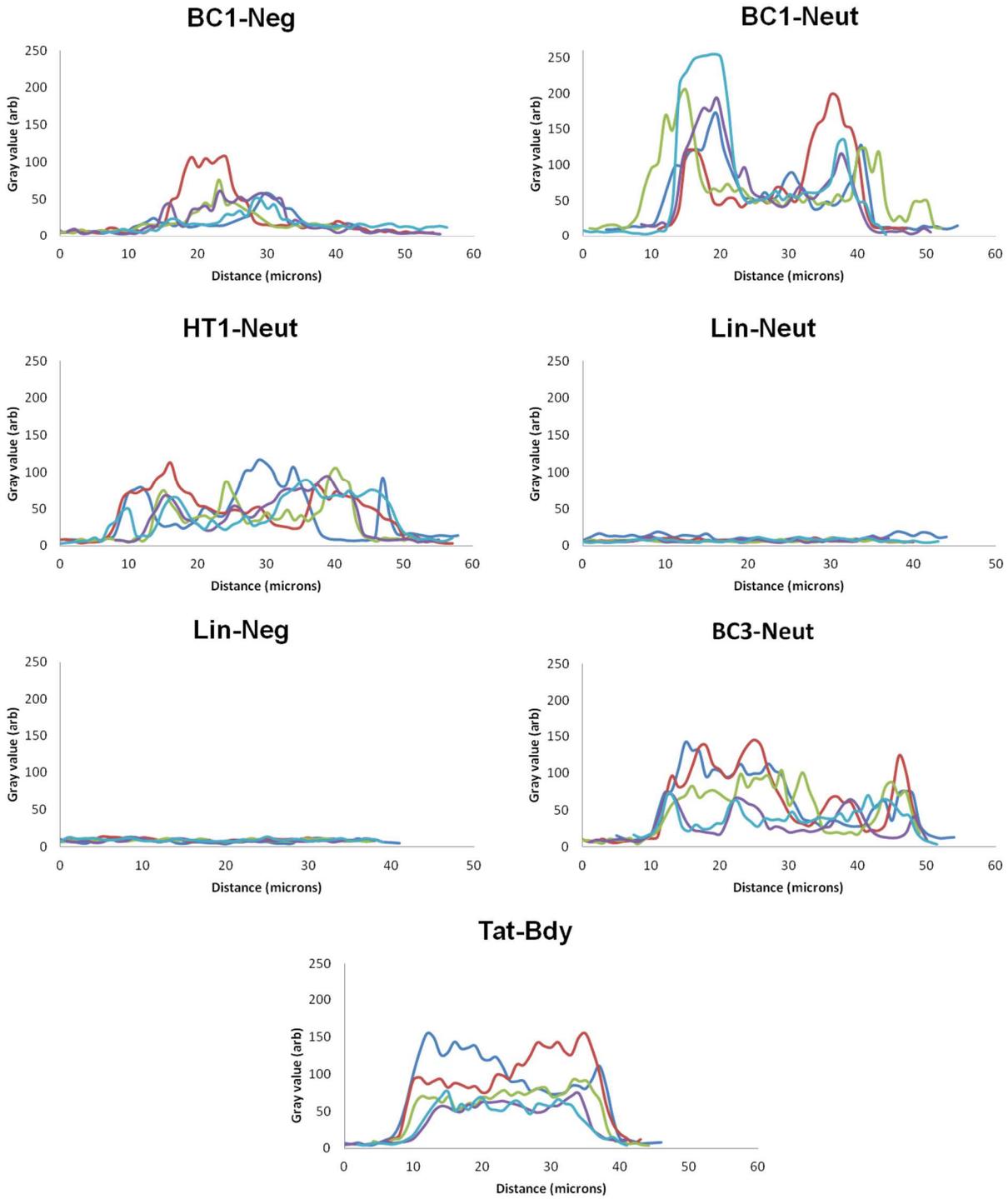


Figure 2.9. Plot profile analysis of confocal fluorescence micrographs of MDA-MB-453 cells incubated with fluorescently-labeled peptide variants in the *presence* of 80  $\mu$ M dynasore. The fluorescence profiles of 5 individually selected cells are overlaid for each peptide.

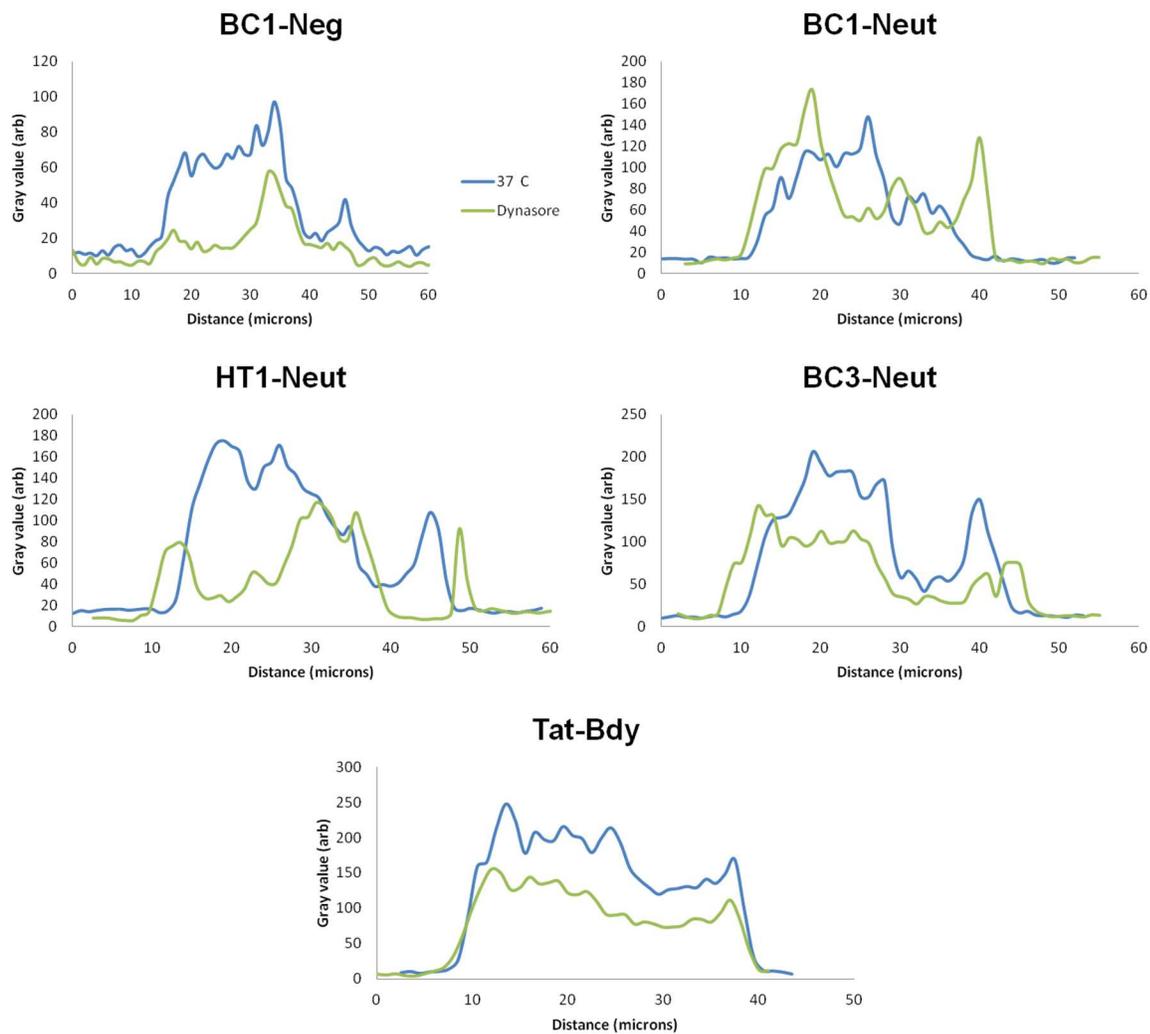


Figure 2.10. Comparison of representative FI distributions of individual cells from confocal fluorescence micrographs recorded from cells incubated with 10  $\mu$ M dye-labeled peptide in the absence (blue) and presence (green) of dynasore.

## Materials and Methods

### *Synthesis of fluorescently-labeled peptide variants*

**BC1** and **HT1**, as well as their cysteine- and lysine-containing analogs, were synthesized using Fmoc/t-bu solid phase peptide synthesis (SPPS) procedures as previously described in Chapter One, Materials and Methods. The methyl ester of BC1 was prepared by reacting BC1 with 10 equivalents of 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (HATU), 10 equivalents of 1-*hydroxy*-7-azabenzotriazole (HOAt) and 0.2 equivalents of 4-dimethylaminopyridine (DMAP) in methanol for 2 hours at room temperature. The resulting ester was then purified by RP-HPLC (see below). Linear variants of HT1 were synthesized using SPPS by attaching Fmoc-Gly-OH to the Rink amide, by first removing the 9-fluorenylmethoxycarbonyl (Fmoc) protecting group using a 20% solution of piperidine in dimethylformamide (DMF, 2 x 8 min). The resin was subsequently washed three times in DMF (3 x 5 min) prior to coupling the amino functionality of the resin to Fmoc-Gly-OH using 5 molar equivalents (relative to resin) of Fmoc-amino acid, 5 equivalents of (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP), 5 equivalents of hydroxybenzotriazole (HOBt) and 13 equivalents of diisopropylethylamine (DIEA) in DMF for 30 minutes at room temperature. The extent of coupling was monitored using the ninhydrin test, and couplings were repeated as necessary until no blue beads were observed. The resin was washed thoroughly with DMF (3 x 5 min) following all coupling and deprotection steps. Elongation of the linear chain continued until all amino acids

of the HT1-linear sequence had been coupled. The N-terminus of Lin-Neut was acetylated by treating the Fmoc-deprotected, resin-bound peptide with a solution of 1:3:3 acetic anhydride:DIEA:DMF for 30 minutes. The side-chain protected peptide was subsequently washed in DMF, dichloromethane (DCM) and methanol prior to global deprotection and cleavage from the Rink resin in trifluoroacetic acid (TFA) solution. A cocktail consisting of 94:2.5:2.5:1 TFA:ethanedithiol (EDT):water:triisopropylsilane (TIS) was used in order to effectively scavenge reactive species liberated during side-chain deprotection. The cleavage reaction proceeded for 2 hours at room temperature. The resin was subsequently filtered using glass wool and washed several times with TFA. The TFA cocktail was concentrated in vacuo and the crude was re-dissolved in 50:50 water:acetonitrile (ACN) prior to purification by RP-HPLC. Peptides were purified on a Varian ProStar HPLC by gradient purification using a water/ACN solvent system containing 0.1% TFA. Preparative scale Vydac C8 columns were used for purification, and all peptides were purified to >95% as judged by analytical scale re-injections of purified products.

#### *Analyzing the anti-proliferative effects of **BC1** and **BC1** variants*

MDA-MB-453 cells were seeded in a 96-well tissue culture-treated microtiter plate at  $10^4$  cells/ml and treated with the indicated concentrations of peptides in complete (FBS-supplemented Leibovitz L-15) media. Control cells were incubated with complete media only. Incubations lasted for 4 days, with one media change after 48 hours. Cells were subsequently washed twice in PBS and

treated with Cell Titer-glo reagent (Promega) to measure viability. Results are shown normalized to the viability of untreated control cells.

#### *Fluorometric Analysis of Cell Penetration*

MDA-MB-453 cells were seeded in a black, tissue culture-treated 96-well plate at  $10^4$  cells per well and washed twice with PBS to remove media. Cells were then incubated with PBS containing dye-labeled peptide at various concentrations for 1 hr at 37 deg C. Following the incubation, cells were washed twice with PBS and subsequently lysed in PBS containing 1% Triton X-100 for 30 min at room temperature. Fluorescence intensity was then measured at 535 nm using a Tecan Infinite F200 Pro plate reader.

#### *Confocal Fluorescence Microscopy*

MDA-MB-453 cells were seeded in a black, tissue culture-treated 24 well glass-bottom (#1.5 cover glass) titer plate at  $10^5$  cells per well and are grown to approximately 95% confluence in FBS-supplemented Leibovitz L15 media (ATCC). Cells were washed twice with PBS and were incubated with 10  $\mu$ M dye-labeled peptide in PBS for 1 hr at 37° C or 4° C. Following the incubation, cells were washed twice with and resuspended in PBS prior to imaging by confocal microscopy using a 63x water immersion lens (Leica Microsystems TCS SP2). For dynamin-inhibition experiments, cells were pre-incubated with 80  $\mu$ M dynasore (Sigma-Aldrich) for 30 min at 37° C followed by the addition of dye-labeled peptide to a final concentration of 10  $\mu$ M. Cells were subsequently

incubated with peptide in the presence of dynasore for 1 hr at 37<sup>0</sup> C prior to washing with PBS and imaging by confocal microscopy.

**Chapter Three: Analyzing the Structural Features of Cyclic  
Phosphotyrosine Mimetics that Promote Target Affinity and Cell  
Penetration by X-ray Crystallography and Nuclear Magnetic Resonance  
Spectroscopy**

**3.1 – Structural investigations of previous phosphotyrosine isosteres that target the Grb2-SH2 domain**

Previous structural campaigns aimed at understanding the interaction between the SH2 domain of Grb2 and phosphopeptides or molecules bearing conventional phosphotyrosine isosteres have revealed much information regarding the location of the phosphotyrosine-binding site of Grb2 as well as the conformational tendencies of high-affinity Grb2 ligands. Seminal efforts by Garcia-Echeverria and colleagues provided the first structural evidence that the SH2 domain of Grb2 binds phosphorylated tyrosine-containing peptides in the context of a  $\beta$ -turn.<sup>62</sup> This observation was the result of crystallization of the Grb2-SH2 domain in complex with a Shc-derived phosphorylated peptide, and has been highly influential in guiding the design of conformational-restricted phosphotyrosine mimetics. X-ray crystallographic investigations pioneered by Garcia-Echeverria and others have revealed that the SH2 domain of Grb2 interacts with anionic phosphotyrosine oxygen atoms via the formation of electrostatic contacts with the positively-charged guanidinium moieties of Arg67 and Arg86. This interaction is further strengthened by the formation of ion-dipole interactions between phosphotyrosine oxygen atoms and backbone amide protons of Ser88 and Ser90. In certain instances, the ionic contacts between a

pTyr residue and the cationic ammonium functionality of Lys109 have also been observed in solved X-ray crystal structures.

The X-ray crystal structure of the SH2 domain of Grb2 in complex with the Shc-derived peptide -pYVNV- has been particularly informative to subsequent medicinal chemistry campaigns, as this structure provided the first explicit evidence of a peripheral binding pocket adjacent to the phosphotyrosine binding site which is defined by Arg67, Arg86, Ser88 and Ser90. This cleft buries approximately 500 square angstroms of hydrophobic surface area and engages in hydrogen bonding with the side chain protons of the asparagine residue within the -pYVNV- sequence. These contacts orient the C $\alpha$  Asn proximal to the side chain of Trp121, which engages in a  $\pi$ :C-H interaction via the electron-rich aromatic indole moiety with the relatively electron-deficient C $\alpha$  proton of Asn. In addition to the identifying electronic characteristics of this binding pocket, this structure has been highly influential to Grb2-SH2 ligand design due to the discovery that the Asn residue that engages this binding site is situated at the i+2 position of a type-I  $\beta$  turn. This turn, nucleated by the valine residue immediately N-terminal to asparagine (position i+1), orients the phosphotyrosine and asparagine epitopes away from one another, such that each pharmacophore extends outward towards the appropriate binding pocket at the Grb2-SH2 domain surface. Subsequent efforts to generate small molecule and peptidomimetic ligands of the Grb2-SH2 domain have used this binding mode as a guide for inhibitor design. Efforts by Terrance Burke and colleagues at the National Institutes of Health to develop peptidic inhibitors of the Grb2-SH2 domain have followed a common paradigm that has resulted in the generation of highly potent

peptidomimetic compounds (figure 1.1 and table 1.1).<sup>52–54,56,57,96–98</sup> Inhibitors engineered using this approach contain three critical features: (i) a non-hydrolyzable phosphotyrosine isostere or other residue capable of mimicking the aromatic and anionic character of phosphotyrosine, (ii) a chemical moiety, often asparagine, bearing a terminal carboxamide (-CONH<sub>2</sub>) and a functional group poised to interact with the aryl system of Grb2 Trp121 via pi-pi or pi-C-H interactions, and (iii) a series of structural elements that pre-organize a  $\beta$ -turn that spatially segregates features (i) and (ii) and promotes the binding of these antigens into the corresponding sites on the Grb2 surface. Turn-stabilizing motifs include amino acids at site i+1 within the  $\beta$ -turn that contain structurally rigidified side chains, such as 1-amino-1-cyclohexane carboxylic acid (Ach), that serve to restrict the  $\phi$  and  $\psi$  torsional angles accessible to the neighboring residue.<sup>50,92</sup> Additional structural features that have been employed to stabilize the presentation of this  $\beta$ -turn scaffold include covalent linkers between the N-terminal phosphotyrosine mimetic and the C-terminal Trp121-engaging residue. Such linkers include all-carbon alkyl and alkenyl tethers generated by ring-closing olefin metathesis. Highly potent Grb2-SH2 inhibitors (IC<sub>50</sub>'s < 10 nM in ELISA experiments) contain 3-carbon linkers between the N- and C-termini of this peptide framework, with a *trans*-olefin located at positions 1 and 2 within the carbon chain. This structural element appears to preferentially stabilize the  $\beta$ -turn appropriate for the presentation of the pTyr-like and Trp121-engaging pharmacophores at the interface with the Grb2-SH2 domain, as subsequent x-ray crystallography investigations by Burke and colleagues of the interaction between the Grb2-SH2 domain and an inhibitor engineered using this paradigm revealed that the two pre-organized epitopes are indeed inserted within the respective binding pockets, while the covalent tether appears to play a structural

role in stabilizing the turn motif, as there is minimal interaction between the alkyl linker and the protein surface (figure 1.1).<sup>57</sup>

Previous investigations have been conducted by Peter Roller and colleagues to characterize the binding modes of the Grb2-SH2 domain with **G1**-derived peptides containing extensive modifications of the ELYDN motif.<sup>47,50,51,58</sup> Much of the chemical derivitization that has been performed on **G1** has been with the aim of enhancing the anionic character of the pTyr-like surface so as to improve the Coulombic interaction with the Ar67 and Arg86 residues of the phosphotyrosine-binding site. An example of this strategy includes the incorporation of a doubly-charged malonic acid functionality at the glutamate position within the ELY motif, as well as the introduction of additional negative charge at the benzylic position of the tyrosine residue. These modifications have furnished peptides with sub-micromolar affinities for the Grb2-SH2 domain. Roller's *in silico* analyses have been carried out by constructing energy-minimized representations of the peptides of interest and subsequently performing molecular dynamics docking simulations to determine the preferred location and orientation of these peptides on the protein surface. These simulations were executed in an unbiased fashion without employing *a priori* knowledge of the phosphotyrosine-binding site and peripheral aromatic binding pocket on the Grb2-SH2 domain surface. The results of these investigations suggest that the anionic functionality installed at the glutamate and tyrosine sites within the ELY sequence do indeed engage the Arg67 and Arg86 guanidinium moieties via ionic interactions, and furthermore that the asparagine residue inserts into the peripheral binding site delineated by the indole system of Trp121. That these structural models of derivatized **G1**

analogs in complex with the Grb2-SH2 domain are in line with reported crystal structures of phosphotyrosine isosteres bound within the pTyr-binding site of Grb2 provided further evidence that the ELY motif indeed constitutes a discontinuous phosphotyrosine mimetic and that proper stabilization of this unmodified epitope can be an effective strategy to engineer Grb2 inhibitors without endowing this compound with further negative charge, a well-documented liability to the generation of cell-penetrant peptides and small molecules.

In light of these previous structural investigations, we sought to characterize the structures of **BC1** and **G1HT1**, both in complex with the Grb2-SH2 domain by X-ray crystallography and alone in solution by employing multidimensional nuclear magnetic resonance spectroscopy techniques. Specifically, we aimed to answer several questions regarding the nature of our cyclic and bicyclic peptide inhibitors: What are the binding modes of these peptides in complex with the SH2 domain of Grb2? What are the atomic-level details that dictate phosphotyrosine mimicry in the context of the ELY motif? By applying a covalent tether to **G1HT1** in generating **BC1**, have we indeed stabilized the bioactive conformation of the molecule? And, in light of the interesting cell-penetrating capacity of **BC1**, what, if any, are the structural features that promote the inferred lipid binding and cell internalization?

The present chapter of this dissertation will document the experiments that have been performed in order to answer these questions. The work discussed herein would not have been possible without the help of collaborating professors at

Tufts Medical Center, Professors Andrew Bohm and James Baleja, who provided assistance and training in the fields of X-ray crystallography and biomolecular NMR, respectively. This chapter will conclude with a discussion of molecular dynamics simulations that were performed to analyze the behavior of **BC1** in the presence of explicit aqueous solvent in order to elucidate structural or electronic features that may promote cell penetration. This work was a collaborative effort with Professor Yu-Shan Lin and post-doctoral fellow Hongtao Yu of the Tufts University Department of Chemistry.

### **3.2 – Crystallization of the Grb2-SH2 domain in complex with cyclic and bicyclic phosphotyrosine mimics**

In order to solve the X-ray crystal structures of the Grb2-SH2 domain in complex with our cyclic and bicyclic peptide ligands, we executed two different strategies in parallel so as to maximize the probability of successful crystal growth. The first was to screen for conditions that favor the formation of crystals of the Grb2-SH2 domain:peptide complex, using previous reports in the literature as a guide for condition selection. The second strategy was to screen for conditions suitable for the crystallization of the apo Grb2-SH2 domain, isolate these crystals and subsequently soak in a concentrated solution of the cyclic or bicyclic inhibitor. In both cases, precedents established in previous Grb2-SH2 domain crystallography investigations were used as starting points for condition screens. In addition, commercial libraries of random crystallization conditions (MCSG

screening suite, Microlytic) were also screened for conditions suitable for the crystallization of both the Grb2-SH2:peptide complex and the apo Grb2-SH2 domain.

### *3.2.1 – Screening crystallization conditions based on literary precedent*

In order to find conditions optimal for the crystallization of the Grb2-SH2 domain in complex with **BC1**, we began by searching the literature for reported conditions that have been used to achieve the successful crystallization of Grb2-SH2 in complex with previous peptidic inhibitors. Martin and colleagues at the University of Texas at Austin have reported crystallization of the Grb2-SH2 domain in complex with the tripeptide AcpYVNV via vapor diffusion by mixing equal volumes of peptide:inhibitor solution at 20 mg/ml in Grb2-SH2 containing a 1.5:1 molar ratio of protein:peptide with a mother liquor comprised of 100 mM sodium cacodylate, 100 mM Ca(OAc)<sub>2</sub>, pH = 6.0.<sup>99</sup> In the same investigation, Martin also reports using vapor diffusion to successfully crystallize the Grb2-SH2 apo protein using a mother liquor containing 100 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer at pH 6.0 and 2.1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Schiering and colleagues have reported crystallization of the Grb2-SH2 domain in complex with the same tripeptide via room temperature hanging drop vapor diffusion by mixing equal volumes of protein peptide complex (containing 1.5:1 molar ratio of protein:peptide complex in 20 mM Tris-HCl, 300 mM NaCl, pH = 7.5) with a mother liquor comprised of 100 mM MES buffer at pH = 5.7, 500 mM NaCl and 11% PEG 3350.<sup>100</sup> Burke and co-workers at the NIH have also reported crystallization of the Grb2-SH2 domain under similar conditions in

complex with a macrocyclic, peptidomimetic inhibitor containing the phosphotyrosine isostere, malonyl tyrosine, and an asparagine residue at the i+2 position of a  $\beta$ -turn stabilized by an olefinic cross-link. In these investigations, crystallization was achieved by pre-treating a concentrated solution of the Grb2-SH2 domain (15 mg/ml in 20 mM Tris-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA) with an aqueous solution of inhibitor to yield a final molar ratio of 1.5:1 ligand:protein. This mixture was subsequently mixed with an equal volume of reservoir solution consisting of either 100 mM NaOAc pH = 5.7, 700 mM NaCl, 12% PEG 3350 or 100 mM NaOAc pH = 5.7, 1.4 M NaCl, 12% PEG 3350 and performing hanging drop vapor diffusion at room temperature.

Given the diversity of these previously reported crystallization conditions, we sought to maximize our search space by optimizing each of these conditions in parallel. In order to find ideal conditions for the crystallization of the Grb2-SH2 domain in complex with **BC1**, the aforementioned conditions were adjusted with respect to buffer pH, salt concentration, precipitant concentration, ambient temperature and the geometry of the vapor diffusion experiment (i.e., sitting drop versus hanging drop format). We reasoned that given the soluble nature of the protein of interest, and in light of the diversity of reported concentrations in the literature, the concentration of the Grb2-SH2 domain also required optimization. We therefore set up crystallization trays so as to systematically alter each of these parameters in order to draw conclusions about the ideal set of parameters that may favor crystallization. Hanging drop vapor diffusion experiments were performed in 24-well trays using 1, 2, or 5  $\mu$ L drops comprised of equal volumes of purified Grb2-SH2 domain and reservoir solution. Sitting drop vapor diffusion

experiments were executed in 96-well plate format using 1  $\mu$ L drops of 1:1 protein:reservoir mixtures. The results of these investigations are detailed in tables 3.1-3.6.

Initial condition optimization experiments failed to produce crystals in our hands. However, the systematic adjustment of each of the above parameters enabled us to draw meaningful conclusions regarding the effect of each of these elements on the behavior of the Grb2-SH2 domain, and served to inform subsequent crystal screens. As the data in tables 3.3-3.6 and images in figure 3.1 reveal, the Grb2-SH2 domain concentration appears to be the most significant factor in dictating protein precipitation, as no significant activity was observed in hanging drops or sitting drops produced from less concentrated protein samples (5 mg/ml). Additionally, at higher Grb2-SH2 domain concentrations (15 mg/ml), there appears to be a narrow range of PEG concentrations that favor crystallization, as fine microcrystals were observed between 18-22% PEG 3350 using Schiering's reported conditions (100 mM MES pH = 5.7, 500 mM NaCl).<sup>100</sup> However, reservoir compositions containing higher concentrations of PEG3350 induced the formation of large, amorphous precipitates, while lower concentrations of precipitant led to clear drops even after days to weeks at room temperature or 4° C.

Unlike the clear correlations observed by altering the concentration of PEG 3350 in mother liquor compositions, gleaned useful information from variations in salt concentration and buffer pH proved to be more challenging. Optimization of Martin's reported conditions<sup>99</sup> for the crystallization of apo Grb2-SH2 domain

revealed a minimal effect of  $(\text{NH}_4)_2\text{SO}_4$  concentration on Grb2-SH2 domain behavior, as the protein appeared to remain soluble even at high concentrations (2.0-2.5 M). Optimization of Burke's reported conditions<sup>57</sup> revealed a narrow pH range favorable to prevent precipitation of Grb2-SH2 in the presence of 1.4 M NaCl, as small amorphous precipitates were observed at pH conditions above 6.1 (6.1-6.9) while clear drops were observed between a pH of 5.1 and 5.9. See tables 3.1-3.6 for complete results.

### *3.2.2 – Screening commercial libraries for crystallization conditions*

In addition to searching for crystallization conditions based on previous reports, we also executed several factorial screens of random libraries of potential mother liquor compositions using commercially available screens. To this end, we began by screening conditions established by the Midwest Center for Structural Genomics (MCSG, a consortium funded by the National Institutes of Health) as favorable for the crystallization of a wide array of eukaryotic and prokaryotic proteins from a diversity of organisms. We performed sitting drop crystallization screens of concentrated apo Grb2-SH2 domain and Grb2-SH2 domain:peptide complex using the MCSG I, II and III crystal screens (Microlytic) for a total of 288 unique crystallization conditions. Each screen was performed in 96-well format at room temperature and 4° C using concentrated Grb2-SH2 domain (15 mg/ml) alone or protein that had been pre-treated with **BC1** to yield a final molar ratio of 1.5:1 peptide:protein.

Of these screens, the MCSG 1 crystallization suite yielded 5 conditions that induced successful crystallization of the apo Grb2-SH2 domain, and 2 of these conditions also enabled the crystallization of the Grb2-SH2:**BC1** complex. Detailed results of the secondary screens to optimize these conditions are provided below. Representative images of these crystals are shown in figure 3.1.

The two reservoir solutions that yielded crystals of the Grb2-SH2:inhibitor complex were MCSG 1 conditions A6 and C2. These conditions differed only in the identity of a cationic counterion; the A6 mother liquor contained 200 mM  $(\text{NH}_4)_2\text{SO}_4$ , 100 mM bis-tris HCl, pH = 5.5, 25% PEG 3350, while the C2 mother liquor contained 200 mM  $\text{Li}_2\text{SO}_4$ , 100 mM bis-tris HCl, pH = 5.5, 25% PEG 3350. These crystals exhibited similar morphology, as both were hexagonal prisms that ranged from 250  $\mu\text{m}$  (A6) to 500  $\mu\text{m}$  (C2) in the longitudinal axis and each approximately 100  $\mu\text{m}$  across the hexagonal face (figure 3.1). Further optimization of these crystallization conditions demonstrated a slight pH tolerance, as crystals of the Grb2-SH2:**BC1** complex were observed in sitting drop secondary screens between a pH of 5.3 to 5.7 under the A6 condition, but even small variations in salt and PEG concentrations resulted in either clear drops or amorphous precipitate. Optimization of the three conditions from the MCSG 1 screen that successfully induced crystallization of the apo Grb2-SH2 domain revealed a slight tolerance to higher PEG concentrations. In the case of MCSG 1 condition B6 (200 mM  $\text{Ca}(\text{OAc})_2$ , 100 mM MES, pH = 6, 20% PEG 8000) tetrahedral crystals were observed in the original crystal screen (approximately 80  $\mu\text{m}$  in the longest dimension). Further optimization of this condition revealed that at the same salt concentration and buffer pH,

incorporating a slightly higher PEG 8000 concentration of 23% resulted initially in visible precipitation. However, after 3-4 days at 4° C, a tetrahedral crystal of approximately 100 μm in the longest dimension grew from this precipitate. This was the only condition in which initial precipitation gradually gave way to protein crystallization, providing evidence that under these conditions the energy barrier for the phase transition from precipitation to crystal nucleation is traversable.

Of the crystal conditions discovered from the MCSG 1 screen, three of these crystals exhibited diffraction on a home source diffractometer (Xcalibur PX Ultra system, Oxford Diffraction, Concord, MA): two were of a Grb2-SH2:peptide complex (MCSG A6 and C2), while one was of the apo Grb2-SH2 domain (MCSG B6). The B6-derived crystal was soaked at 4° C in a concentrated solution of **BC1** (0.5 mM) dissolved in the B6 reservoir solution (200 mM Ca(OAc)<sub>2</sub>, 100 mM MES, pH = 6, 20% PEG 8000) plus 25% ethylene glycol, which was incorporated as a cryoprotectant to prevent the formation of ice on the crystal exterior following the soaking procedure (see Materials and Methods). A6- and C2-derived crystals were also immersed in mother liquor containing 25% ethylene glycol prior to storage of the crystals in liquid nitrogen. These three crystals were subsequently subject to X-ray diffraction at the National Synchrotron Light Source (Brookhaven National Laboratory, Brookhaven, NY) and 360° data sets were obtained for each (Appendix 1).

Table 3.1 Optimization of Literature Condition for Crystallization of Grb2-SH2:BC1 Complex

Original Condition: 100 mM MES, pH 5.7, 500 mM NaCl, 11% PEG 3350

[Grb2-SH2] (mg/ml)	[NaCl] (mM)	Buffer pH	[PEG 3350] (%) w/v)	Result
5	200	5.7	11	Clear droplet
5	300	5.7	11	Clear droplet
5	400	5.7	11	Clear droplet
5	500	5.7	11	Clear droplet
5	600	5.7	11	Clear droplet
5	700	5.7	11	Clear droplet
5	800	5.7	11	Light precipitation
5	500	5.1	11	Clear droplet
5	500	5.2	11	Clear droplet
5	500	5.3	11	Clear droplet
5	500	5.4	11	Clear droplet
5	500	5.6	11	Clear droplet
5	500	5.7	11	Clear droplet
5	500	5.8	11	Clear droplet
5	500	5.9	11	Clear droplet
5	500	5.7	10	Clear droplet
5	500	5.7	14	Clear droplet
5	500	5.7	18	Clear droplet
5	500	5.7	20	Clear droplet
5	500	5.7	22	Clear droplet
5	500	5.7	24	Clear droplet
5	500	5.7	26	Clear droplet
15	200	5.7	11	Clear droplet
15	300	5.7	11	Clear droplet
15	400	5.7	11	Clear droplet
15	500	5.7	11	Light precipitation
15	600	5.7	11	Light precipitation
15	700	5.7	11	Heavy precipitation
15	800	5.7	11	Heavy precipitation
15	500	5.1	11	Clear droplet
15	500	5.2	11	Clear droplet
15	500	5.3	11	Clear droplet
15	500	5.4	11	Clear droplet
15	500	5.6	11	Clear droplet
15	500	5.7	11	Clear droplet
15	500	5.8	11	Clear droplet
15	500	5.9	11	Clear droplet
15	500	5.7	10	Clear droplet
15	500	5.7	14	Clear droplet

15	500	5.7	18	Small microcrystals
15	500	5.7	20	Small microcrystals
15	500	5.7	22	Small microcrystals
				Large, amorphous
15	500	5.7	24	precipitate
15	500	5.7	26	Heavy precipitation

Table 3.2 Optimization of Literature Condition for Crystallization of Apo Grb2-SH2 domain

Original Condition: 100 mM MES, pH 6.0, 2.1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>

[Grb2-SH2] (mg/ml)	Buffer pH	[Ammonium sulfate] (M)	Result
5	5.5	2.1	Clear droplet
5	5.6	2.1	Clear droplet
5	5.7	2.1	Clear droplet
5	5.8	2.1	Clear droplet
5	5.9	2.1	Clear droplet
5	6.0	2.1	Clear droplet
5	6.1	2.1	Light precipitation
5	6.2	2.1	Light precipitation
5	6.3	2.1	Light precipitation
5	6.4	2.1	Light precipitation
5	6.5	2.1	Light precipitation
5	6.0	1.3	Clear droplet
5	6.0	1.5	Clear droplet
5	6.0	1.8	Clear droplet
5	6.0	2.1	Clear droplet
5	6.0	2.4	Clear droplet
15	5.5	2.1	Clear droplet
15	5.6	2.1	Clear droplet
15	5.7	2.1	Clear droplet
15	5.8	2.1	Clear droplet
15	5.9	2.1	Clear droplet
15	6.0	2.1	Clear droplet
15	6.1	2.1	Clear droplet
15	6.2	2.1	Clear droplet
15	6.3	2.1	Clear droplet
15	6.4	2.1	Light precipitation
15	6.5	2.1	Light precipitation
15	6.0	1.3	Heavy precipitation
15	6.0	1.5	Heavy precipitation

15	6.0	1.8	Clear droplet
15	6.0	2.1	Clear droplet
15	6.0	2.4	Clear droplet

Table 3.3 Optimization of Literature Condition for Crystallization of Grb2-SH2:BC1 Complex  
Original Condition: 100 mM NaOAc, pH 5.7, 1.4 M NaCl, 12% PEG 3350

[Grb2-SH2] (mg/ml)	[NaCl] (M)	Buffer pH	[PEG 3350] (% w/v)	Result
5	0.7	5.7	12	Clear droplet
5	0.9	5.7	12	Clear droplet
5	1.1	5.7	12	Clear droplet
5	1.2	5.7	12	Clear droplet
5	1.3	5.7	12	Clear droplet
5	1.4	5.7	12	Clear droplet
5	1.5	5.7	12	Light precipitation
5	1.7	5.7	12	Light precipitation
5	1.9	5.7	12	Light precipitation
5	2.1	5.7	12	Clear droplet
5	1.4	5.1	12	Clear droplet
5	1.4	5.3	12	Clear droplet
5	1.4	5.5	12	Clear droplet
5	1.4	5.7	12	Clear droplet
5	1.4	5.9	12	Clear droplet
5	1.4	6.1	12	Clear droplet
5	1.4	6.3	12	Clear droplet
5	1.4	6.5	12	Clear droplet
5	1.4	6.7	12	Clear droplet
5	1.4	6.9	12	Clear droplet
5	1.4	6.0	8	Clear droplet
5	1.4	6.0	12	Clear droplet
5	1.4	6.0	14	Clear droplet
5	1.4	6.0	18	Clear droplet
5	1.4	6.0	22	Light precipitation
5	1.4	6.0	26	Light precipitation
15	0.7	5.7	12	Clear droplet
15	0.9	5.7	12	Clear droplet
15	1.1	5.7	12	Clear droplet
15	1.2	5.7	12	Clear droplet
15	1.3	5.7	12	Clear droplet
15	1.4	5.7	12	Clear droplet

15	1.5	5.7	12	Mostly clear droplet, slight precipitation
15	1.7	5.7	12	Light precipitation
15	1.9	5.7	12	Clear droplet
15	2.1	5.7	12	Clear droplet
15	1.4	5.1	12	Clear droplet
15	1.4	5.3	12	Clear droplet
15	1.4	5.5	12	Clear droplet
15	1.4	5.7	12	Clear droplet
15	1.4	5.9	12	Clear droplet
15	1.4	6.1	12	Light precipitation
15	1.4	6.3	12	Heavy precipitation
15	1.4	6.5	12	Phase separation, precipitation
15	1.4	6.7	12	Heavy precipitation
15	1.4	6.9	12	Heavy precipitation
15	1.4	6.0	8	Clear droplet
15	1.4	6.0	12	Clear droplet
15	1.4	6.0	14	Clear droplet
15	1.4	6.0	18	Light precipitation
15	1.4	6.0	22	Light precipitation
15	1.4	6.0	26	Phase separation, precipitation

Table 3.4 Optimization of MCSG1 A6 Condition for Crystallization of Grb2-SH2:**BC1** Complex

Original Condition: 100 mM bis-Tris-HCl, pH 5.5, 200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 25% PEG 3350

[Grb2-SH2] (mg/ml)	[(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ] (mM)	Buffer pH	[PEG 3350] (%) w/v)	Result
15	50	5.5	25	Clear droplet
15	100	5.5	25	Clear droplet
15	150	5.5	25	Clear droplet
15	200	5.5	25	Hexagonal crystals
15	250	5.5	25	Slight precipitation
15	300	5.5	25	Amorphous brown precipitate
15	350	5.5	25	Heavy precipitation
15	400	5.5	25	Heavy precipitation
15	200	5.0	25	Clear droplet
15	200	5.1	25	Clear droplet
15	200	5.2	25	Slight precipitation
15	200	5.3	25	Very small Hexagonal crystals
15	200	5.4	25	Small (50 μm) hexagonal crystals
15	200	5.5	25	Large (200 μm) hexagonal crystals

15	200	5.6	25	Small hexagonal crystals
15	200	5.7	25	Small hexagonal crystals
15	200	5.8	25	Slight precipitation
15	200	5.9	25	Amorphous brown precipitate
15	200	6.0	25	Large, amorphous precipitate
15	200	6.0	18	Clear droplet
15	200	6.0	22	Clear droplet
15	200	6.0	24	Slight precipitation
15	200	6.0	26	Heavy precipitation
15	200	6.0	28	Large, amorphous precipitate

Table 3.5 Optimization of MCSG1 C2 Condition for Crystallization of Grb2-SH2:BC1 Complex

Original Condition: 100 mM bis-Tris-HCl, pH 5.5, 200 mM Li<sub>2</sub>SO<sub>4</sub>, 25% PEG 3350

[Grb2-SH2] (mg/ml)	[Li <sub>2</sub> SO <sub>4</sub> ] (mM)	Buffer pH	[PEG 3350] (%) w/v)	Result
15	50	5.5	25	Clear droplet
15	100	5.5	25	Clear droplet
15	150	5.5	25	Small microcrystals
15	200	5.5	25	Hexagonal crystals
15	250	5.5	25	Clear droplet
15	300	5.5	25	Slight precipitation
15	350	5.5	25	Heavy precipitation
15	400	5.5	25	Heavy precipitation
15	200	5.0	25	Phase separation
15	200	5.1	25	Phase separation
15	200	5.2	25	Brown precipitate
15	200	5.3	25	Hexagonal microcrystals
15	200	5.4	25	Small hexagonal microcrystals
15	200	5.5	25	100 μm hexagonal crystals
15	200	5.6	25	Small hexagonal crystals
15	200	5.7	25	Slight precipitation
15	200	5.8	25	Amorphous brown precipitate
15	200	5.9	25	Slight precipitation
15	200	6.0	25	Slight precipitation
15	200	6.0	18	Clear droplet
15	200	6.0	22	Clear droplet
15	200	6.0	24	Slight precipitation
15	200	6.0	26	Heavy precipitation
15	200	6.0	28	Heavy precipitation

Table 3.6 Optimization of MCSG1 B6 Condition for Crystallization of Apo Grb2-SH2 Domain

Original Condition: 100 mM MES-NaOH, pH 6.0, 200 mM Ca(OAc)<sub>2</sub>, 20% PEG 8000

[Grb2-SH2] (mg/ml)	[Ca(OAc) <sub>2</sub> ] (mM)	Buffer pH	[PEG 8000] (%) w/v	Result
15	50	6.0	25	Clear droplet
15	100	6.0	25	Clear droplet
15	150	6.0	25	Clear droplet
15	200	6.0	25	Small, tetrahedral microcrystals
15	250	6.0	25	Slight precipitation
15	300	6.0	25	Phase separation and precipitation
15	350	6.0	25	Amorphous brown precipitate
15	400	6.0	25	Heavy precipitation
15	200	5.5	25	Clear droplet
15	200	5.6	25	Clear droplet
15	200	5.7	25	Clear droplet
15	200	5.8	25	Slight precipitation
15	200	5.9	25	Small, tetrahedral crystals
15	200	6.0	25	Tetrahedral crystals
15	200	6.1	25	Small, tetrahedral microcrystals
15	200	6.2	25	Slight precipitation
15	200	6.3	25	Slight precipitation
15	200	6.4	25	Heavy precipitation
15	200	6.5	25	Heavy precipitation
15	200	6.0	17	Clear droplet
15	200	6.0	19	Clear droplet
15	200	6.0	21	Slight precipitation
15	200	6.0	23	Precipitation, tetrahedral crystal
15	200	6.0	15	Heavy precipitation

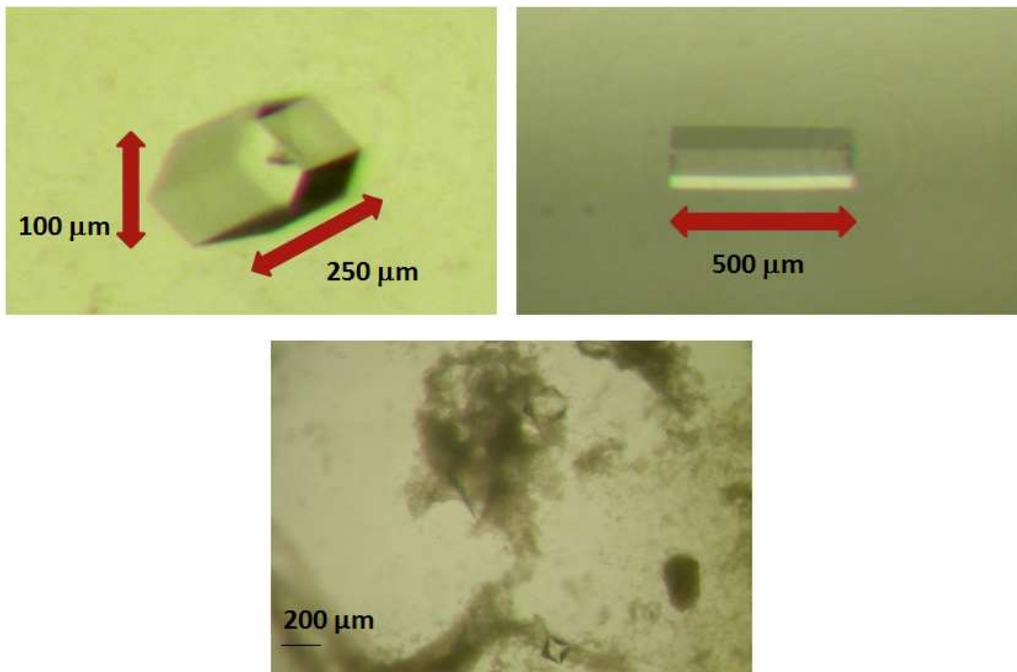


Figure 3.1. Crystals of the Grb2-SH2 domain in complex with peptide **BC1** from the MCSG1 A6 condition (top left) and C2 conditions (top right). Apo Grb2-SH2 domain crystals were grown from the MCSG1 B6 condition (bottom). These tetrahedral crystals were soaked in concentrated solutions of the peptide inhibitor before X-ray diffraction (see Materials and Methods, below).

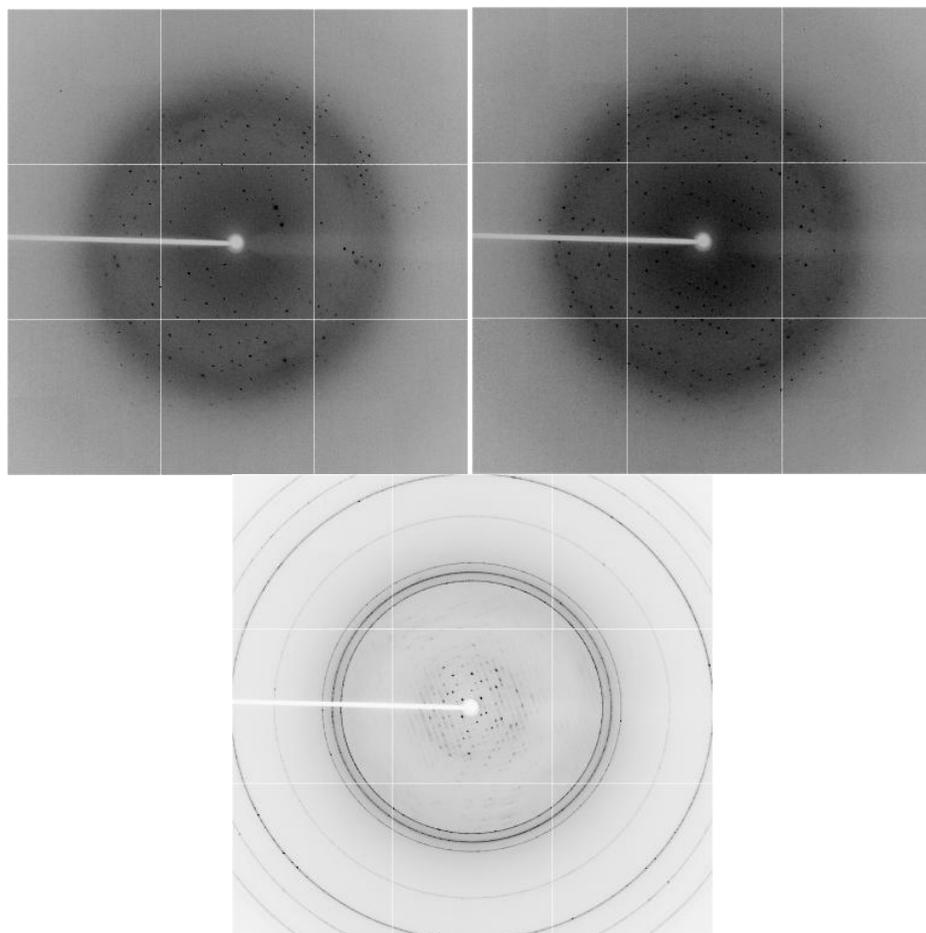


Figure 3.2. Representative diffraction patterns of crystals of the Grb2-SH2 domain in complex with peptide **BC1** from the MCSG1 A6 condition (top left), C2 conditions (top right) and B6 condition (bottom, following soaking in procedure). Maximum resolution was 2.8 angstroms, 3.2 angstroms and 3.4 angstroms for the A6, C2 and B6 crystals, respectively.

While the diffraction patterns of B6-derived crystals reveal significant ice formation on the crystal surface as evidenced by the presence of concentric ice diffraction rings dispersed about the beamstop, the A6- and C2-derived crystal exhibited no such ice diffraction, indicating successful cryoprotection of this crystal following immersion in liquid nitrogen. In all cases, the crystal structures of the Grb2-SH2 domain:peptide complexes were solved via molecular replacement methods using the apo Grb2-SH2 domain (PDB ID: 2AOA) as a search ensemble (see Materials and Methods). The resolutions of our refined structures are detailed in table 3.7, along with the space group and unit cell parameters. Final electron density maps and atomic coordinates are shown in figures 3.4 and 3.5.

In each case, we observed an interesting domain-swapping phenomenon that became apparent early in the phasing process. In solving the optimal unit cell parameters, we observed optimal spatial packing and water composition for a repeating unit of 4 protein chains per cell. However, upon comparison of early electron density composite maps to our atomic coordinates, we observed a continuous stretch of electron density originating from the C-terminal  $\alpha$ -helix of one SH2 domain monomer to a sheet comprised of three antiparallel  $\beta$ -strands of a neighboring monomer. Indeed, we also observed a twin region of electron density stretching from the C-terminal helix of the second monomer terminating at the  $\beta$ -sheet of the first monomer (figure 3.2). A visual inspection of the discrepancy between the electron density map and the atomic coordinates at Val122 within the Grb2 chain revealed the point of origin of this domain swapped dimeric Grb2-SH2 domain. This phenomenon has been reported previously,<sup>99,100</sup>

and it has been established that this dimerization is particularly favored by heating the protein:peptide complex to 50° C prior to setting up the vapor diffusion experiment. After making this observation and reflecting on the literary precedent more carefully, we recalled that our solution of protein:inhibitor complex was indeed heated briefly prior to executing the vapor diffusion experiment. This heating step may provide the energy necessary to overcome the kinetic barrier to domain swapping in order to form a more thermodynamically stable dimeric SH2 domain complex.

In each of our solved crystal structures, we observed incomplete saturation of the phosphotyrosine-binding site by peptide **BC1**. In the case of our C2-derived crystal, the residual electron density observed in our  $2F_o - F_c$  composite map indicated occupancy of the putative pTyr-binding site defined by Arg67, Arg86, Ser88, Ser90 and Lys109 by a sulfate counterion. This was also observed in the final composite map of the B6 crystal structure. However, the composite electron density map of our A6 crystal revealed residual electron density in each of the four pTyr-binding sites in the unit cell of the Grb2-SH2 domain structure. This excess electron density does not correspond to the atomic coordinates of the Grb2-SH2 domain, and indeed the coordinates of the Arg, Lys and Ser residues that delineate the pTyr-binding site are in close agreement with the electron density assigned to these sites. It was originally hypothesized that the residual electron density in the pTyr-binding site may be attributed to local deformations in the macroscopic crystal, or perhaps that this was an artifact of the observed ice diffraction (figure 3.2). However, this excess electron density appears in each of the pTyr-binding sites of the domain-swapped Grb2-SH2 domain complex, and in

all four pTyr-binding sites per unit cell. Moreover, the volume of this residual electron density is exceedingly large to be accounted for by a simple sulfate counterion (as was the case in our C2 crystal structure, figure 3.4), and it lacks the spherical distribution that would be expected for a symmetric, tetrahedral species.

Taken together, these observations provide evidence that this residual electron density can be attributed to the inhibitor, **BC1**. As shown in our most refined structure, it is reasonable that this residual electron density corresponds to the ELY region of the **BC1** surface. However, there is insufficient electron density to assign the atomic coordinates of the remainder of the peptide. This phenomenon may be the result of one or more molecular-level events within the Grb2-SH2:**BC1** crystal. It is possible that the peptide engages the pTyr-binding pocket via a rigidified ELY pharmacophore that is confined to a single spatial orientation upon establishing ion-dipole contacts and hydrogen bonds between the phenolic OH and glutamate carboxyl moiety of **BC1** with the cationic ammonium and guanidinium functionalities at the Grb2 surface, while the remainder of the peptide that does not directly bind the protein surface is flexible and adopts multiple conformations. This would account for the inability to detect electron density corresponding to the distal end of the peptide, and the conformational flexibility of **BC1** in solution is a property we sought to explicitly evaluate using multidimensional NMR techniques (*vide infra*). Another possible explanation for the lack of complete **BC1** electron density may also be incomplete occupancy of the pTyr-binding sites in the Grb2-SH2 crystal lattice. This inability to saturate the binding site may be due to the dimerization of the SH2 domain, as the observed

domain-swapping may partially preclude binding of **BC1** to the pTyr-binding surface on Grb2.<sup>99,100</sup> This domain-swapping mode of dimerization was only observed upon heating the protein:peptide sample prior to crystallization. It follows that this heating step was perhaps necessary to overcome a kinetic barrier to dimerization, which suggests that the folding of the domain-swapped Grb2 dimer is the slow step in the formation of the dimeric Grb2:**BC1** complex. If the kinetics of crystal nucleation are such that crystallization proceeds rapidly following the proper folding of the domain-swapped Grb2 dimer, it is reasonable that a substantial portion of this dimeric complex may crystallize without a bound **BC1** ligand in the pTyr-binding site. Finally, another potential reason for the insufficient electron density at this site may also be the that crystal diffraction was punctuated by ice diffraction at various areas within the ensuing diffractograms. The resulting ice rings impose an upper limit to the resolution of the calculated electron density map, as spots within the diffraction pattern of the A6 crystal that are near these ice rings are often not indexed so as to avoid capturing ice diffraction and introducing artifacts into the subsequent Fourier transform.

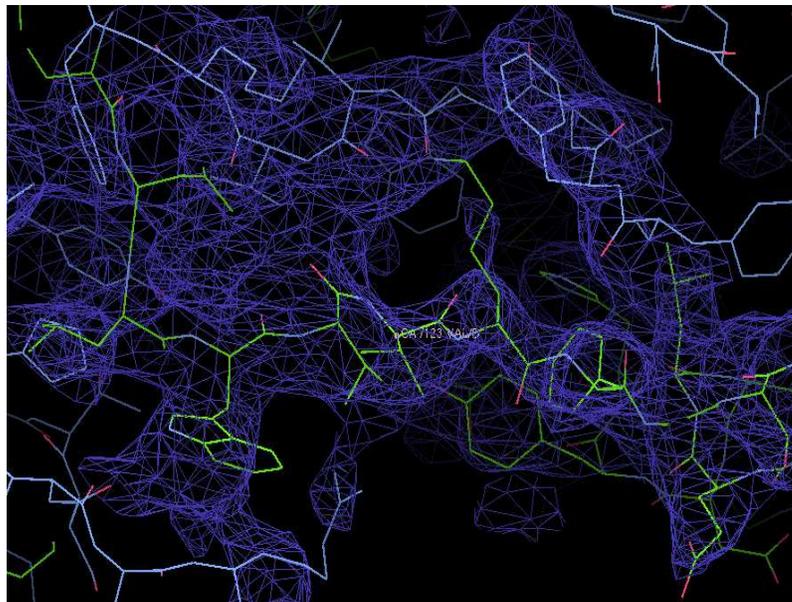
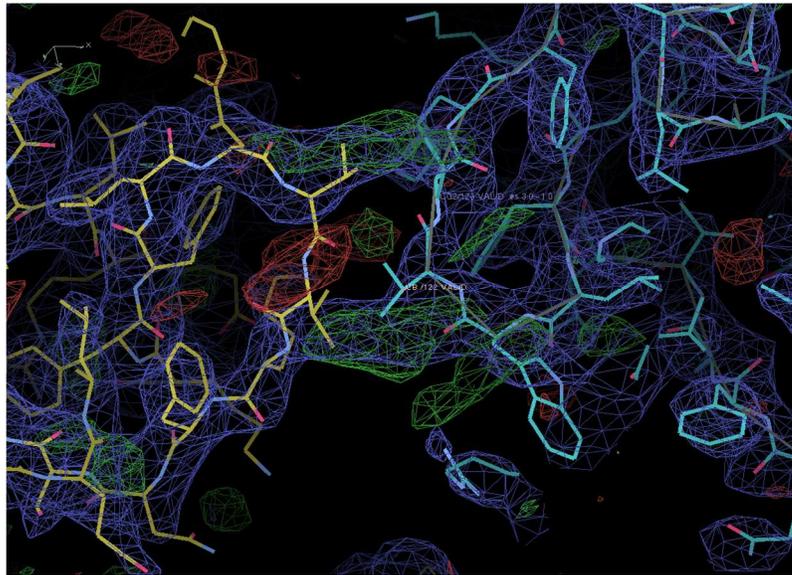


Figure 3.3. Domain-swapping in the formation of dimeric Grb2-SH2 domain complexes. Early phasing of electron density amplitudes to the molecular replacement search ensemble (PDB ID 1AOA, monomeric Grb2-SH2 domain) revealed a domain-swapped dimer. *Top*: Continuous electron density spans the gap between the  $\beta$ -sheets of two individual Grb2 monomers indicated by the atomic coordinates. *Bottom*: Manual construction of a domain-swapped search ensemble and phasing to this model restores alignment between electron density and atomic coordinates in the dimeric structure.

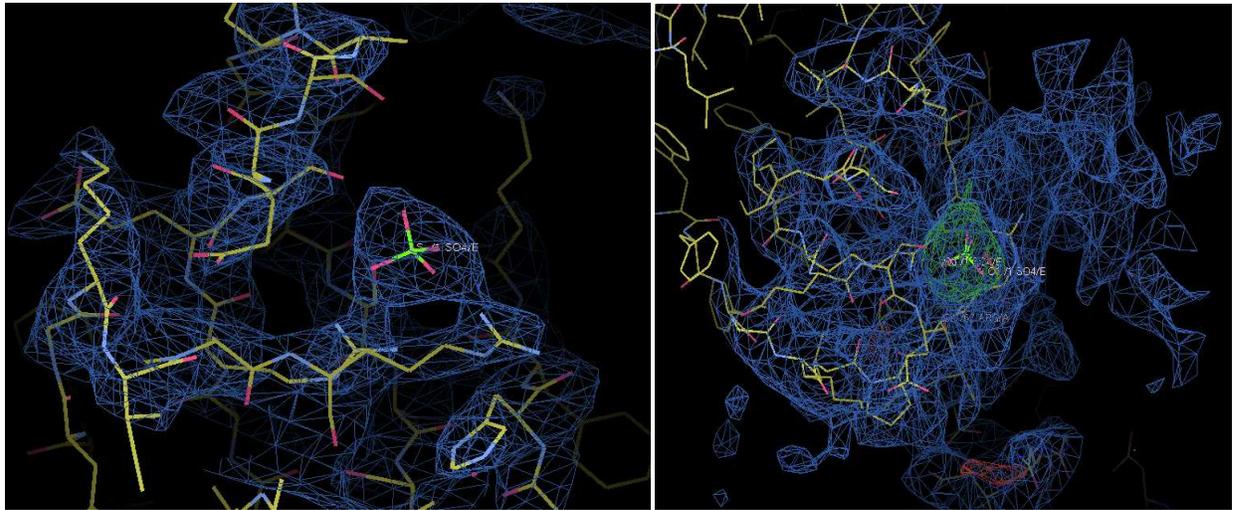


Figure 3.4. Phosphotyrosine-binding sites of Grb2-SH2 domain crystal structures determined from C2 (left) and A6 (right) crystals and electron density maps. Residual electron density in A6 crystal structure may correspond to incomplete occupancy **BC1**. Electron density in C2 structure maps well to a sulfate counterion.

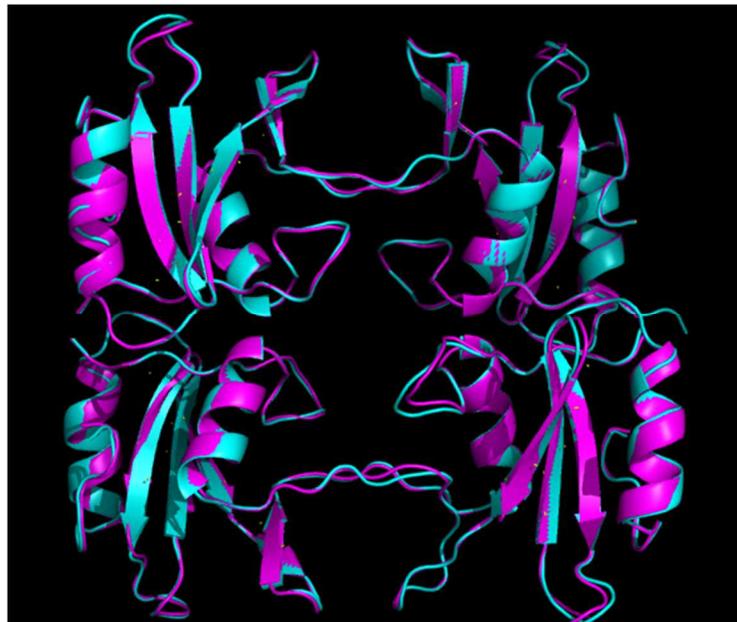


Figure 3.5. Overlay of cartoon representations of the A6 (teal) and C2 (purple) crystal structures. Both structures demonstrate Grb2 dimerization by domain-swapping.

## Materials and Methods

### *Expression and Purification of Grb2-SH2 Domain*

Recombinant Grb2-SH2 domain was expressed and purified as previously described (see section 1.4, Materials and Methods), with several notable exceptions. High efficiency BL21(DE3) (New England Biolabs) were transformed with pGEX vector DNA encoding residues 59-152 of Grb2 (the Grb2-SH2 domain) as a C-terminal glutathione-S-transferase (GST) fusion protein. Single colonies were picked and overnight solutions prepared as described previously. A total of 6 L of 2XYT containing 100 µg/ml ampicillin were inoculated with 60 mL saturated overnight culture and were grown to an OD<sub>600</sub> of 0.5 at 37° C under gentle agitation. Protein production was induced by the addition of IPTG to a final concentration of 400 µM, and cultures were subsequently incubated at 30° C for 3 hours. Cells were harvested by centrifugation at 3,000 x g and pellets were resuspended in 50 mM Trish-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA containing 10 mM β-mercaptoethanol and protease inhibitor cocktail (Sigma-Aldrich). Cells were lysed by rod sonication using four consecutive on/off cycles of 1 min each. Lysis was carried out on ice. Cell lysates were cleared by centrifugation at 9,000 x g, and clarified cell lysate was subsequently diluted 1:2 in 50 mM Trish-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA to reduce viscosity. The resulting solution was incubated with 10 mL aqueous, swollen glutathione-agarose matrix (Gold Biotechnology) for 2 hours at 4° C. Glutathione-agarose matrix was subsequently washed with 10 column volumes of column buffer. Fusion protein was then cleaved on-resin by treatment of the solid support with

column buffer containing 100 units of purified thrombin (Sigma-Aldrich) for 12-16 hours at 4° C. The filtrate was collected and the resin washed several times with thrombin digestion buffer. These fractions were pooled and passed over a *p*-aminobenzamidine column so as to remove thrombin from the mixture. Filtrate was then dialyzed into 50 mM Tris-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA and concentrated to 15 mg/mL as measured by nanodrop UV-Vis (MW = 11.19 kDa,  $\epsilon = 13,463 \text{ M}^{-1} \text{ cm}^{-1}$  at 280 nm).

#### *Hanging Drop Vapor Diffusion Experiments*

Hanging drop vapor diffusion experiments were carried out in 24-well Linbro crystallization plates (Hampton Research) according to established protocols.<sup>57</sup> Briefly, a concentrated aqueous solution of peptide inhibitor (**BC1** or **G1HT1**) was added to a purified Grb2-SH2 domain in 50 mM Tris-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA to yield a solution of protein:inhibitor complex with a final molar ratio of peptide:protein of 1:1.5 and a final concentration of Grb2-SH2 domain of 5 or 15 mg/ml. Equal volumes of this solution (or a solution of the apo Grb2-SH2 domain) and mother liquor were mixed and applied by pipetting in the center 22 mm x 22 mm siliconized glass cover slips (Hampton Research). Well reservoirs were filled with 1 mL mother liquor solution. High vacuum grease was applied along the outer rim of each well, and the glass cover slip was subsequently inverted and applied to the greased well exterior such that protein:peptide drops were suspended above the mother liquor solution. The tray was subsequently sealed and stored at room temperature or 4° C for 24 hours to

a week. Plates were monitored every 24-36 hours to check for favorable crystallization conditions.

#### *Sitting Drop Vapor Diffusion Experiments*

Sitting drop vapor diffusion experiments were carried out in 96-well and 24-4 well Intelliplate crystallization plates (Hampton Research) according to established protocols.<sup>99</sup> Briefly, a concentrated aqueous solution of peptide inhibitor (**BC1** or **G1HT1**) was added to a purified Grb2-SH2 domain in 50 mM Tris-HCl, pH = 7.5, 200 mM NaCl, 1 mM EDTA to yield a solution of protein:inhibitor complex with a final molar ratio of peptide:protein of 1:1.5 and a final concentration of Grb2-SH2 domain of 5 or 15 mg/ml. Equal volumes of this solution (or a solution of the apo Grb2-SH2 domain) and mother liquor were mixed and applied by pipetting to a raised platform within each well designed to hold 1-5  $\mu$ L sample. Well reservoirs were subsequently filled with 100  $\mu$ L mother liquor solution. The tray was then sealed and stored at room temperature or 4° C for 24 hours to a week. Plates were monitored every 24-36 hours to check for favorable crystallization conditions.

#### *Data collection, structure determination and refinement*

Final X-ray diffraction data was collected at the National Synchrotron Light Source at Brookhaven National Labs. Diffraction patterns were indexed and Fourier Transform analysis was executed using the iMosflm software package. Structures were solved using molecular replacement, and previously reported

SH2 domain crystal structures were used as search ensembles (PDB ID: 1A0A).<sup>57</sup> Phasing and molecular replacement were executed using the HKL-2000 software suite. Manual refinement of atomic coordinates was carried out using the WinCoot software tool. Two-stage simulated annealing and refinement was performed in iMosflm.

### 3.3 – Probing the structures of **BC1** and **G1HT1** by multidimensional NMR

The proposed binding mode for **HT1** and **BC1** involved a discontinuous epitope across Glu1 and Tyr3 that mimics phosphotyrosine (pTyr), and Asn5 that could bind a nearby selectivity pocket on Grb2-SH2.<sup>47,50,51</sup> Our prior work demonstrated that this epitope is highly sensitive to small alterations in macrocycle conformation,<sup>44</sup> but no structural details of the proposed conformation-dependent pTyr mimic were available. To obtain structural insight into this conformation-dependent epitope, we determined the NMR structure of **BC1** in aqueous solution. TOCSY, NOESY and <sup>1</sup>H-<sup>13</sup>C HSQC NMR spectra were obtained for 0.76 mM **BC1** in 9:1 H<sub>2</sub>O:*d*<sub>6</sub>-DMSO. The recorded spectra allowed complete resonance assignment. The H $\alpha$  and amide proton chemical shifts of most residues differ significantly from typical values for random-coil structures, indicating the peptide adopts a well-defined structure. 90 NOE-derived distance restraints were compiled. Many of the strongest NOEs were between protons of Glu1 and Tyr3, providing direct evidence that these residues form a contiguous surface. The distance restraints were used in multi-stage simulated annealing calculations to produce a structural ensemble of thirty lowest-energy structures..

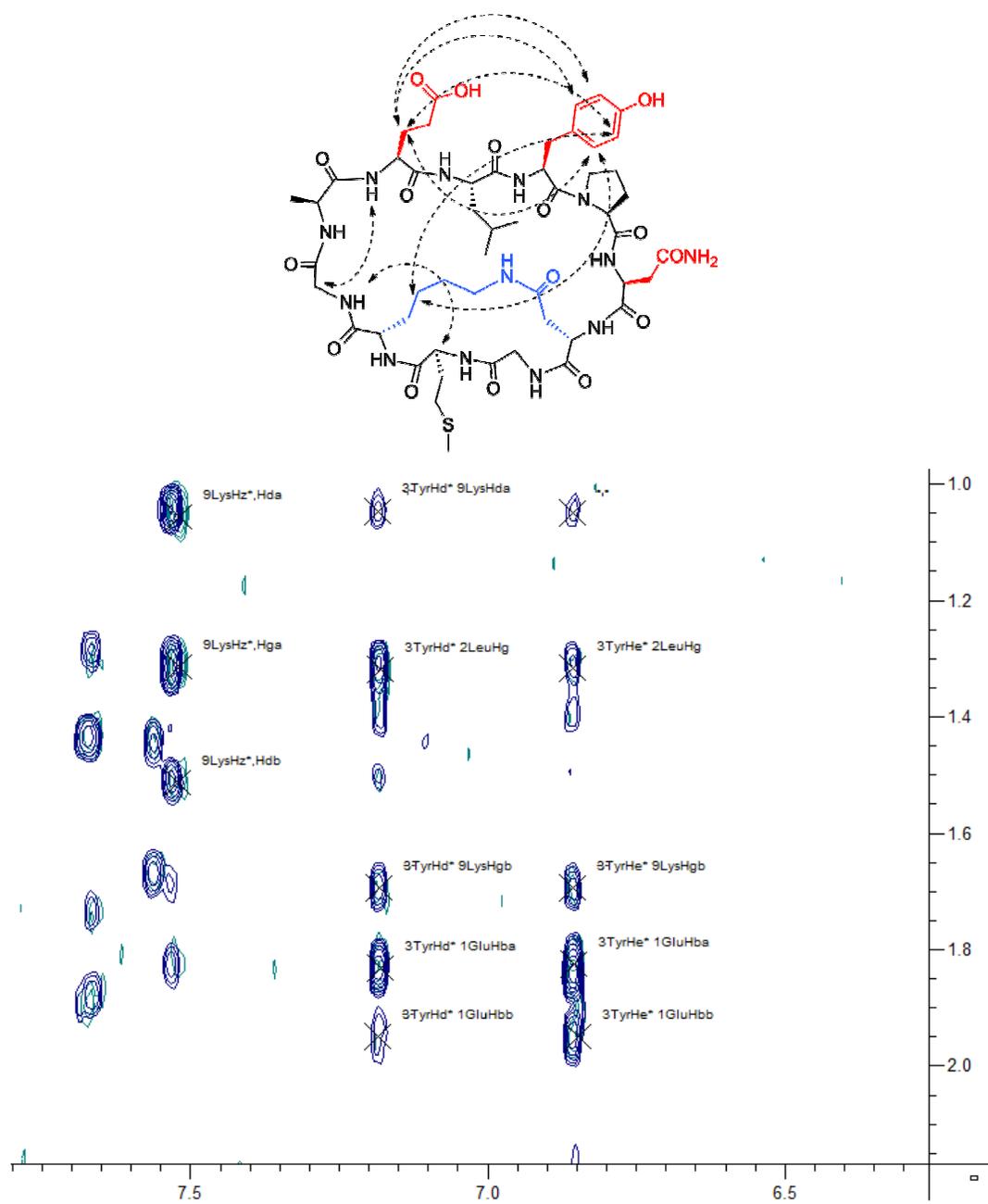


Figure 3.6. *Top*: Select long-range NOEs observed in NMR analysis of **BC1**. *Bottom*: Portion of 2D-NOESY spectrum recorded for **BC1** illustrating several long range NOEs between Glu1-Tyr3 and Tyr3-Lys9 pairs.

Table 3.8. **BC1** Chemical Shifts

<b>BC1</b>	<i>NH</i>	<i>HA</i>	<i>HB</i>	<i>HG</i>	<i>HD</i>	<i>HE</i>	<i>HZ</i>
Glu1	8.02	4.09	1.85/1.95	2.23/2.17			
Leu2	8.09	4.31	1.38/1.41	1.31	0.77/0.82		
Tyr3	8.24	4.91	2.88/3.08		7.18	6.85	
Pro4		4.34	2.02/2.30	1.91	3.75		
Asn5	8.89	4.51	2.87/2.94				
Asp6	7.95	4.82	2.59/2.69				
Gly7	8.37	3.77/4.12					
Met8	8.63	4.68	2.06/1.99	2.51		2.25	
Lys9	8.41	4.46	1.82/1.70	1.31/1.69	1.04/1.51	2.85/3.33	7.52
Gly10	8.4	3.71/4.31					
Ala11	8.75	4.30	1.40				

**Table 3.9.** NOE-derived distance constraints for **BC1**. A total of 73 distance constraints were generated from the 2D NOESY experiment performed on the 600 MHz spectrometer, and 17 unique, additional constraints were generated from the same experiment performed on the 800 MHz spectrometer. Resonance nomenclature is as follows: Residue number, residue name, proton designation, followed by “a” or “b” for diastereotopic methylene protons or magnetically inequivalent aromatic protons. Ambiguous protons are denoted with an asterisk. Residues are numbered 1 through 11 in the order Glu-Leu-Tyr-Pro-Asp-Gly-Met-Lys-Gly-Ala.

Restraint	Resonances	Value (Ang)	Upper Limit	Lower Limit	Difference	Mean Violation	Fraction of structures violated
1	6AspH-7GlyH	3.020	3.624	2.416	1.208	0.051	1
2	2LeuHda*-3TyrHa	3.631	4.357	2.905	1.452	0.062	1
3	8MetHa-10GlyH	2.761	3.313	2.209	1.104	0.071	1
4	2LeuHa-2LeuHba	3.297	3.957	2.638	1.319	0.096	1
5	1GluH-2LeuHa	2.907	3.489	2.326	1.163	0.221	1
6	10GlyHab-11AlaH	2.704	3.245	2.163	1.082	0.232	1
7	4ProHa-4ProHbb	3.291	3.949	2.633	1.316	0.266	1
8	4ProHa-5AsnH	2.559	3.071	2.047	1.024	0.281	1
9	4ProHga-5AsnH	3.652	4.382	2.922	1.461	0.051	0.967
10	2LeuHdb*-3TyrHa	3.918	4.702	3.134	1.567	0.071	0.967
11	1GluH-10GlyHab	3.411	4.093	2.729	1.364	0.877	0.967
12	3TyrHbb-3TyrHa	3.202	3.842	2.561	1.281	0.071	0.9
13	10GlyHaa-11AlaH	2.931	3.518	2.345	1.173	0.105	0.867
14	6AspHba-9LysHz*	2.917	3.501	2.334	1.167	0.060	0.767
15	9LysHdb-9LysHea	3.225	3.870	2.580	1.290	0.264	0.733
16	3TyrHe*-9LysHgb	3.787	4.544	3.029	1.515	0.012	0.633

17	8MetH-7GlyHab	2.755	3.306	2.204	1.102	0.020	0.567
18	4ProHba-4ProHda	2.870	3.444	2.296	1.148	0.030	0.5
19	3TyrHd*-1GluHba	3.410	4.093	2.728	1.364	0.009	0.467
20	9LysHda-9LysHgb	3.576	4.291	2.860	1.430	0.083	0.467
21	9LysHz*-9LysHga	3.306	3.967	2.645	1.322	0.055	0.367
22	9LysHz*-9LysHda	3.289	3.947	2.631	1.316	0.035	0.3
23	6AspHbb-9LysHz*	2.876	3.451	2.301	1.150	0.059	0.3
24	9LysH-9LysHga	3.194	3.833	2.555	1.278	0.023	0.267
25	5AsnH-5AsnHbb	3.686	4.424	2.949	1.475	0.095	0.267
26	9LysHz*-9LysHeb	3.234	3.881	2.587	1.294	0.100	0.267
27	3TyrHa-3TyrHd*	2.981	3.577	2.384	1.192	0.020	0.233
28	11AlaHa-1GluHba	3.980	4.776	3.184	1.592	0.015	0.133
29	9LysHba-9LysH	3.223	3.868	2.579	1.289	0.130	0.133
30	6AspHbb-7GlyH	3.318	3.982	2.655	1.327	0.004	0.1
31	3TyrHe*-1GluHba	3.179	3.815	2.543	1.272	0.010	0.1
32	5AsnHa-5AsnHba	3.285	3.942	2.628	1.314	0.040	0.1
33	3TyrHd*-9LysHgb	3.871	4.645	3.097	1.548	0.005	0.067
34	8MetH-7GlyH	3.449	4.139	2.759	1.380	0.008	0.067
35	8MetHba-8MetH	2.907	3.488	2.326	1.163	0.008	0.067
36	2LeuHg-3TyrHd*	3.417	4.100	2.734	1.367	0.015	0.067
37	6AspH-5AsnH	3.050	3.660	2.440	1.220	0.021	0.067
38	2LeuHdb*-3TyrHe*	4.422	5.306	3.537	1.769	0.060	0.067
39	3TyrHba-3TyrH	3.109	3.730	2.487	1.243	0.106	0.067
40	3TyrHe*-1GluHbb	3.627	4.352	2.901	1.451	0.004	0.033
41	3TyrHd*-1GluHbb	4.706	5.647	3.765	1.882	0.007	0.033

42	2LeuH-2LeuHg	3.074	3.689	2.460	1.230	0.045	0.033
43	2LeuHa-3TyrH	2.633	3.159	2.106	1.053	0.233	0.033
44	6AspH-6AspHba	3.223	3.868	2.579	1.289	None	None
45	10GlyH-10GlyHaa	2.714	3.257	2.171	1.086	None	None
46	9LysHz*-9LysHdb	3.606	4.327	2.885	1.442	None	None
47	8MetH-7GlyHaa	3.048	3.658	2.439	1.219	None	None
48	3TyrHe*-2LeuHg	3.963	4.755	3.170	1.585	None	None
49	3TyrHba-3TyrHa	3.576	4.292	2.861	1.431	None	None
50	6AspH-6AspHbb	3.106	3.727	2.485	1.242	None	None
51	5AsnH-5AsnHba	3.427	4.112	2.741	1.371	None	None
52	3TyrHa-3TyrH	3.220	3.864	2.576	1.288	None	None
53	3TyrHa-4ProHda	2.562	3.074	2.049	1.025	None	None
54	3TyrHba-4ProHda	3.830	4.595	3.064	1.532	None	None
55	11AlaHa-11AlaH	3.202	3.843	2.562	1.281	None	None
56	9LysHba-9LysHga	3.169	3.803	2.535	1.268	None	None
57	1GluH-11AlaHb*	3.277	3.932	2.621	1.311	None	None
58	5AsnHa-5AsnHbb	3.177	3.813	2.542	1.271	None	None
59	9LysHea-9LysHga	3.534	4.241	2.827	1.414	None	None
60	8MetHa-8MetH	3.055	3.665	2.444	1.222	None	None
61	4ProHga-4ProHda	3.594	4.313	2.876	1.438	None	None
62	3TyrHbb-4ProHda	3.480	4.176	2.784	1.392	None	None
63	9LysHda-9LysHea	3.310	3.972	2.648	1.324	None	None
64	3TyrHba-4ProHda	3.918	4.702	3.134	1.567	None	None
65	7GlyHab-7GlyH	2.909	3.490	2.327	1.163	None	None
66	9LysHz*-9LysHea	2.870	3.445	2.296	1.148	None	None

67	3TyrHbb-4ProHda	3.463	4.156	2.770	1.385	None	None
68	3TyrHbb-3TyrHd*	2.793	3.352	2.235	1.117	None	None
69	1GluH-11AlaH	2.935	3.522	2.348	1.174	None	None
70	6AspH-5AsnH	3.098	3.718	2.478	1.239	None	None
71	2LeuH-2LeuHa	2.529	3.034	2.023	1.011	None	None
72	7GlyHaa-7GlyH	2.888	3.466	2.310	1.155	None	None
73	2LeuHdb*-3TyrHd*	5.053	6.063	4.042	2.021	None	None
74	5AsnHa-6AspH	3.250	3.899	2.600	1.300	None	None
75	1GluH-1GluHba	3.431	4.118	2.745	1.373	None	None
76	9LysHeb-9LysHga	3.477	4.173	2.782	1.391	None	None
77	9LysHdb-9LysHga	2.802	3.362	2.242	1.121	None	None
78	3TyrHba-3TyrHd*	2.717	3.260	2.173	1.087	None	None
79	6AspHa-7GlyH	3.077	3.693	2.462	1.231	None	None
80	2LeuHa-2LeuHdb*	3.501	4.201	2.800	1.400	None	None
81	4ProHbb-4ProHda	3.638	4.366	2.910	1.455	None	None
82	1GluH-1GluHbb	3.200	3.840	2.560	1.280	None	None
83	9LysH-9LysHa	2.667	3.200	2.133	1.067	None	None
84	1GluH-11AlaH	3.054	3.665	2.443	1.222	None	None
85	2LeuHa-2LeuHda*	3.432	4.119	2.746	1.373	None	None
86	3TyrHd*-3TyrH	3.375	4.050	2.700	1.350	None	None
87	11AlaHb*-11AlaH	2.851	3.422	2.281	1.141	None	None
88	4ProHbb-5AsnH	3.833	4.599	3.066	1.533	None	None
89	5AsnHa-5AsnH	2.676	3.211	2.141	1.070	None	None
90	9LysHba-9LysHda	3.524	4.229	2.819	1.410	None	None

Secondary chemical shifts are defined as the typical value for that chemical shift in a random coil peptide, minus the observed chemical shift. Large secondary chemical shifts for NH and H $\alpha$  protons are indicative of a well-ordered backbone structure. Some of the largest NH secondary chemical shifts were observed for the Glu1 and Ala11 residues, suggesting that the local structure at this site is particularly rigid and atypical. Molecular dynamics simulations of **BC1** in explicit solvent revealed that these NH protons are part of a high-occupancy water-binding pocket, which would explain the observed secondary chemical shifts (see section 3.4). The diastereotopic H $\alpha$  protons of Gly10 exhibit secondary chemical shifts of opposite sign, indicating that these atoms experience very distinct electronic environments. Molecular dynamics simulations showed the backbone NH of Gly10 contributes to an adjacent water-binding site, consistent with the high degree of order of this region. Taken together, these large secondary chemical shifts provide direct evidence for well-defined structure as well as support for the hypothesis, suggested by molecular dynamics simulations, of distinct water-binding pockets formed by the peptide backbone.

The solution structure of **BC1** revealed that the side-chains of Glu1 and Tyr3 form a pTyr-mimicking epitope. This is a unique epitope that mimics pTyr using naturally-occurring amino acids, and without the polyanionic character of traditional pTyr isosteres. Pro4 organizes a turn that positions Asn5 on the same face as Glu1 and Tyr3, consistent with established binding modes for ligands of Grb2-SH2.<sup>62</sup> A representative **BC1** structure was docked into the pTyr-binding site of Grb2 (figure 3.8D). This model demonstrates that **BC1** is pre-organized in a manner consistent with direct binding to the pTyr-binding pocket of Grb2-SH2,

and is capable of making many of the same contacts observed in crystal structures with pTyr-containing peptides.<sup>62</sup>

Analysis of our NMR solution structure of **BC1** in comparison with established Grb2-SH2 domain ligands, such as the -pYVNV- sequence reveals distinct similarities in structure between our bicyclic peptide ligand and pTyr-containing molecules. The earliest crystal structure of the Grb2-SH2 domain in complex with a pTyr-containing peptide by Garcia-Echeverria demonstrated that phosphotyrosine and asparagine are situated at positions *i* and *i*+2 of a type-I  $\beta$ -turn that spatially separates the phenol and carboxamide side-chains of these residues so as to promote binding of each epitope into a distinct pocket on the Grb2 surface. Analysis of our NMR structure of **BC1** reveals that Pro4 exhibits  $\phi$  and  $\psi$  angles of  $-33.1^\circ$  and  $-58.7^\circ$ , characteristic backbone dihedral angles of residues situated at position *i*+1 of a type-I  $\beta$  turn (figure 3.9). This was a gratifying result, as Pro4 is located between the phosphotyrosine mimic defined by the Glu1 and Tyr3 residues and the Asn5 residue, indicating that Pro4 is indeed nucleating a turn at this region of the macrocycle. This observation validated our original hypothesis that insertion of a proline residue at this region in the original monocyclic peptide **G1HT1** would stabilize a turn element at this site. Additionally, visual inspection of Garcia-Echeverria's crystal structure of Grb2-SH2 in complex with -pYVNV- reveals that in addition to this turn, the side-chain of pTyr is oriented such that it extends outward and away from the Asn carboxamide. This is manifested in a side-chain  $\chi_1$  angle (dihedral angle defined by N-C $\alpha$ -C $\beta$ -C $\gamma$ ) of  $-38.5^\circ$ . The  $\chi_1$  angle of the tyrosine side-chain in **BC1** is in good agreement with this value ( $\chi_1 = -33.1^\circ$ ), and a visual overlay of the two

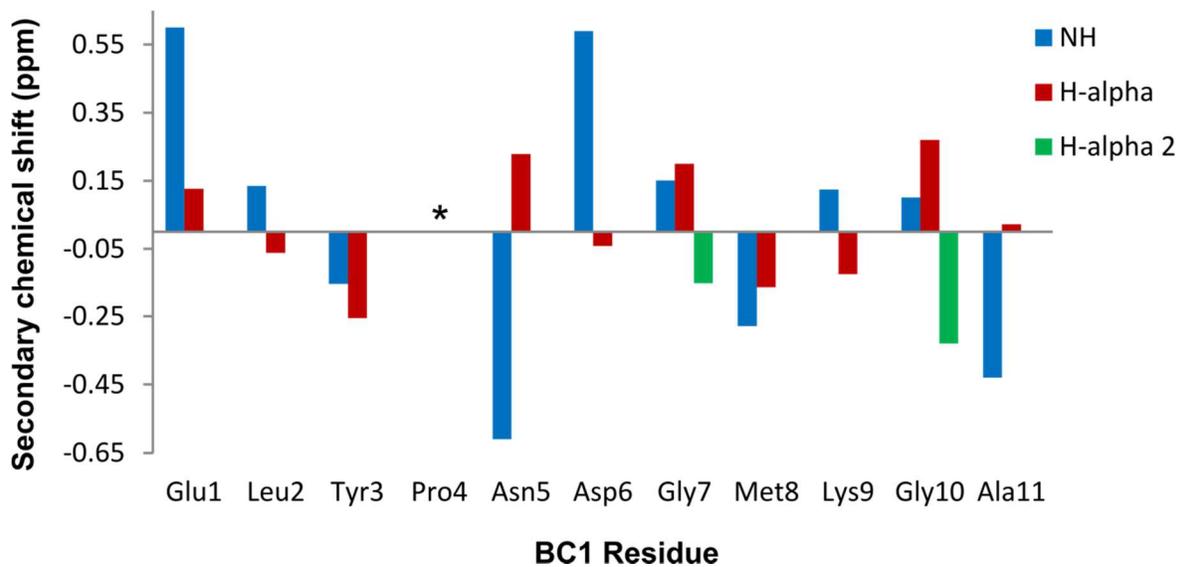


Figure 3.7. Secondary chemical shifts of backbone NH and H $\alpha$  protons. Sequence-corrected chemical shifts of random coil peptides were calculated using the method of Kjaergard, et. al. (See Materials and Methods).

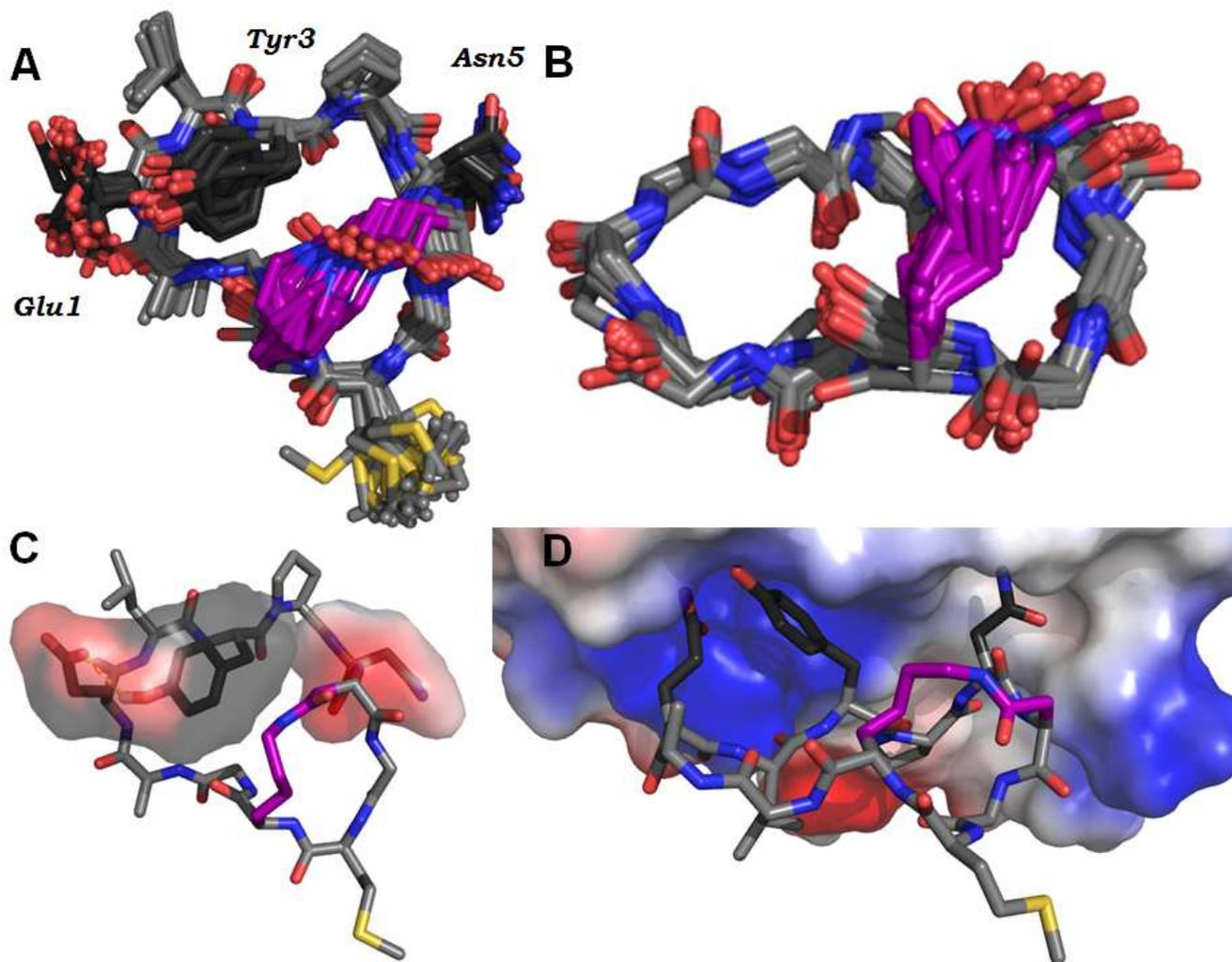


Figure 3.8. (A) Thirty lowest-energy structures calculated by two-stage simulated annealing using 90 NOE-derived distance constraints. (B) Thirty-structure ensemble showing only backbone and cross-link atoms (backbone RMSD = 0.33 angstroms). (C) The NMR structural ensemble reveals the Glu and Tyr residues are in close proximity, forming a discontinuous phosphotyrosine mimetic (shown as a surface). (D) Representative structure from NMR ensemble was docked into the pTyr-binding site of the Grb2-SH2 domain (Molecular Operating Environment software package). This model suggests the Glu, Tyr and Asn residues bind distinct surfaces at the Grb2 interface.

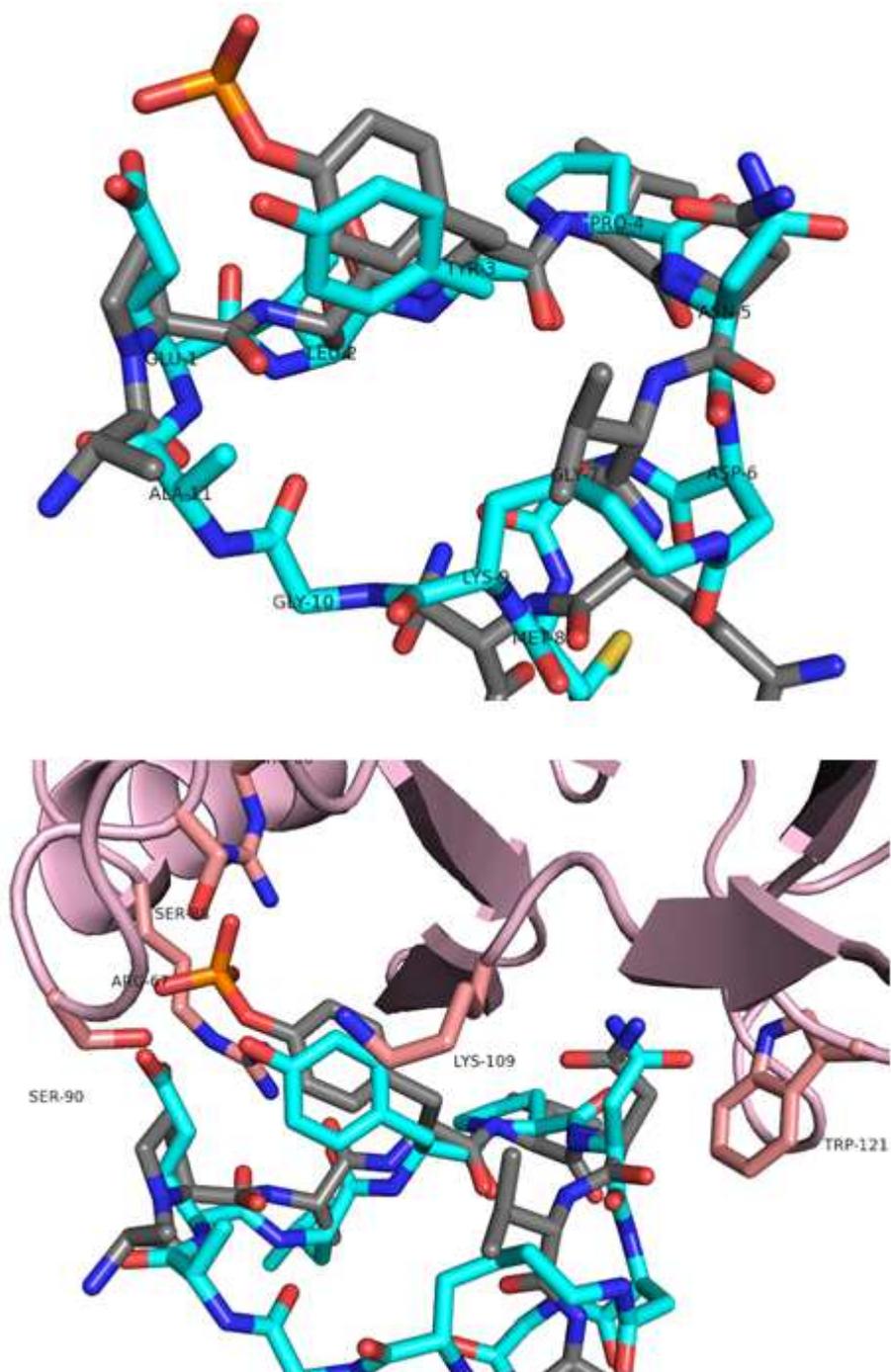


Figure 3.9. Overlay of representative structure of **BC1** from NMR structural ensemble with pYVNV Grb2-SH2 ligand from pYVNV:Grb2-SH2 domain crystal structure (PDB ID: 1TZE). For clarity, the overlay is shown in the absence (top) and presence (bottom) of the protein. The orientation of the backbone and side chains at the phosphotyrosine and turn sites align well in the two structures, indicating **BC1** indeed stabilizes a pTyr-like epitope and a turn element for the presentation of a competent Grb2-SH2-binding pharmacophore.

structures reveals clearly that the orientation of the phosphotyrosine-like pharmacophore of **BC1** with respect to the peripheral asparagine is similar to that observed in the pYVNV peptide (figure 3.9). Moreover, NMR solution structure ensemble is highly convergent in this regions, indicating the peptide is particularly well-ordered at this site within the bicycle. Collectively, these structural data represent explicit evidence validating our hypothesis that the careful application of successive conformational constraints within this system can effectively stabilize a bioactive geometry. That this geometry is reminiscent of that observed in an inhibitor containing an explicit phosphotyrosine residue despite the absence of this functionality in **BC1** bespeaks the unique ability of this peptide to present a novel, discontinuous phosphotyrosine-like epitope.

In addition to inferring peptide structure using NOE cross peaks to derive inter-proton distance constraints, NOESY and ROESY spectra also provided direct evidence of the presence of discrete, bound water molecules at the surface of **BC1**. Our NMR structural model reveals the presence of several unique structural elements, including a continuous region of inward-facing amide N-H protons. Molecular dynamics simulations suggest that the bicyclic scaffold binds distinct water molecules at the peptide surface (*vide infra*), a phenomenon that effectively buries Hydrogen-bond donors in the molecule's interior and may provide an explanation for the ability of the peptide to be internalized by mammalian cells. The occurrence of individual water molecules that persist at the surface of biomolecules has been observed in the case of large, globular proteins,<sup>101-103</sup> but is rare for smaller peptides. We further analyzed the results of NOESY and ROESY experiments and looked for cross-peaks between **BC1**

protons and the water OH resonance at 4.87 ppm. Distinct peptide-water cross-peaks in NOESY and ROESY spectra arise from a transfer of magnetization between peptide protons and solvent protons due to either (1) rapid chemical exchange between labile peptide protons and bulk solvent or (2) a direct NOE between a peptide proton and the proton of a long-lasting water molecule.<sup>104,105</sup> In general, while NOESY experiments provide direct evidence of a NOE between a non-exchangeable peptide proton and the proton of a distinct water molecule, these experiments are unable to distinguish between a direct NOE and an exchange-mediated magnetization transfer for a solute proton that is subject to chemical exchange. ROESY spectra, however, are capable of differentiating between a direct NOE and a chemical exchange-mediated relay of magnetization: chemical exchange gives rise to ROESY cross-peaks of the same sign (+) as the diagonal cross-peaks, while peaks that arise from a direct NOE are of the opposite sign (-).<sup>104,105</sup> Analysis of the NOESY and ROESY spectra of **BC1** revealed cross peaks between water protons and peptide protons that are implicated in water-binding sites on the peptide surface. Regions of the NOESY and ROESY spectra containing these cross-peaks are shown in figure 3.10.

In order to validate the proposed structural model for **BC1**-Grb2 binding, bicyclic variants of **BC1** were synthesized in order to test the effects of distorting the hypothesized phosphotyrosine-mimicking region of the peptide. In one such variant, glutamic acid was substituted with aspartic acid in order to impede the ability of the phosphotyrosine-like epitope to extend into the pTyr-binding pocket of Grb2, while in the other, tyrosine was substituted with phenylalanine so as to eliminate several key interactions proposed by the binding model (figure 3.8D).

Table 3.10 NMR structural data and refinement statistics.

	BC1
Experimental restraints	
Total distance restraints from NOEs	90
Dihedral restraints	0
NOE classes	
Intra-residue	49
Sequential	27
Medium-range	7
Long-range	7
Total no. of experimental restraints	90
Restraint violations	
> 0.5 Å	0
> 0.3 Å	15
> 0.1 Å	62
≤ 0.1 Å	703
Rms deviations from ideal stereochemistry	
Bonds (Å)	0.0065 ± 0.00032
Angles (deg)	0.6286 ± 0.0186
Impropers (deg)	0.4464 ± 0.0356
Ramachandran analysis of structures	
Residues in most favored regions	17.8%
Residues in additionally allowed regions	33.9%
Residues in generously allowed regions	27.8%
Residues in disallowed regions	20.6%
Coordinate precision (Å)	
Backbone	0.316 ± 0.1673
All heavy atoms	0.785 ± 0.2341

Our structural model reveals an important Hydrogen bond between the tyrosine hydroxyl and LyS119 in the cationic pTyr binding site of Grb2, as well as an intramolecular ion-dipole interaction between the tyrosine hydroxyl and the glutamic acid carboxylate. We hypothesize that this constraint stabilizes the overall conformation of **BC1**, and that removing this hydrogen bond will distort the global conformation of **BC1** and diminish the peptide's capacity for phosphotyrosine mimicry. In order to explicitly test this hypothesis, we used a fluorescence polarization competitive binding assay in order to monitor the ability of these peptides to displace a fluorescently-labeled phosphotyrosine-containing peptide from purified, recombinant Grb-SH2 domain. The assay was set up as previously described (See Chapter One, Materials and Methods). As expected, these variants exhibited little affinity for the Grb2-SH2 domain in this assay (figure 3.11), providing evidence for the importance of the chemical functionality and the relative orientations of the glutamate and tyrosine residues of **BC1** as proposed by the binding model.

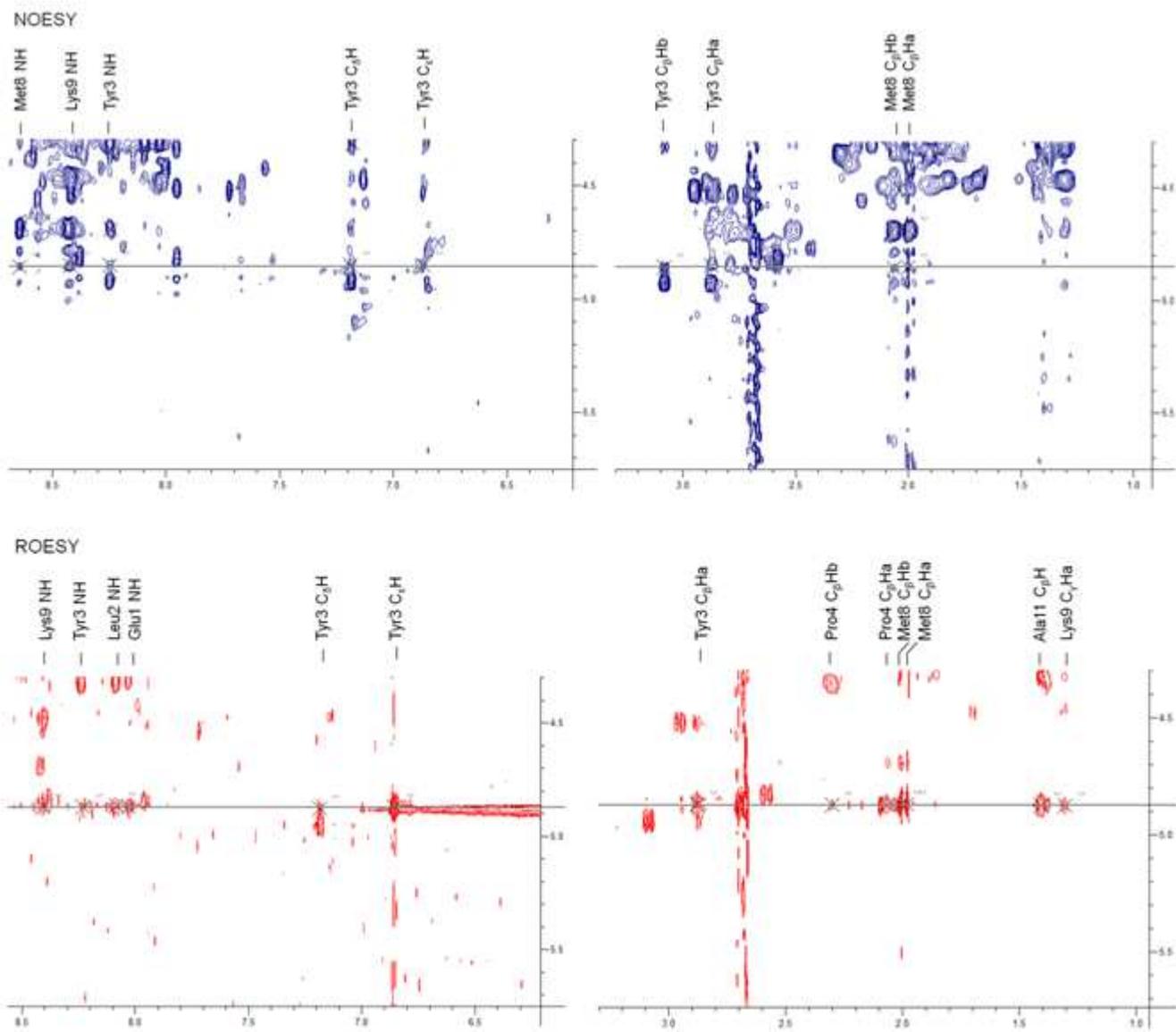


Figure 3.10. NOESY and ROESY spectra recorded for **BC1** showing cross-relaxation between distinct backbone NH and CH protons with the water line (4.78 ppm in 10% DMSO).

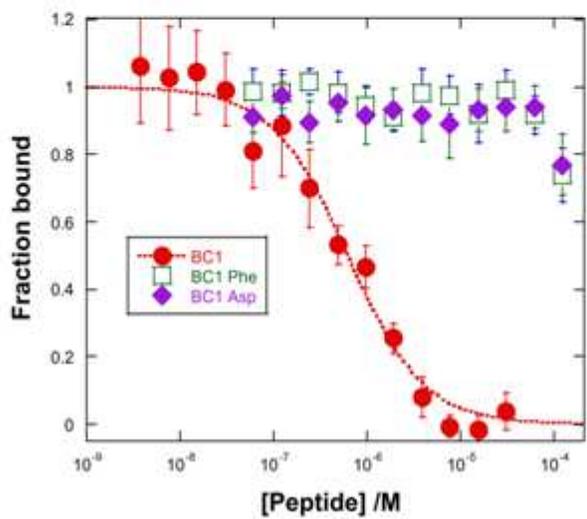
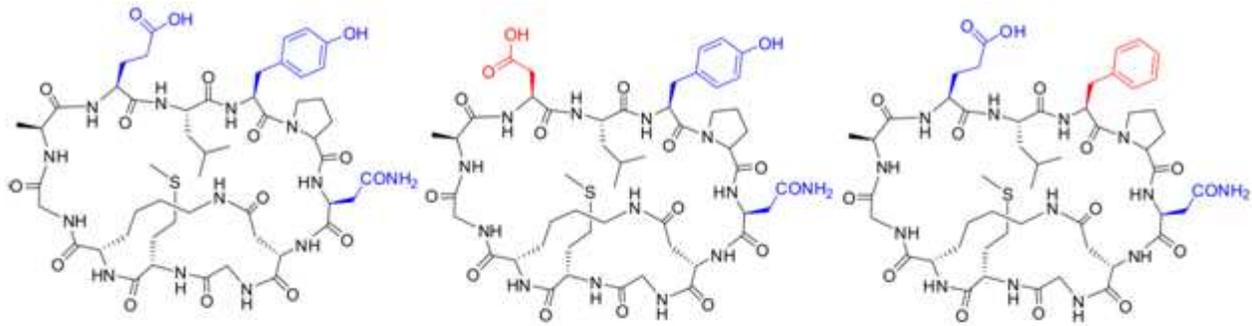


Figure 3.11. FP competitive binding experiments performed with **BC1** variants containing altered phosphotyrosine mimetics. Disrupting the hydrogen bonding ability of this epitope leads to a dramatic reduction in inhibitory activity.

## Materials and Methods

**BC1** was synthesized as described previously using Fmoc/t-bu solid phase peptide synthesis with an orthogonal protecting group strategy. The sample was purified to >95% purity by RP-HPLC using a water/acetonitrile/0.1% TFA solvent system. Lyophilized **BC1** was dissolved in 500  $\mu$ L 9:1 H<sub>2</sub>O:*d*<sub>6</sub>-DMSO buffered at pH 6 with deuterated acetic acid/sodium acetate to yield a 0.76 mM solution. DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) was added to a final concentration of 100 $\mu$ M to provide an internal reference standard, and the sample was titrated to a pH of 6.0 using aqueous NaOH.

All experiments were conducted at 15° C. 2D TOCSY, NOESY, ROESY and <sup>1</sup>H-<sup>13</sup>C HSQC experiments were conducted on a Bruker Avance 600 MHz spectrometer (Tufts University Department of Developmental, Molecular, and Chemical Biology, Boston, MA), and additional NOESY experiments were conducted on Bruker Avance 800 MHz spectrometer with a 5 mm TCI cryoprobe with z-axis pulsed magnetic field gradients (Brandeis University, Waltham, MA). For all 2D experiments, suppression of the water resonance was achieved using gradient suppression with selective irradiation at the  $\omega_1$  frequency of 2821 Hz (600 MHz spectrometer) or 3778 Hz (800 MHz spectrometer). TOCSY spectra were acquired with a spectral width of 7716 Hz in each dimension, a 1.3 s recycle delay and acquisition times of 266 ms and 33 ms along F2 and F1, respectively. A mixing time of 40 ms was used. Signal averaging of 36 scans resulted in a total acquisition time of 19 hours. <sup>1</sup>H-<sup>13</sup>C HSQC spectra were collected with a spectral width of 7788 Hz along the F2 dimension and 11318 Hz

along F1. A recycle delay of 1.3 s was used between successive scans, and acquisition times of 132 ms and 13 ms along F2 and F1, respectively. A total time of 8 hours was required for spectral collection due to signal averaging of 160 scans. NOESY spectra collected on the 600 MHz spectrometer were acquired with a spectral width of 7716 Hz, a recycle delay of 1.3 s and acquisition times of 133 ms and 26 ms along F2 and F1, respectively, with mixing applied for 500 ms. ROESY spectra were acquired with a spectral width of 7716 Hz, a recycle delay of 1.3 s and acquisition times of 266 ms and 26 ms along F2 and F1, respectively, with mixing applied for 500 ms. NOESY spectra acquired on the 800 MHz spectrometer were collected with a spectral width of 8012 Hz, a recycle delay of 2 s and acquisition times of 130 ms and 16 ms along F2 and F1, respectively, with mixing applied for 500 ms. In each experiment, a total of 128 scans were averaged, resulting in a total collection time of 24 hours.

All spectra were processed using the Topsin software package (Bruker). Processed spectra were imported into the CCPNMR program, which was used for making peak assignments and generating NOE-derived distance restraints. Molecular dynamics simulated annealing was performed in the CNSolve software suite. 90 non-redundant distance restraints were used in simulated annealing from 10 random initial structures. Simulated annealing was executed using a high temperature annealing stage of 1,000 steps, followed by two slow-cool annealing stages, each of 2,000 steps. These were followed by a 10 cycles of a 200 step energy minimization phase. The resulting 30 lowest-energy structures were visualized using PyMOL. Calculated distance restraints are given above. Molecular structures were visualized using PyMOL.

### 3.4 – The interaction of BC1 with aqueous solvent: a molecular dynamics simulation

At the onset of this investigation, we hypothesized that specific structural features of **BC1** promote its association with the plasma membrane, leading to uptake via endocytosis. The backbone of Ala11 through Tyr3 forms an unusual turn structure that points amide protons inward and carbonyl oxygens out towards solvent (figure 3.8B). This effectively shields the polar hydrogen bond donors from aqueous solvent and may promote peptide-membrane interactions. To provide evidence for this hypothesis, we performed all-atom molecular dynamics simulations on a tethered model of the **BC1** solution structure. During a 100 ns MD simulation in TIP4P explicit water<sup>106</sup> using the OPLS/AA forcefield,<sup>107</sup> we observed three distinct water molecules that bind three separate pockets formed by backbone amides and carbonyls (figure 3.12). Discrete, bound waters are typical for protein surfaces,<sup>101–103</sup> but this behavior is highly unusual for small, constrained peptides. The most tightly bound water molecule binds in a pocket defined by the inward-facing amide protons of Ala11, Glu1, Leu2, and Tyr3, providing a unique explanation for the unusual conformation of this loop. Over the course of the 100 ns simulation, this pocket exhibits 99% occupancy, with an average water molecule residence time of 1.4 ns and a calculated binding  $\Delta\Delta G$  of -2.74 kcal/mol. The amide NH protons of Glu1 and Ala11 exhibit chemical shifts that differ greatly from expected values for a random coil (figure 3.7), providing further evidence that this pocket represents a unique electronic environment. We hypothesize that these discrete water-binding pockets

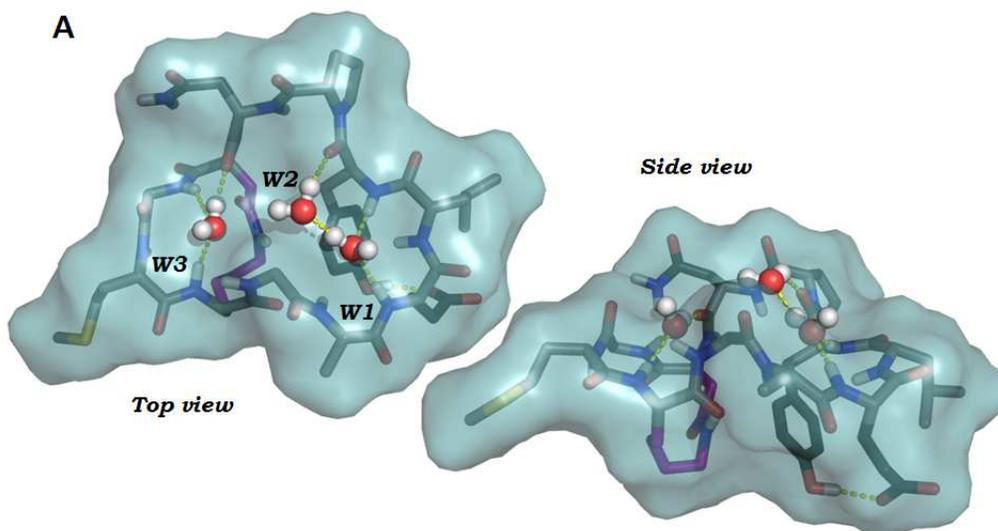
effectively shield multiple backbone amides from the bulk solvent, increasing the effective lipophilicity of the peptide.

In our evaluation of explicit solvent-peptide interactions, we calculated the density of water molecules within 3 Å of the peptide. Three bound water molecules are identified and located in the shallow pockets formed by the peptide backbone atoms (water W1, W2, and W3 in figure 3.12). All of these bound water molecules are partially buried, with some solvent-exposed surface area. The residence times are ~1400 ps, ~60 ps, and ~60 ps for water W1, W2 and W3 respectively. One measure of the bound water stability is the occupancy of the binding pocket. In this study, the occupancy of a pocket was calculated based on the minimal distances of the solvent molecules to the center of the pocket. The center of each pocket was defined based on the oxygen atom of the water that occupies the pocket in a typical conformation with all three binding pockets hydrated. A pocket was considered occupied if the minimal distance was smaller than 1.2 Å. The occupancies for W1, W2 and W3 were 0.99, 0.70 and 0.34 respectively. The free energy differences (occupied relative to empty) calculated from the equation  $-kT \ln(\text{occup}/(1-\text{occup}))$  were -2.74 kcal/mol for W1, -0.51 kcal/mol for W2 and 0.40 kcal/mol for W3.

For insight into the interaction patterns between the bound water molecules and the peptide, a hydrogen-bond analysis was performed. We used the geometric definition of a hydrogen bond: donor-acceptor distance is less than 3.5 Å and the angle between the donor-hydrogen vector and acceptor-hydrogen vector is between 150° and 180°. W1 typically formed two hydrogen bonds with the Glu1

and Tyr3 amide protons, and the fractions of the simulation during which these two hydrogen bonds exist are 0.78 and 0.47 respectively. Although W2 formed only one hydrogen bond to the Tyr3 carbonyl oxygen with a relatively small fractional time of 0.40, W2 at this position was stabilized by hydrogen bonding with W1. The fractional time for the hydrogen bond between W1 and W2 was 0.82. In its binding pocket, W3 could make hydrogen bonds with several partners. However, the fractional times for all these hydrogen bonds were small. This observation is consistent with the lower occupancy and positive free energy difference for W3.

All together, our structural and computational data provide original explanations for the structural mimicry and cell uptake of the bicyclic scaffold **BC1**. Initial results with **HT1** and **BC3** indicate such scaffolds may be more readily designed than previously assumed. Future work will examine how varied these scaffolds can be and still be efficiently internalized, and will optimize these and other scaffolds for maximal uptake and endosomal escape. These data will also be used to apply scaffolds towards inhibition of SH2 domains and tyrosine phosphatases, which have been difficult to target effectively and selectively in living systems, as well as other difficult-to-target protein-protein interactions.



Water Site	Residence Time	Fractional Occupancy	$\Delta\Delta G$ bound/unbound (kcal/mol)
W1	1.4 ns	0.99	-2.74
W2	60 ps	0.70	-0.51
W3	60 ps	0.34	0.40

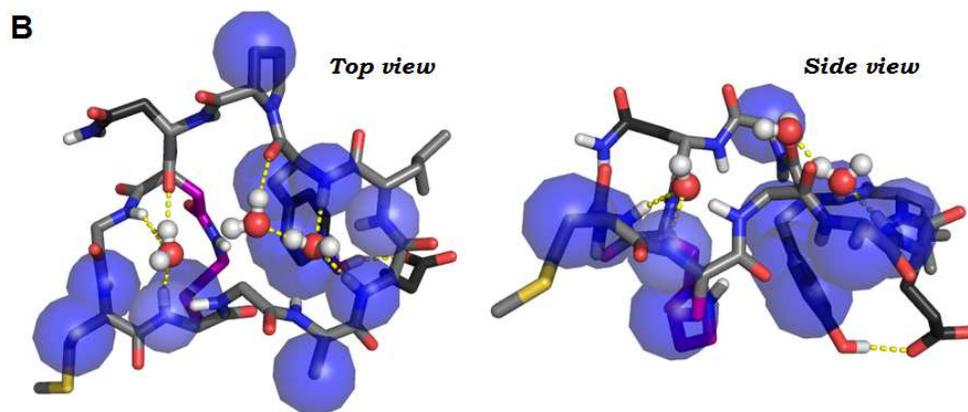


Figure 3.12. Molecular dynamics simulations revealed the presence of distinct water-binding sites on the **BC1** surface (top). Interestingly, these sites align well with residues along the **BC1** backbone and certain side chains that show distinct NOE cross-relaxation with water (these residues are highlighted in blue spheres).

## Materials and Methods

The molecular dynamics (MD) simulation was performed using the Gromacs 4.6.1 suite of programs in conjunction with the OPLS/AA force field<sup>107</sup> and TIP4P water model.<sup>106</sup> The NMR structure was used as the starting structure for the MD simulation. To maintain the NMR structure, all the heavy atoms in the peptide were restrained by a harmonic potential during the entire study. Following energy minimization in vacuum, the peptide was solvated with a cubic box containing 2289 water molecules. The dimension of the water box was chosen such that the minimum distance between any atom of the peptide and the box walls is 1.0 nm. One Na<sup>+</sup> ion was also added to neutralize the system. The resulting system was further optimized by steepest descent algorithm to remove unphysical contacts. Before the production simulation, the solvated and minimized structure was heated gradually in an isobaric-isothermal (NPT) ensemble to the target temperature of 300K within 500 ps to adjust the volume of the system.

The production simulation was run in the NPT ensemble at a temperature of 300K and a pressure of 1 bar. The temperature was maintained using the v-rescale thermostat<sup>108</sup> with a coupling time constant of 0.1 ps. The pressure was regulated using an isotropic Parrinello-Rahman barostat with a coupling time of 2.0 ps and a compressibility of  $4.5 \times 10^{-5} \text{ bar}^{-1}$ . All bonds were constrained with the LINCS algorithm. A 2 fs time step with the leap-frog algorithm was used to evolve the dynamics of the system. The non-bonded interactions (Lennard-Jones and electrostatic) were truncated at 0.9 nm. Long-range electrostatic interactions beyond the cut-off distance were calculated using the Particle Mesh Ewald

summation method<sup>109</sup> with a Fourier spacing of 0.15 nm and an interpolation order of 4. A long-range analytic dispersion correction was applied to both the energy and pressure to account for the truncation of Lennard-Jones interactions. The production run was performed for 102 ns with a data collection frequency of 10 ps. Only the last 100 ns simulations were used for further analysis.

## **Chapter Four: Expanding Conformational Phosphotyrosine Mimicry to Target Protein Tyrosine Phosphatase 1B**

### **4.1 – Protein Tyrosine Phosphatase 1B as a Drug Target for Obesity and Type-II Diabetes**

Protein tyrosine phosphatase 1B (PTP1B) belongs to a large family of essential enzymes whose proper biological functioning and regulation is paramount to directing healthy cellular activity. Protein phosphatases, along with their protein kinase counterparts, comprise a substantial proportion of the eukaryotic proteome. The human proteome includes 518 protein kinases and over 130 protein phosphatases.<sup>110</sup> Members of these enzyme superfamilies play an essential role in the propagation of signal transduction pathways. Due to the quantity and nature of the events triggered by signal transduction pathways, it is clear that the cellular response to an external signal must be highly coordinated; indeed, the aberrant functioning of signaling pathways mediated by PTP1B has been implicated in host of human disease states, including growth and developmental disorders, cancer, autoimmune diseases, obesity, diabetes and others.<sup>111</sup>

Protein tyrosine phosphatase 1B (PTP1B) is a 321-residue protein that plays important roles in the signal transduction pathways of insulin and leptin. PTP1B is expressed in nearly all tissues, and the proper regulation of the activity and concentration of this enzyme is vital to the healthy functioning of liver, muscle and adipose tissue. In general, PTP1B is regarded as a negative regulator of

insulin and leptin signaling. Insulin, a peptide hormone comprised of 51 amino acids, is released from the pancreas in response to elevated concentrations of blood glucose. Upon arrival at the surface of a target cell, insulin binds the insulin receptor (IR), a receptor tyrosine kinase, so named for its intrinsic tyrosine autophosphorylation activity.<sup>7</sup>

Binding of insulin to the IR induces autophosphorylation events within the cytoplasmic tails of the protein complex. The resulting pTyr residues serve as binding sites for a host of insulin substrate proteins (IRS), which are themselves phosphorylated and serve as adaptor proteins in that they subsequently bind the next protein involved in the signaling cascade. The effects of insulin signaling are numerous, and include glucose uptake and up-regulation of gluconeogenesis in the liver. PTP1B is known to dephosphorylate the IR, as well as one of the IRS proteins (IRS-1), thus reducing cellular sensitivity to insulin. In fact, diminished levels of PTP1B have been shown to enhance insulin sensitivity *in vivo*. PTP1B also plays a pivotal role in leptin signal transduction. Leptin, a 16 kDa protein hormone, is synthesized chiefly by adipose tissue and is generally released in response to elevated triglyceride content.

Leptin is a known appetite suppressant, and acts on cells of the hypothalamus to induce a state of satiation.<sup>112</sup> As leptin binds its receptor in target tissue, Janus-associated kinase 2 (JAK2) is phosphorylated upstream in the signaling pathway. It is here that PTP1B serves as a negative regulator of leptin functioning, as PTP1B catalyzes the dephosphorylation of JAK2 and thus reduces leptin sensitivity.<sup>113</sup> Leptin and insulin are known to play a critical role in the

maintenance of healthy body weight. A landmark study which demonstrated the gross obesity of mice deficient in the functional leptin receptor (*ob/ob* mice) has implicated reduced leptin sensitivity in obesity. Moreover, genetic evidence has linked the over-expression of PTP1B in humans to obesity and type-2 (insulin-independent) diabetes. Consistent with this link, mice deficient in PTP1B (*PTP1B*<sup>-</sup>/*PTP1B*<sup>-</sup>) exhibit enhanced sensitivity to leptin and insulin, and are indeed immune to obesity even upon subjection to excessively fatty diets.<sup>112</sup>

One challenge medicinal chemists are faced with is that any inhibitor of neuronal PTP1B must be able to penetrate the blood-brain barrier. Though this has proven difficult by traditional means, the anti-obesity effects of reducing PTP1B *in vivo* are well-established, and this molecule is indeed an attractive target for obesity and type-2 diabetes<sup>7,114–116</sup>. The pages that follow will review the structure and function of PTP1B, as well as the approaches that have been employed to develop inhibitors of PTP1B activity.

To date, much is known regarding the catalytic mode of action of PTP1B. This phosphatase binds with high affinity to phosphotyrosine residues and utilizes an activated cysteine, C215, to cleave the phosphoryl group. The reaction is believed to proceed as follows: The Cys215 residue (pKa approximately 5.4)<sup>117</sup> is ionized by a nearby aspartate, generating a nucleophilic thiolate moiety within the active site. Upon binding of phosphotyrosine to PTP1B, a loop consisting of tryptophan, proline and the catalytic aspartate (the so-called WPD loop) closes over the substrate and encapsulates the target within the enzyme active site. Upon deprotonation of Cys215, this aspartic acid residue, Asp181, forms a

Hydrogen bond to the phenolic oxygen of the substrate. This serves to further polarize and enhance the electrophilicity of the phosphate group. The thiolate functionality of Cys215 then engages in a nucleophilic attack on the phosphorylated substrate, resulting in phosphoryl transfer from the substrate to the cysteine side chain to furnish a thiophosphate intermediate. Serving as a general acid-base catalyst, Asp181 then activates a water molecule by forming a hydrogen bond to, and ultimately capturing, one of the water protons. This polarized water then carries out a nucleophilic attack on the phosphate, thus yielding inorganic orthophosphate and restoring the enzyme active site to its initial state.

The binding of the phosphorylated substrate protein to the PTP1B active site is mediated by a host of intermolecular attractions. It stands to reason that the active site is highly cationic in nature, as it engages in polar and ionic contacts with phosphotyrosine, a residue that formally carries two negative charges at physiological pH. The interactions that govern this binding will be reviewed in the sections that follow.

In 1994, Barford and coworkers at Cold Spring Harbor Laboratories reported the first crystal structure of PTP1B.<sup>118</sup> Shortly thereafter, Barford reported the crystal structure of PTP1B in complex with phosphotyrosine-containing substrates.<sup>119</sup> To render the enzyme catalytically-inactive, a PTP1B mutant (PTP1B C215S) was expressed in which the active site Cys215 residue was exchanged for serine. Crystal structures were solved for two distinct complexes – one in which PTP1B was co-crystallized with phosphotyrosine, and another in which the protein was

complexed with a phosphotyrosine-containing hexapeptide. The peptide, DADEpYL-NH<sub>2</sub>, was modeled after a specific autophosphorylation site (residues 988-993) within the cytoplasmic tail of the epidermal growth factor receptor (EGFR).

X-ray data revealed a great deal of information regarding the binding of pTyr to the PTP1B active site. As pTyr settles within the catalytic domain, the so-called WPD loop, consisting of residues Trp179 to Ser187, closes around the substrate. This serves to bring the catalytic Asp181 into the active site, where it engages in ionic interactions with the side chain of Lys120. Pi-stacking interactions are believed to assist in stabilizing the ligand-bound structure, as closing of the WPD loop appropriately positions the phenyl rings of Phe182 and pTyr for hydrophobic interaction. The pTyr recognition motif is made up of several hydrophobic and aromatic residues, including Ala217, Ile219, Tyr46, Val49, Phe182 and Gln262. Tyr46 is believed to interact extensively with the phosphotyrosine target, and is itself stabilized in the ligand-bound conformation by Hydrogen bonding to the side chain hydroxyl of Ser216. The enzyme active site is comprised of the amino acids Cys215 to Arg221, which are held just beneath the phosphotyrosine binding pocket. Cys215 contains the catalytically active thiolate necessary for dephosphorylation. The side chain oxygen of Ser215 was reported to be 3.2 angstroms away from the tetrahedral phosphorus center of the substrate, suggestive that sulfur at this position would be appropriately poised for nucleophilic attack on the phosphate functionality. Moreover, the side chain hydroxyl of Ser215 engages in hydrogen bonding with the side chain carboxylate of Asp181, providing structural evidence for the role of Asp181 as a general acid-

base catalyst in the mechanism of dephosphorylation. The backbone nitrogen atoms of these residues are believed to engage in extensive hydrogen bonding with the peripheral phosphate oxygen atoms of the pTyr substrate. Ionic contacts between these oxygen atoms and the side chain nitrogen atoms of Arg221 further stabilize the ligand-bound conformation.

In 1997, Zhang and coworkers at Albert Einstein College of Medicine reported the discovery of a second aryl phosphate binding site within PTP1B.<sup>120</sup> This finding suggested a paradigm for developing selective modulators of PTP1B activity. One of the difficulties in developing PTP1B inhibitors is achieving target specificity, as the active sites of PTPs are generally conserved, and finding an inhibitor with specificity to PTP1B has been a challenge to medicinal chemists throughout the last decade, as off-target mediated toxicity has been a significant source of attrition of PTP1B therapeutics in clinical trials.<sup>7</sup> Not surprisingly, the discovery of a new aryl phosphate binding site adjacent to but distinct from the active site had set the stage for a new methodology in the search for novel inhibitors of PTP1B. In early analyses of these binding pockets on the enzyme surface, catalytically-inactive PTP1B C215S was co-crystallized with bisphosphophenylmethane (BPPM). X-ray crystallography data indicated that BPPM binds PTP1B in two distinct sites: the enzyme active site, as well as an adjacent non-catalytic site. Crystal structures of two distinct PTP1B:BPPM complexes that differed in the location of the ligand revealed electron densities of the distal portions of the BPPM molecule in each complex overlay with one another in space, an indication that it is not possible for two BPPM molecules to bind PTP1B simultaneously. This finding suggested that the discovery of

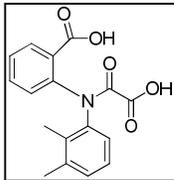
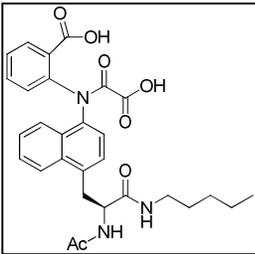
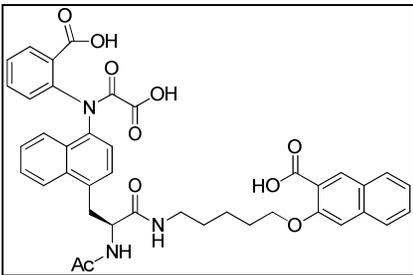
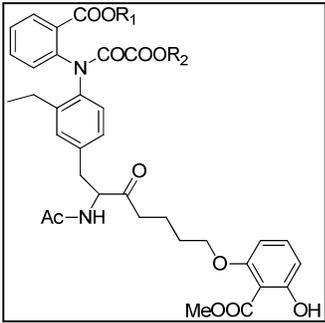
molecules capable of engaging both of these sites on the pTyr surface may represent a promising approach to selective inhibitor development.

In light of this structural data, one of the more common methods employed recently for identifying PTP1B inhibitors is a linked fragment approach. This strategy takes advantage of the discovery of the non-catalytic aryl phosphate binding pocket. Inhibitors which are able to bridge the gap between these two sites and bind strongly to each pocket are expected to demonstrate improved affinity and specificity. Szczepankiewicz and coworkers reported on the successful application of this approach in 2003.<sup>121</sup> The approach was straightforward: combinatorial libraries of small molecules containing moieties isosteric to phosphotyrosine were screened for PTP1B active site binding via two-dimensional <sup>1</sup>H-<sup>13</sup>C and <sup>1</sup>H-<sup>15</sup>N HSQC NMR titration experiments. For the purposes of this investigation, a shortened 292-residue version of the PTP1B protein was used. PTP1B was either labeled entirely with <sup>15</sup>N (prepared via expression in *E. coli* grown in <sup>15</sup>N-enriched NH<sub>4</sub>Cl) or the side chain δ-methyl groups of Ile were labeled with <sup>13</sup>C (via expression in *E. coli* to which a cocktail of <sup>13</sup>C-γ-methionine, 3,3'-<sup>13</sup>C-α-ketovalerate and 3-<sup>13</sup>C-α-ketobutyrate was administered prior to induction). Using PTP1B labeled in this way, binding to pTyr surrogates at the active site was monitored by observation of changes in the chemical shifts of active site cross peaks upon titration with potential inhibitors. Inhibition of PTP1B activity was also monitored via *in vitro* enzymatic competition assays in which PTP1B and inhibitors were incubated with *p*-nitrophenylphosphate (pNPP), a substrate which yields the strong chromophore *p*-nitrophenol following cleavage.

Using this design, Szczepankiewicz reported the discovery of a diaryloxamic acid that binds to the PTP1B active site and induces chemical shift changes similar to those effected by pTyr. In order to optimize this hit ( $K_d = 100 \mu\text{M}$ ,  $K_i = 293 \mu\text{M}$ ), the lead compound was further embellished via installation of a naphthyloxamic acid functionality so as to increase the contact surface area at protein surface. This was indeed the case, as this naphthyloxamate exhibited a lower  $K_d$  in NMR titration experiments ( $26 \mu\text{M}$ ) and a lower  $K_i$  ( $39 \mu\text{M}$ ) than 1. Fitting of kinetic data to Lineweaver-Burk plots revealed that both compounds function as reversible competitive inhibitors of PTP1B.

A crystal structure of naphthyloxamic acid inhibitor bound to PTP1B revealed that it binds the protein active site in a similar fashion to pTyr, as expected from HSQC titration data. The objective then became the installation of a linker capable of reaching from this molecule out of the enzyme active site and towards the adjacent non-catalytic pTyr-binding site. To this end, compound 3 (see table above) was synthesized by the installation of a diamide-containing chain at the 4-naphthyl position of 2. Though it was hypothesized that this chain could extend out of the active site and provide a linkage to the second pTyr-binding site, it was also found that this linker imparted additional inhibitory activity ( $K_i = 1.1 \mu\text{M}$ ). X-ray data revealed evidence that this is due to favorable hydrogen bonds made between the diamide function and the side chain carboxylate of Asp48.

**Table 4.1 Representative PTP1B Inhibitors**

Entry	Inhibitor	EC <sub>50</sub>	K <sub>i</sub>	Ref
1			293 μM	118
2			1.1 μM	118
3			22 nM	118
4		R1 = H, R2 = H	18 nM	119
5		R1 = Me, R2 = Et		

6		25 $\mu$ M	120
7		2.4 nM	121
8		12 $\mu$ M	122

\*Note: EC<sub>50</sub> determined in whole cell insulin receptor phosphorylation assays

\*\*Note: IC<sub>50</sub> values determined in enzymatic p-nitrophenylphosphate assays

Table 4.1 provides a sample of various small molecule PTP1B inhibitors from the literature. Examples include species containing substructures that are isoelectronic with phosphotyrosine (entries 1-3), esterified prodrugs (entries 5-6), as well as non-hydrolyzable phosphotyrosine isosteres (entry 7). None of these molecules have entered clinical trials. The most recent PTP1B-related therapeutic to have entered clinical trials was an anti-sense siRNA oligonucleotide, which dropped out at phase II due to efficacy.<sup>113</sup>

Moreover, this structure revealed that the pentyl chain extends from the active site and points directly towards the second pTyr-binding site.

In order to complete the linked-fragment approach, similar HSQC screens were performed on the 10,000-molecule combinatorial library, now with an aim at finding ligands for the second aryl phosphate binding site. Initial hits suggested that a small naphthoic acid might bind with modest affinity ( $K_d$  values for site 2 hits generally exceeded 1 mM). These hits were primarily small aromatic acids consisting of fused rings. As such, methyl 3-hydroxy-2-naphthoate was suggested as a compound that could be readily conjugated to the existing inhibitor skeleton.

Attachment of the naphthoic acid to the inhibitor and diamido linker yielded an inhibitor of 50-fold greater potency in enzymatic inhibition experiments ( $K_i = 22$  nM). In addition to showing improved potency, the selectivity of this species was demonstrated by comparing the empirical  $K_i$  values of the synthesized inhibitors for PTP1B with those obtained against other common PTPs. As expected, inhibitors 1, 2 and 3 demonstrated little selectivity against a panel of tyrosine phosphatases (TCPTP, LAR, SHP-2, CD45). Compound 5, however, exhibited marked selectivity (>30-fold) against all of the PTPs studied, save for the T-cell tyrosine phosphatase (TCPTP), for which 2-fold selectivity was observed. The use of TCPTP is the gold standard for measuring PTP1B specificity, as the two proteins are greater than 77% identical.

The previous case study by Szczepankiewicz represents a seminal effort in the development of PTP1B inhibitors, and established the so-called SAR-by-NMR approach to optimizing linked fragments. Shortly thereafter, Liu and coworkers at Abbott Laboratories reported on the application of this method to develop inhibitors of not only improved potency, but that also exhibit cell permeability and inhibitory effects in HepG2 cell culture.<sup>122,123</sup> By applying the NMR screen of a small molecule library for affinity to the PTP1B active site and the non-catalytic aryl phosphate binding site and installing an appropriate linker, compound 4 was synthesized and exhibited a  $K_i$  of 18 nM. This inhibitor also demonstrated a 4-fold selectivity against TCPTP. However, this compound failed to exhibit substantial cell permeability, presumably due to the unit negative charges that the carboxylate functions bear at physiological pH. One strategy to overcoming this challenge is to synthesize methyl or ethyl esters to mask the carboxylic acid functionality. One such prodrug, compound 5, has been shown to exhibit cell permeability and inhibitory activity. Liu and coworkers also applied this strategy to the discovery of compound 6, which exhibits cell permeability and a 30-fold selectivity against TCPTP. Interestingly, it appears the *ortho*-fluorine is necessary for traversing the lipid bilayer, as the substitution of this fluorine for hydrogen (9) results in loss of cell permeability.

In addition to the SAR-by-NMR methodology, small molecules have also been screened in high-throughput fashion via traditional ELISA-based assays. In 2001, Zhang and coworkers<sup>116</sup> reported on the optimization of this assay, which systematically screens terminal elements in combination with a variety of linkers for the ability to inhibit binding of a PTP1B C215S mutant to the biotinylated

peptide DADEpYL-NH<sub>2</sub> immobilized on Neutravidin-coated microtiter plates.<sup>18</sup>

One key difference between the library used for this screen and those used in the previously described investigations is that the compounds in this library contained aryl phosphate functionalities. Specifically, the library under study was composed of a phosphotyrosine moiety tethered to one of 8 diverse aryl acids via one of 23 different linker elements. Using this screening approach, the compound which exhibited the greatest affinity was identified, and a  $K_d = 30$  nM was measured via equilibrium dialysis methods. The non-hydrolyzable bisdifluorophosphonate analog of this compound was synthesized, furnishing a potent competitive inhibitor of PTP1B activity ( $K_i = 2.4$  nM in *p*NPP enzymatic cleavage assays).

In addition to potency, compound this phosphotyrosine isostere exhibited 10-fold selectivity against TCPTP. Structural data produced from X-ray crystallography studies revealed the binding of this compound to PTP1B is indeed mediated by intermolecular interactions within the active site as well as an adjacent non-catalytic site. It was reported that the C-terminal pTyr mimic in the above molecule engages in extensive ionic interactions and Hydrogen bonding with the PTP1B active site, while the N-terminal surrogate binds to a neighboring non-catalytic pocket and forms hydrophobic and polar interactions with Lys41, Arg47 and Asp48 at this site.

Another approach employed for obtaining PTP1B inhibitors is the rational design of pTyr mimics strategically engineered to interact with the PTP1B active and an adjacent non-catalytic site in bidentate fashion. This approach is made possible

by the abundance of crystal structure data available for PTP1B and its complexes with known inhibitors. Once again the aim here is to improve potency and selectivity, and a variety of successful scaffolds have been utilized to accomplish this end. For instance, Zhang and coworkers have recently explored the application of an aryl diketoacid scaffold for targeting PTP1B in response to the success this scaffold has exhibited in targeting HIV-1 integrase.<sup>115</sup> By starting with this general model and optimizing substituents about the aromatic system, several inhibitors of micromolar  $K_i$  were reported, including compound 8 (table 4.1). X-ray data reveal this compound binds the active site and an adjacent binding pocket, and validates aryl ketoacids as potential pTyr surrogates. Fitting of enzymatic data from *p*NPP hydrolysis studies to Lineweaver-Burk plots revealed this inhibitor is indeed functioning as a non-competitive inhibitor, consistent with X-ray data that confirms binding to a non-catalytic site in addition to the active site. In this fashion, the enzyme is stabilized in a conformation in which the WPD loop is held in an open, catalytically-inert state.

In consideration of the state of the art of PTP1B inhibition, we recognize that the challenge of generating selective inhibitors against this phosphatase is indeed a pressing and historically intractable task in light of the high degree of sequence and structural homology with other phosphatases, particularly TCPTP. Currently, the PTP1B antagonist that features the highest specificity over TCPTP is 10-fold selective. There are currently no anti-PTP1B therapeutic candidates in clinical trials. We propose that cyclic and bicyclic peptides may succeed where small molecules have failed by harnessing a larger topology in order to establish contacts with peripheral sites on the PTP1B surface that will enable peptidic

inhibitors to discriminate one phosphatase from another, despite conservation at the phosphotyrosine-binding site. Moreover, the introduction of successive conformational constraints to such an inhibitor may further serve to increase potency and selectivity by stabilizing a single bioactive conformation and reducing the conformational flexibility necessary to bind multiple targets. Our structural pre-organization platform has already been shown to effectively engender a phosphotyrosine-like surface in the presence of surrounding unrelated chemical functionality. In addition, we have also shown that our cyclic and bicyclic scaffolds are uniquely cell-penetrant due to the presence of distinct water-binding sites on the peptide surface that sequester hydrogen-bond donors and acceptors and shield these atoms from bulk solvent. We hypothesized that replacement of certain chemical moieties in the **G1HT1** or **BC1** scaffold that are not essential for the presentation of a phosphotyrosine-like epitope with residues that have been shown to favor PTP1B binding may represent a general strategy for transforming our Grb2-SH2 domain inhibitors to modulators of another medically relevant phosphotyrosine-binding protein. The sections that follow will detail our efforts to discover novel inhibitors of PTP1B via rational design as well as through the screening of synthetic peptide libraries.

#### **4.2 – Engineering Cyclic Peptide PTP1B Inhibitors Based on G1HT1**

Previous SAR data from our work and from other investigators demonstrate that the ELY motif in the **G1** and **G1HT1** scaffold is critical for binding the pTyr-

binding site on the Grb2-SH2 domain. However, many of the remaining residues are not explicitly necessary for pTyr surrogacy. We proposed replacement of some of these residues with amino acids bearing side chain functionalities that have been shown to promote PTP1B binding. This strategy was particularly attractive in light of the wealth of structural information and consensus peptide sequence data that is available for this enzyme. Work by Dehua Pei and others<sup>114</sup> over the last decade and a half examining the functional preferences of peptidic PTP1B ligands provided a strong foundation for this project, as much information has been elucidated regarding the nature of PTP1B consensus sequences. Unlike Grb2-SH2 domain peptide ligands in which the specificity-determining sequence is located C-terminal to the pTyr residue, reported PTP1B-binding consensus sequences typically feature selectivity elements N-terminal to pTyr. Several high-affinity consensus sequences were selected as a starting point for inhibitor design. Among the reported consensus sequences, we selected several that exhibit various physicochemical properties. Included in this series are the AKFEDpY, DADEpY, and KAVFIpY sequences. We designed a series of peptides based on **G1HT1** that integrate these sequences N-terminal to the pTyr-mimicking ELY epitope. While we screened peptides containing each of these consensus sequences, we chose to focus our inhibitor design efforts on peptides containing the KAVFI sequence. This decision was based on the higher lipophilic character of these residues, as well as the absence of local negative charges, which have been shown to often be poor indicators of cell penetration.

The results of our screens revealed several interesting results. From the results of the AKFED- and DADE-containing peptides, it is clear that while **AKFED1**

exhibits double digit-micromolar inhibitory activity in a *p*-nitrophenylphosphate-based enzymatic assay ( $IC_{50} = 29 \mu\text{M}$ , see Materials and Methods), **DADE1** shows a dramatic reduction in inhibitory potency ( $IC_{50} > 100 \mu\text{M}$ , figure 4.1). This disparity in activity in light of the similar physicochemical nature of these peptides – both terminate with consecutive anionic residues N-terminal to the ELY sequence – was gratifying, as it demonstrated that the presence of several local negative charges adjacent to the ELY epitope is insufficient for binding activity, and provided the first evidence in our hands that a distinct orientation and neighboring functionality was necessary for a competent PTP1B pharmacophore.

The AVFI-containing series of cyclic peptides revealed further information regarding the utility of the **G1HT1** scaffold in targeting this enzyme. Peptides **AVFI1-4** represent a systematic variation of several residues external to the ELY motif that have been shown to be non-essential for pTyr-mimicry. However, we recognize that these residues likely play important structural roles in stabilizing a bioactive conformation of the peptide macrocycle. It is for this reason that we sought to minimally perturb the **G1HT1** sequence while introducing the AVFI consensus sequence. We varied the identity of amino acid AA4 within the parent peptide sequence, as our NMR data indicated Pro4 likely stabilizes a turn motif in this scaffold. However, we had no reason to believe a turn element at this site is desirable for bioactivity, as this was a unique feature of Grb2-SH2 domain ligands. We also varied the identity of amino acid AA6 in our designed peptides between methionine and lysine. This was done in order to probe the physicochemical preference of this position, as we sought to introduce Lys as part of the AVFI consensus sequence, and Met provides a convenient, nearly

isosteric control for Lys with reduced polarity and no ionic charge, reminiscent of the electronic environment of the parent peptide **G1HT1** at the distal end of the molecule.

The results of this initial screen revealed that while **AVFI1**, **2**, and **4** show very mild anti-PTP1B activities, **AVFI3** exhibits an  $IC_{50}$  of 37  $\mu$ M. This inhibitory activity is only about 1-2 orders of magnitude lower than the activities exhibited by PTP1B inhibitors of similar charge, and given the propensity of this scaffold for cellular internalization, we reasoned that **AVFI3** represents an ideal starting point for further derivitization and optimization. The results of subsequent analysis of cyclic peptides based on **AVFI3** revealed that inhibition is indeed predicated on the presence of the pTyr-mimicking ELY element, yet another gratifying result that helps to rule out the possibility of non-specific binding due to the hydrophobic nature of the AVFI region of the macrocycle. The results of screening still other **AVFI3** derivatives indicated a conformation-dependant inhibition response, providing further evidence that the stabilization of a distinct

<b>Analog</b>	<b>Sequence</b>
AVFI1	ELYANMGAVFI
AVFI2	ELYPNMGAVFI
AVFI3	ELYANKGAVFI
AVFI4	ELYPNKGAVFI
AVFI3.1	ELYANK $\beta$ AVFI
AVFI3.2	ELYANK $\gamma$ AVFI
AVFI3.3	ELYANKGGVFI
AVFI3.4	ALYANKGAVFI
AKFED1	ELYANGAKFED
DADE1	ELYANGGDADE

Table 4.2. Designed cyclic peptides based on **G1HT1** to target the active site of PTP1B. (**G1HT1** sequence: cyclo-ELYPNVGMVGA)

bioactive conformation of this scaffold will further enhance target affinity and hence inhibitory activity.

Peptide	Sequence	IC50 ( $\mu\text{M}$ )
AVFI1	ELYANMGAVFI	> 100
AVFI2	ELYPNMGAVFI	> 100
AVFI3	ELYANKGAVFI	37
AVFI4	ELYPNKGAVFI	> 100
AVFI3.1	ELYANK $\beta$ AVFI	> 100
AVFI3.2	ELYANK $\gamma$ AVFI	> 400
AVFI3.3	ELYANKGGVFI	> 50
AVFI3.4	ALYANKGAVFI	no inhibition
AKFED1	ELYANGAKFED	29
DADE1	ELYANGGDADE	> 100

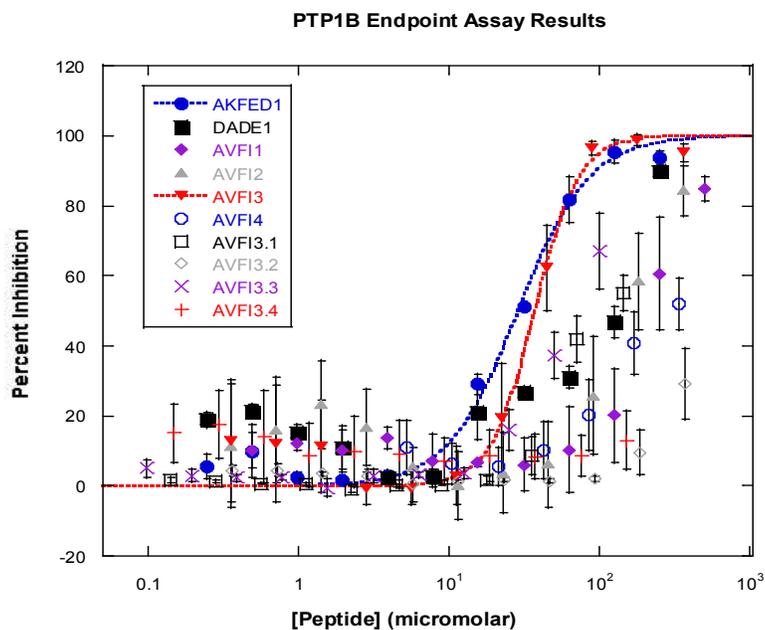
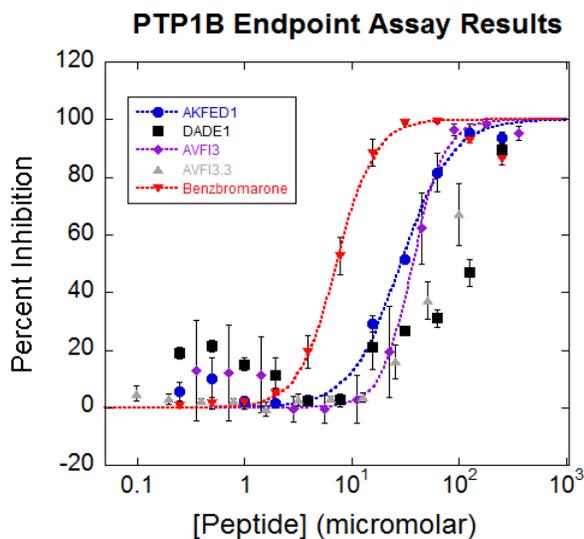


Figure 4.1. Results of endpoint competition assay for **G1HT1**-derived cyclic peptides incorporating the AVFI, AKFED or DADE PTP1B-binding consensus sequences. The small molecule inhibitor benzbromarone was used included as a positive control (see Materials and Methods). For clarity, negative peptides are omitted in one graph (top) and included for completeness in the other (bottom).

## **Materials and Methods**

### *Cyclic Peptide Synthesis and Purification*

Cyclic peptides were synthesized entirely on solid support using Wang resin loaded with orthogonally-protected Fmoc-Glu-ODmab. The synthetic scheme outlined previously (Materials and Methods, Chapter One) was used for the synthesis of PTP1B-targeting peptides. Peptides were purified by HPLC as previously described.

### *Enzymatic Assay*

Inhibition of PTP1B activity was monitored using a competitive endpoint assay essentially as described.<sup>121</sup> PTP1B (80 nM) was incubated with p-nitrophenylphosphate (10 mM) and various concentrations of peptide inhibitor in 10 mM Tris-HCl, pH = 7.6, 150 mM NaCl in flat, black-bottom 96-well plates (Corning). Plates were incubated at 37° C for 45 minutes. Absorbance at 405 nm was subsequently determined using a Tecan Infinite 200 plate reader.

Peptide	Calculated MW	Observed m/z	Species
AKFED1	1238.3	1237.1	[M-H]
DADE1	1135.1	1134.3	[M-H]
AVFI1	1209.4	1208.2	[M-H]
AVFI2	1235.4	1234.7	[M-H]
AVFI3	1206.4	1205.3	[M-H]
AVFI4	1232.4	1231.2	[M-H]
AVFI3.1	1218.4	1217.6	[M-H]
AVFI3.2	1230.4	1229.2	[M-H]
AVFI3.3	1194.4	1193.2	[M-H]
AVFI3.4	1148.3	1147.2	[M-H]

Table 4.3. Characterization of **G1HT1**-derived cyclic peptides to target the active site of PTP1B.

#### 4.3 – Applying Cysteine Bis-alkylation to the Development of Bicyclic Peptide Inhibitors of PTP1B

In order to install conformational constraints within the **AVFI3** scaffold, we began adopted an intramolecular cross-linking approach that has been used by a variety of independent investigators in order to constrain structurally diverse peptide scaffolds. The strategy is relatively straightforward, as cysteine residues are substituted for amino acids that are spatially proximal to one another. These nucleophilic moieties are then cross-linked with bifunctional electrophilic species that span an array of sizes, geometries and that impose different conformational preferences on the resulting tethered molecule. This chemistry has been successfully applied to the stapling of  $\alpha$ -helical peptides, as well as to stabilizing non-canonical turns in large (>30 amino acid) peptide sequences.<sup>34,85,86</sup> We reasoned that this chemical cross-linking strategy represents an improvement over our previous lactam- and olefin-based bicyclization approaches, as this

platform enables the rapid sampling of a large conformational space by taking advantage of a late-stage divergent synthetic scheme. We proposed that by synthesizing just one monocyclic peptide precursor, we can access a multitude of structurally distinct bicyclic derivatives by virtue of the abundance of aromatic and aliphatic bifunctional electrophiles that are already commercially available. Moreover, the kinetics of these cross-linking reactions are such that these processes occur rapidly (bicyclization of **AVFI3** in this fashion often required less than 2 hours at room temperature) and can be carried out on unprotected peptides due to the high nucleophilicity of cysteine thiolates. This is in contrast to our previous approaches, which required the installation of expensive olefinic amino acids or Lys/Asp residues that can only yield one final product upon lactam formation and that often required extended reaction times (24-72 hours) or complex orthogonal protection schemes.

We applied this cross-linking strategy to **AVFI3**, and the results are shown in figure 4.2. Unlike some of the previously-reported peptides that have been cross-linked in this fashion, **AVFI3** is already a constrained macrocycle, and so we reasoned that not all of the linkers we sampled would successfully cross-link this peptide. Energy minimization of the **AVFI3** peptide (Molecular Operating Environment software package, Chemical Computing Group) suggested amino acid positions external to the ELY sequence that contain spatially proximal side chains. These residues were replaced with cysteine and subjected to a variety of bis(bromomethyl)-containing electrophilic linkers in order to screen for linkers capable of cross-linking the peptide. The peptide was treated with various linkers and reaction progress was monitored by MALDI-TOF mass spectrometry at

distinct intervals. Reactions were run at room temperature instead of elevated temperatures in order to favor kinetic control of reactivity and to prevent the possibility of cross-linking the peptide so as to stabilize a biologically inactive conformation. Of the aromatic and aliphatic electrophiles sampled, the xylene and pyridine derivatives were the only linkers that induced successful bicyclization of the scaffold.

Upon screening the bicyclic peptides listed in table 4.5 in our endpoint phosphatase activity assay, a trend similar to the Grb2-SH2-targeting bicyclic peptides was observed: seemingly small structural changes resulted in dramatic changes in inhibitory activity. Of the bicyclic **AVFI3** derivatives sampled, one of the compounds, **BC-AVFI3.2**, exhibited an  $IC_{50}$  of 18  $\mu$ M. This represents a 2-fold improvement in affinity over the cyclic precursor peptide, and is within the affinity regime for which cellular phenotypes of PTP1B inhibitors have been reported. Moreover, this peptide is interesting from a structural perspective. Unlike previous bicyclic peptides discussed that target the Grb2-SH2 domain feature cross-linkers that serve only to stabilize the structure of the macrocycle, the aromatic cross-linking moiety in **BC-AVFI3.2** represents an opportunity to further engage the PTP1B surface, either through pi-pi interactions with the adjacent aryl phosphate binding site or through more general Van der Waals contacts with hydrophobic residues peripheral to the enzyme active site. As the following section will discuss, we are actively following up on this compound in order to determine the structural attributes of this unique scaffold that promote PTP1B inhibition. We are also anxious to test this compound in mammalian cell models of PTP1B misregulation (HepG2 hepatocytes) in order to determine the potential

effect of this peptide on rescuing the phosphorylation of the insulin receptor and other known PTP1B substrate proteins.

Table 4.5 AVF13 (Cys) Sequence: cyclo(ELYCNKGACFI)

Peptide	Calculated MW	Observed m/z: 1 hour	Observed m/z: 2 hour	Observed m/z: 24 hour	Species
AVF13 (Cys)	1242.4				[M-H]
+ $\alpha,\alpha'$ -Dibromo- <i>o</i> -xylene	1344.6	1242.4	1242.4, 1344.6	1344.6	[M-H]
+ $\alpha,\alpha'$ -Dibromo- <i>m</i> -xylene	1344.6	1242.4, 1344.6	1344.6	1344.6	[M-H]
+ $\alpha,\alpha'$ -Dibromo- <i>p</i> -xylene	1344.6	1242.4	1242.4, 1344.6	1242.4, 1344.6	[M-H]
+ 2,6-Bis(bromomethyl)pyridine	1345.6	1344.6	1344.6	1344.6	[M-H]
+ 1,8-Bis(bromomethyl)naphthalene	1394.62	1242.4	1242.4	1242.4	[M-H]
+ 2,6-Bis(bromomethyl)naphthalene	1394.62	1242.4	1242.4	1242.4	[M-H]
+ 2,7-Bis(bromomethyl)naphthalene	1394.62	1242.4	1242.4	1242.4	[M-H]
+ 4,4'-Bis(bromomethyl)biphenyl	1420.65	1242.4	1242.4	1242.4	[M-H]
+ 4,4''-Bis(Bromomethyl)-(1,1':3,1'')terphenyl	1496.76	1242.4	1242.4	1242.4	[M-H]
+ 1,2-dibromobutane	1264.4	1242.4	1242.4	1242.4	[M-H]
+ 1,2-dibromoethane	1288.4	1242.4	1242.4	1242.4	[M-H]

### PTP1B Endpoint Competition Assay

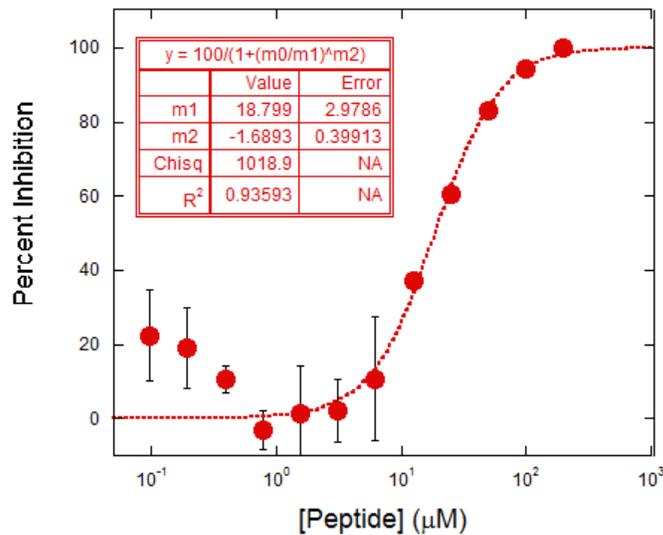
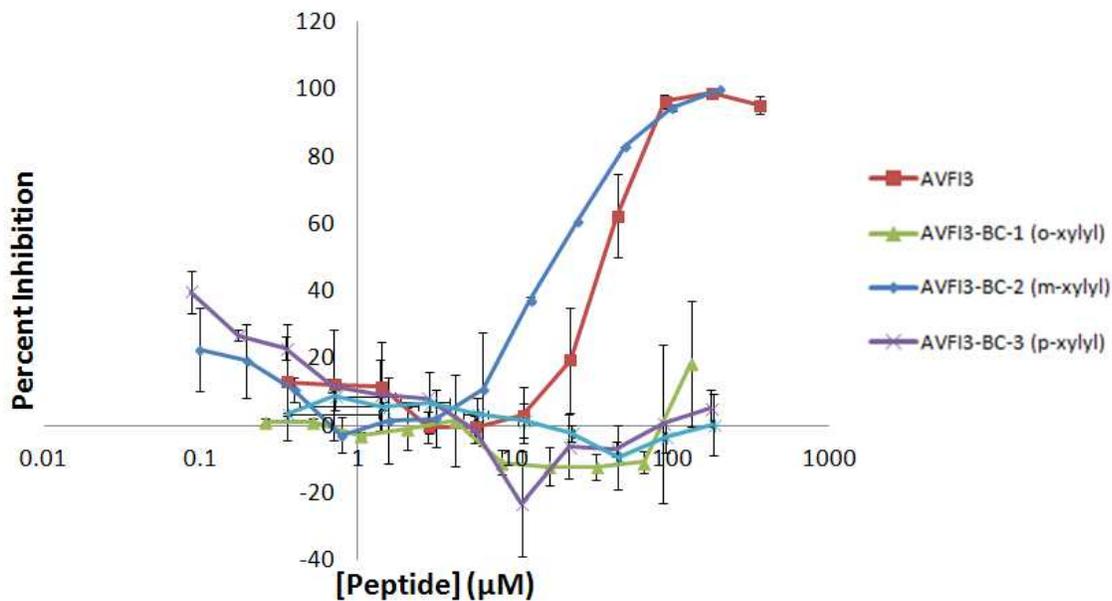


Figure 4.2. Results of competitive endpoint assay for bicyclic **AVF13** derivatives. *Top*: As has been observed in previous rounds of inhibitor design, seemingly subtle structural changes lead to large changes in target affinity in this scaffold. *Bottom*: Non-linear regression analysis was performed for **AVF13-BC-2**, which exhibited a 2-fold improvement in potency relative to the parent peptide (IC<sub>50</sub>=18 μM vs. 37 μM).

## **4.4 – Future Directions: Structural Investigations and Cellular Activities of Cyclic and Bicyclic PTP1B Inhibitors**

### *4.4.1 – Structure-based design of PTP1B inhibitors.*

As was the case for the design of our Grb2 inhibitors, our lab plans to transition this project from a model-based, iterative design strategy at its onset to an approach that is informed by detailed structural information. Many co-crystal structures of PTP1B and synthetic ligands have been reported, and we expect that PTP1B will be readily crystallized in our hands.<sup>39-41</sup> We are currently working in collaboration with Prof. Andrew Bohm to solve crystal structures of PTP1B in complex with cyclic and bicyclic peptide inhibitors. Furthermore, we are actively working to solve the NMR solution structure of our PTP1B inhibitors. Together, this atomic-level structural information will guide the design of bicyclic peptides that inhibit PTP1B and allow us to correlate structural features with desirable *in vitro* and cellular activity.

### *4.4.2 – Evaluate the cell-penetrating capabilities and cellular activities of our PTP1B inhibitors.*

In order to monitor the cell-penetrating properties of PTP1B inhibitors, fluorescently labeled variants of these peptides will be synthesized and incubated with insulin-responsive HepG2 hepatocytes and 3T3-L1 adipocytes. Many previously reported PTP1B inhibitors, including those with significant anionic character, have been shown to penetrate these cell types, likely via an anionic

transport mechanism.<sup>42-43</sup> Following incubation, cells will be imaged by confocal fluorescent microscopy and, in a parallel assay, will be lysed and subjected to fluorometric analysis (as was performed for Grb2-SH2 domain inhibitors, Chapter Two). Furthermore, to probe the specificity of our inhibitors, we will conjugate variants of these peptides to agarose beads and perform pull down experiments using HepG2 and 3T3-L1 lysates. We will then perform Western blots against PTP1B and homologous PTPs, including SHP1, SHP2 and TCPTP (approximately 70% homologous). In addition, we will test the inhibitory activity of cell-penetrant inhibitors by assessing their affect on insulin signaling. Inhibitors will be incubated with HepG2 and 3T3-L1 cells, both in the absence and presence of insulin, for set time intervals. Cells will then be lysed and Western blots will be conducted to effect of bicyclic PTP1B inhibitors on phosphorylated insulin receptor and IRS-1, validated substrates of PTP1B.<sup>42-44</sup>

#### *4.4.3 – Outlook*

The work presented in this dissertation represents a long-term effort to establish a paradigm for transforming unmodified peptides into constrained compounds capable of disrupting clinically-relevant protein-protein interactions and that feature more promising pharmacological profiles than their conformationally flexible precursors. Specifically, the target proteins of interest in this work have been phosphotyrosine-binding proteins, the misregulation of which has been well-documented in various human pathologies. We have demonstrated that carefully applying conformational constraints to an established PPI antagonist using rational inhibitor design can lead to bicyclic peptides with improved target

affinity, selectivity, protease resistivity and potentially improved bioavailability, four important barometers for gauging the likelihood of success in pre-clinical studies. It is our vision that the molecules reported in this dissertation and the strategies employed to discover these compounds can be starting points for the development of constrained peptide inhibitors of a broad array of protein-protein interactions, especially those mediated by phosphotyrosine. Subsequent structural data, obtained via X-ray crystallography and/or NMR analyses, will inform the design of improved optimized inhibitors that will then be evaluated in live cells. The work discussed in this dissertation provides a series of standard operating procedures and several active collaborations with experts in structural biology that will aid in executing this project strategy.

The use of bicyclic peptides in inhibitor development has grown in popularity over the last five years, as other labs have also begun establishing screening methods for mining this unique region of chemical space. As the field advances and the concept of the “druggable” proportion of the proteome expands, it is evident that constrained cyclic and bicyclic peptides will constitute an important part of the movement to target protein-protein interactions that have been historically intractable. The novel aspects of rational bicyclic inhibitor design and characterizing distinct structural elements that favor enhanced target affinity and cellular internalization in the scaffolds discussed herein provide a series of general rules and observations to aid in the discovery of new, bioactive bicyclic peptides. The research platform established as a result of this work provides another tool for the field of non-traditional drug discovery and development, and we look forward to the application of this strategy in future settings and the

development of new, promising molecules that target problematic areas of human biology.

Appendix 1: Crystal parameters of A6 and C2 crystals of **BC1**-Grb2-SH2 domain complexes

**X-ray Diffraction Data Collection:**

**MCSG1 A6 and C2 Crystals of Grb2-SH2:BC1 Complex**

	<b>A6 Crystal</b>	<b>C2 Crystal</b>
<b>Data Collection</b>		
X-ray source	NLS	NLS
Space group	P3 <sub>1</sub> 21	P3 <sub>1</sub> 21
Unit cell dimensions (Å)	a=93.4, b=93.4, c=83.5	a=96.5, b=96.5, c=96.3
Resolution (Å)	58-2.8	28-3.2
Unique reflections	3,455	12,724
Completeness (%)	91.3	93.5
Redundancy	8.1	11.2

## References

1. Makley, L. N. & Gestwicki, J. E. Expanding the Number of 'Druggable' Targets: Non-Enzymes and Protein-Protein Interactions: **Expanding the Number of 'Druggable' Targets**. *Chem. Biol. Drug Des.* **81**, 22–32 (2013).
2. Smith, M. C. & Gestwicki, J. E. Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev. Mol. Med.* **14**, e16 (2012).
3. Berg, T. Small-molecule inhibitors of protein-protein interactions. *Curr. Opin. Drug Discov. Devel.* **11**, 666–674 (2008).
4. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).
5. Machida, K. & Mayer, B. J. The SH2 domain: versatile signaling module and pharmaceutical target. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1747**, 1–25 (2005).
6. Combs, A. P. Recent Advances in the Discovery of Competitive Protein Tyrosine Phosphatase 1B Inhibitors for the Treatment of Diabetes, Obesity, and Cancer. *J. Med. Chem.* **53**, 2333–2344 (2010).
7. Zhang, S. & Zhang, Z. PTP1B as a drug target: recent developments in PTP1B inhibitor discovery. *Drug Discov. Today* **12**, 373–381 (2007).
8. Yue, P. Targeting STAT3 in cancer: how successful are we? *Expert Opin. Investig. Drugs* **18**, 45–56 (2009).
9. Burke, T. R. Development of Grb2 SH2 Domain Signaling Antagonists: A Potential New Class of Antiproliferative Agents. *Int. J. Pept. Res. Ther.* **12**, 33–48 (2006).
10. Walensky, L. D. Activation of Apoptosis in Vivo by a Hydrocarbon-Stapled BH3 Helix. *Science* **305**, 1466–1470 (2004).
11. Bernal, F., Tyler, A. F., Korsmeyer, S. J., Walensky, L. D. & Verdine, G. L. Reactivation of the p53 Tumor Suppressor Pathway by a Stapled p53 Peptide. *J. Am. Chem. Soc.* **129**, 2456–2457 (2007).
12. Walensky, L. D. *et al.* A Stapled BID BH3 Helix Directly Binds and Activates BAX. *Mol. Cell* **24**, 199–210 (2006).
13. Walensky, L. D. & Bird, G. H. Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress: Miniperspective. *J. Med. Chem.* 140306151941006 (2014). doi:10.1021/jm4011675
14. Lian, W., Upadhyaya, P., Rhodes, C. A., Liu, Y. & Pei, D. Screening Bicyclic Peptide Libraries for Protein–Protein Interaction Inhibitors: Discovery of a

- Tumor Necrosis Factor- $\alpha$  Antagonist. *J. Am. Chem. Soc.* **135**, 11990–11995 (2013).
15. Kritzer, J. Stapled Peptides: Magic bullets in nature's arsenal. *Nat. Chem. Biol.* **6**, 566–567 (2010).
  16. Rezai, T. *et al.* Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J. Am. Chem. Soc.* **128**, 14073–14080 (2006).
  17. Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P. & Lokey, R. S. Testing the Conformational Hypothesis of Passive Membrane Permeability Using Synthetic Cyclic Peptide Diastereomers. *J. Am. Chem. Soc.* **128**, 2510–2511 (2006).
  18. Adessi, C. & Soto, C. Converting a peptide into a drug: Strategies to improve stability and bioavailability. *Curr. Med. Chem.* **9**, 963–978 (2002).
  19. Ovadia, O. *et al.* Improvement of drug-like properties of peptides: the somatostatin paradigm. *Expert Opin. Drug Discov.* **5**, 655–671 (2010).
  20. Kim, Y.-W., Grossmann, T. N. & Verdine, G. L. Synthesis of all-hydrocarbon stapled  $\alpha$ -helical peptides by ring-closing olefin metathesis. *Nat. Protoc.* **6**, 761–771 (2011).
  21. Verdine, G. L. & Hilinski, G. J. in *Methods in Enzymology* **503**, 3–33 (Elsevier, 2012).
  22. Frank, A. O. *et al.* Discovery of a Potent Stapled Helix Peptide That Binds to the 70N Domain of Replication Protein A. *J. Med. Chem.* 140219084536007 (2014). doi:10.1021/jm401730y
  23. Joullie, M. & Richard, D. Cyclopeptide alkaloids: chemistry and biology. *Chem. Commun.* **18**, 2011–2015 (2004).
  24. Temal-Laib, T., Chastanet, J. & Zhu, J. A convergent approach to cyclopeptide alkaloids: Total synthesis of Sanjoinine G1. *J. Am. Chem. Soc.* **124**, 583–590 (2002).
  25. Craik, D., Daly, N., Mulvenna, J., Plan, M. & Trabi, M. Discovery, structure and biological activities of the cyclotides. *Curr. Protein Pept. Sci.* **5**, 294–315 (2004).
  26. Tam, J., Lu, Y., Yang, J. & Chiu, K. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci.* **96**, 8913–8918 (1999).
  27. Tan, N. & Zhou, J. Plant cyclopeptides. *Chem. Rev.* **106**, 840–895 (2006).

28. Kobayashi, J., Suzuki, H., Shimbo, K., Takeya, K. & Morita, H. Celogentins A-C, new antimitotic bicyclic peptides from the seeds of *Celosia argentea*. *J. Org. Chem.* **66**, 6626–6633 (2001).
29. Lindell, T. J., Weinberg, F., Morris, P. W., Roeder, R. G. & Rutter, W. J. Specific inhibition of nuclear RNA polymerase II by  $\alpha$ -amanitin. *Science* **170**, 447–449 (1970).
30. Craik, D., Cemazar, M. & Daly, N. The cyclotides and related macrocyclic peptides as scaffolds in drug design. *Curr. Opin. Drug Discov. Devel.* **9**, 251–260 (2006).
31. Craik, D., Simonsen, S. & Daly, N. The cyclotides: Novel macrocyclic peptides as scaffolds in drug design. *Curr. Opin. Drug Discov. Devel.* **5**, 251–260
32. Valkov, E. Targeting Protein-Protein Interactions and Fragment-Based Drug Discovery. *Fragm.-BASED DRUG Discov. X-RAY Crystallogr.* **317**, 145–179 (2012).
33. Spokoiny, A. M. *et al.* A Perfluoroaryl-Cysteine S<sub>N</sub>Ar Chemistry Approach to Unprotected Peptide Stapling. *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).
34. Angelini, A. *et al.* Bicyclic Peptide Inhibitor Reveals Large Contact Interface with a Protease Target. *ACS Chem. Biol.* **7**, 817–821 (2012).
35. Baeriswyl, V. & Heinis, C. Polycyclic Peptide Therapeutics. *ChemMedChem* **8**, 377–384 (2013).
36. Chen, S. *et al.* Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries. *J. Am. Chem. Soc.* **135**, 6562–6569 (2013).
37. Korsinczky, M. L. *et al.* Solution structures by <sup>1</sup>H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J. Mol. Biol.* **311**, 579–591 (2001).
38. Zabłotna, E. *et al.* Chemical Synthesis and Kinetic Study of the Smallest Naturally Occurring Trypsin Inhibitor SFTI-1 Isolated from Sunflower Seeds and Its Analogues. *Biochem. Biophys. Res. Commun.* **292**, 855–859 (2002).
39. Jiang, S. *et al.* Design and Synthesis of Redox Stable Analogues of Sunflower Trypsin Inhibitors (SFTI-1) on Solid Support, Potent Inhibitors of Matriptase. *Org. Lett.* **9**, 9–12 (2007).
40. Li, P. *et al.* Design and Synthesis of Novel and Potent Inhibitors of the Type II Transmembrane Serine Protease, Matriptase, Based upon the Sunflower Trypsin Inhibitor-1. *J. Med. Chem.* **50**, 5976–5983 (2007).

41. Marx, U. C. *et al.* Enzymatic Cyclization of a Potent Bowman-Birk Protease Inhibitor, Sunflower Trypsin Inhibitor-1, and Solution Structure of an Acyclic Precursor Peptide. *J. Biol. Chem.* **278**, 21782–21789 (2003).
42. Yao, N. *et al.* Discovery of targeting ligands for breast cancer cells using the one-bead one-compound combinatorial method. *J. Med. Chem.* **52**, 126–133 (2008).
43. McBride, J. D., Freeman, N., Domingo, G. J. & Leatherbarrow, R. J. Selection of chymotrypsin inhibitors from a conformationally-constrained combinatorial peptide library. *J. Mol. Biol.* **259**, 819–827 (1996).
44. Quartararo, J. S., Wu, P. & Kritzer, J. A. Peptide Bicycles that Inhibit the Grb2 SH2 Domain. *ChemBioChem* **13**, 1490–1496 (2012).
45. Downward, J. The Grb2/sem-5 adapter protein. *FEBS Lett.* **338**, 113–117 (1992).
46. Lowenstein, E. The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* **70**, 431–442 (1992).
47. Lung, F.-D. *et al.* Functional preference of the constituent amino acid residues in a phage-library-based nonphosphorylated inhibitor of the Grb2-SH2 domain. *J. Pept. Res.* **57**, 447–454 (2001).
48. Lung, F.-D. *et al.* Novel peptide inhibitors for Grb2 SH2 domain and their detection by surface plasmon resonance. *J. Pept. Res.* **60**, 143–149 (2002).
49. Shi, Y. *et al.* Binding affinity difference induced by the stereochemistry of the sulfoxide bridge of the cyclic peptide inhibitors of Grb2-SH2 domain: NMR studies for the structural origin. *Biochem. Biophys. Res. Commun.* **330**, 1254–1261 (2005).
50. Li, P. *et al.* Structural basis for a non-phosphorus-containing cyclic peptide binding to Grb2-SH2 domain with high affinity. *Biochem. Biophys. Res. Commun.* **307**, 1038–1044 (2003).
51. Jiang, S. *et al.* Discovery of thioether-bridged cyclic pentapeptides binding to Grb2-SH2 domain with high affinity. *Bioorg. Med. Chem. Lett.* **19**, 2693–2698 (2009).
52. Gao, Y., Voigt, J., Wu, J., X., Yang, D. & Burke, T. R. Macrocyclization in the Design of a Conformationally Constrained Grb2 SH2 Domain Inhibitor. *Bioorg. Med. Chem. Lett.* **11**, 1889–1892 (2001).
53. Gao, Y., Wei, C.-Q. & Burke, T. R. Olefin Metathesis in the Design and Synthesis of a Globally Constrained Grb2 SH2 Domain Inhibitor. *Org. Lett.* **3**, 1617–1620 (2001).

54. Gao, Y. *et al.* Inhibition of Grb2 SH2 Domain Binding by Non-Phosphate-Containing Ligands. 2. 4-(2-Malonyl)phenylalanine as a Potent Phosphotyrosyl Mimetic †. *J. Med. Chem.* **43**, 911–920 (2000).
55. Liu, B. A. *et al.* The Human and Mouse Complement of SH2 Domain Proteins—Establishing the Boundaries of Phosphotyrosine Signaling. *Mol. Cell* **22**, 851–868 (2006).
56. Lee, K., Zhang, M., Liu, H., Yang, D. & Burke, T. R. Utilization of a  $\beta$ -Aminophosphotyrosyl Mimetic in the Design and Synthesis of Macrocyclic Grb2 SH2 Domain-Binding Peptides. *J. Med. Chem.* **46**, 2621–2630 (2003).
57. Phan, J., Shi, Z.-D., Burke, T. R. & Waugh, D. S. Crystal Structures of a High-affinity Macrocyclic Peptide Mimetic in Complex with the Grb2 SH2 Domain. *J. Mol. Biol.* **353**, 104–115 (2005).
58. Lou, Y.-C. *et al.* Solution structure and dynamics of G1TE, a nonphosphorylated cyclic peptide inhibitor for the Grb2 SH2 domain. *Arch. Biochem. Biophys.* **372**, 309–314 (1999).
59. Machida, K. *et al.* High-Throughput Phosphotyrosine Profiling Using SH2 Domains. *Mol. Cell* **26**, 899–915 (2007).
60. Tureček, F. NC  $\alpha$  Bond Dissociation Energies and Kinetics in Amide and Peptide Radicals. Is the Dissociation a Non-ergodic Process? *J. Am. Chem. Soc.* **125**, 5954–5963 (2003).
61. Jorgensen, W. L. & Gao, J. Cis-trans energy difference for the peptide bond in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **110**, 4212–4216 (1988).
62. Rahuel, J. *et al.* Structural basis for specificity of GRB2-SH2 revealed by a novel ligand binding mode. *Nat. Struct. Biol.* **3**, 586–589 (1996).
63. Thakkar, A., Trinh, T. B. & Pei, D. Global Analysis of Peptide Cyclization Efficiency. *ACS Comb. Sci.* **15**, 120–129 (2013).
64. Karlström, A. & Undén, A. A new protecting group for aspartic acid that minimizes piperidine-catalyzed aspartimide formation in Fmoc solid phase peptide synthesis. *Tetrahedron Lett.* **37**, 4243–4246 (1996).
65. Palasek, S. A., Cox, Z. J. & Collins, J. M. Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. *J. Pept. Sci.* **13**, 143–148 (2007).
66. Songyang, Z. Platelet-derived Growth Factor-induced Formation of Tensin and Phosphoinositide 3-Kinase Complexes. *J. Biol. Chem.* **271**, 23452–23457 (1996).
67. Wavreille, A.-S. & Pei, D. A Chemical Approach to the Identification of Tensin-Binding Proteins. *ACS Chem. Biol.* **2**, 109–118 (2007).

68. Zhang, Y., Zhou, S., Wavreille, A.-S., DeWille, J. & Pei, D. Cyclic Peptidyl Inhibitors of Grb2 and Tensin SH2 Domains Identified from Combinatorial Libraries. *J. Comb. Chem.* **10**, 247–255 (2008).
69. Korsinczky, M., Schirra, H. J. & Craik, D. Sunflower trypsin inhibitor-1. *Curr. PROTEIN Pept. Sci.* **5**, 351–364 (2004).
70. Bechara, C. & Sagan, S. Cell-penetrating peptides: 20years later, where do we stand? *FEBS Lett.* **587**, 1693–1702 (2013).
71. Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R. & Brock, R. A Comprehensive Model for the Cellular Uptake of Cationic Cell-penetrating Peptides. *Traffic* **8**, 848–866 (2007).
72. Futaki, S. Arginine-rich Peptides. AN ABUNDANT SOURCE OF MEMBRANE-PERMEABLE PEPTIDES HAVING POTENTIAL AS CARRIERS FOR INTRACELLULAR PROTEIN DELIVERY. *J. Biol. Chem.* **276**, 5836–5840 (2001).
73. Lian, W., Jiang, B., Qian, Z. & Pei, D. Cell-Permeable Bicyclic Peptide Inhibitors against Intracellular Proteins. *J. Am. Chem. Soc.* **136**, 9830–9833 (2014).
74. Wallbrecher, R. *et al.* Exploration of the Design Principles of a Cell-Penetrating Bicyclic Peptide Scaffold. *Bioconjug. Chem.* 140416143342001 (2014). doi:10.1021/bc500107f
75. Okamoto, T. *et al.* Stabilizing the Pro-Apoptotic BimBH3 Helix (BimSAHB) Does Not Necessarily Enhance Affinity or Biological Activity. *ACS Chem. Biol.* **8**, 297–302 (2013).
76. Bird, G. H., Bernal, F., Pitter, K. & Walensky, L. D. in *Methods in Enzymology* **446**, 369–386 (Elsevier, 2008).
77. Bird, G. H., Christian Crannell, W. & Walensky, L. D. in *Current Protocols in Chemical Biology* (eds. Arkin, A. P. *et al.*) (John Wiley & Sons, Inc., 2011). at <<http://doi.wiley.com/10.1002/9780470559277.ch110042>>
78. Yu, P., Liu, B. & Kodadek, T. A high-throughput assay for assessing the cell permeability of combinatorial libraries. *Nat. Biotechnol.* **23**, 746–751 (2005).
79. Cascales, L. *et al.* Identification and Characterization of a New Family of Cell-penetrating Peptides: CYCLIC CELL-PENETRATING PEPTIDES. *J. Biol. Chem.* **286**, 36932–36943 (2011).
80. Vives, E., Brodin, P. & Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* **272**, 16010–16017 (1997).

81. Wadia, J. S., Stan, R. V. & Dowdy, S. F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* **10**, 310–315 (2004).
82. Stanzi, E. G., Trantow, B. M., Vargas, J. R. & Wender, P. A. Fifteen Years of Cell-Penetrating, Guanidinium-Rich Molecular Transporters: Basic Science, Research Tools, and Clinical Applications. *Acc. Chem. Res.* **46**, 2944–2954 (2013).
83. Koren, E. & Torchilin, V. P. Cell-penetrating peptides: breaking through to the other side. *Trends Mol. Med.* **18**, 385–393 (2012).
84. Appelbaum, J. S. *et al.* Arginine Topology Controls Escape of Minimally Cationic Proteins from Early Endosomes to the Cytoplasm. *Chem. Biol.* **19**, 819–830 (2012).
85. Jo, H. *et al.* Development of  $\alpha$ -Helical Calpain Probes by Mimicking a Natural Protein–Protein Interaction. *J. Am. Chem. Soc.* **134**, 17704–17713 (2012).
86. Muppidi, A. *et al.* Design of antiviral stapled peptides containing a biphenyl cross-linker. *Bioorg. Med. Chem. Lett.* **24**, 1748–1751 (2014).
87. Muppidi, A., Wang, Z., Li, X., Chen, J. & Lin, Q. Achieving cell penetration with distance-matching cysteine cross-linkers: a facile route to cell-permeable peptide dual inhibitors of Mdm2/Mdmx. *Chem. Commun.* **47**, 9396 (2011).
88. Oehlke, J. *et al.* Cellular uptake of an alpha-helical amphipathic model peptide with the potential to deliver polar compounds into the cell interior non-endocytically. *Biochim. Biophys. Acta* **1414**, 127–139 (1998).
89. Mäe, M. *et al.* A stearylated CPP for delivery of splice correcting oligonucleotides using a non-covalent co-incubation strategy. *J. Controlled Release* **134**, 221–227 (2009).
90. Heitz, F., Morris, M. C. & Divita, G. Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *Br. J. Pharmacol.* **157**, 195–206 (2009).
91. Rydström, A. *et al.* Direct Translocation as Major Cellular Uptake for CADY Self-Assembling Peptide-Based Nanoparticles. *PLoS ONE* **6**, e25924 (2011).
92. Wei, C.-Q. *et al.* Macrocyclization in the Design of Grb2 SH2 Domain-Binding Ligands Exhibiting High Potency in Whole-Cell Systems. *J. Med. Chem.* **46**, 244–254 (2003).
93. Räägel, H., Säälük, P. & Pooga, M. Peptide-mediated protein delivery—Which pathways are penetrable? *Biochim. Biophys. Acta BBA - Biomembr.* **1798**, 2240–2248 (2010).

94. Macia, E. *et al.* Dynasore, a Cell-Permeable Inhibitor of Dynamin. *Dev. Cell* **10**, 839–850 (2006).
95. Ochocki, J. D., Mullen, D. G., Wattenberg, E. V. & Distefano, M. D. Evaluation of a cell penetrating prenylated peptide lacking an intrinsic fluorophore via in situ click reaction. *Bioorg. Med. Chem. Lett.* **21**, 4998–5001 (2011).
96. Furet, P. *et al.* Structure-based design and synthesis of high affinity tripeptide ligands of the Grb2-SH2 domain. *J. Med. Chem.* **41**, 3442–3449 (1998).
97. Furet, P. *et al.* Structure-based design and synthesis of phosphinate isosteres of phosphotyrosine for incorporation in Grb2-SH2 domain inhibitors. Part 1. *Bioorg. Med. Chem. Lett.* **10**, 2337–2341 (2000).
98. Burke, T. R. J. *et al.* Nonhydrolyzable phosphotyrosyl mimetics for the preparation of phosphatase-resistant SH2 domain inhibitors. *Biochemistry (Mosc.)* **33**, 6490–6494 (1994).
99. Benfield, A. P., Whiddon, B. B., Clements, J. H. & Martin, S. F. Structural and energetic aspects of Grb2-SH2 domain-swapping. *Arch. Biochem. Biophys.* **462**, 47–53 (2007).
100. Schiering, N., Casale, E., Caccia, P., Giordano, P. & Battistini, C. Dimer Formation through Domain Swapping in the Crystal Structure of the Grb2-SH2-Ac-pYVNV Complex ‡. *Biochemistry (Mosc.)* **39**, 13376–13382 (2000).
101. Dastidar, S. & Mukhopadhyay, C. Structure, dynamics, and energetics of water at the surface of a small globular protein: A molecular dynamics simulation. *Phys. Rev. E* **68**, (2003).
102. Smolin, N. & Winter, R. Molecular Dynamics Simulations of Staphylococcal Nuclease: Properties of Water at the Protein Surface. *J. Phys. Chem. B* **108**, 15928–15937 (2004).
103. Sheu, S.-Y. & Yang, D.-Y. Determination of Protein Surface Hydration Shell Free Energy of Water Motion: Theoretical Study and Molecular Dynamics Simulation. *J. Phys. Chem. B* **114**, 16558–16566 (2010).
104. Otting, G. & Wuethrich, K. Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules. *J. Am. Chem. Soc.* **111**, 1871–1875 (1989).
105. Otting, G. NMR studies of water bound to biological molecules. *J. Prog. Nucl. Magn. Reson. Spectrosc.* **31**, 259–285 (1997).
106. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).

107. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
108. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
109. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577 (1995).
110. Yang, C.-Y. *et al.* PhosphoPOINT: a comprehensive human kinase interactome and phospho-protein database. *Bioinformatics* **24**, i14–i20 (2008).
111. Vintonyak, V. V., Antonchick, A. P., Rauh, D. & Waldmann, H. The therapeutic potential of phosphatase inhibitors. *Curr. Opin. Chem. Biol.* **13**, 272–283 (2009).
112. Zabolotny, J. M. *et al.* PTP1B regulates leptin signal transduction in vivo. *Dev. Cell* **2**, 489–495 (2002).
113. Zinker, B. A. *et al.* PTP1B antisense oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves insulin sensitivity in diabetic mice. *Proc. Natl. Acad. Sci.* **99**, 11357–11362 (2002).
114. Garaud, M. & Pei, D. Substrate Profiling of Protein Tyrosine Phosphatase PTP1B by Screening a Combinatorial Peptide Library. *J. Am. Chem. Soc.* **129**, 5366–5367 (2007).
115. Liu, S. *et al.* Targeting Inactive Enzyme Conformation: Aryl Diketoacid Derivatives as a New Class of PTP1B Inhibitors. *J. Am. Chem. Soc.* **130**, 17075–17084 (2008).
116. Shen, K. *et al.* Acquisition of a Specific and Potent PTP1B Inhibitor from a Novel Combinatorial Library and Screening Procedure. *J. Biol. Chem.* **276**, 47311–47319 (2001).
117. Dixon, J. Roles of Aspartic Acid-181 and Serine-222 in Intermediate Formation and Hydrolysis of the Mammalian Protein-Tyrosine-Phosphatase PTP1. *Biochemistry (Mosc.)* **36**, 4568–4575 (1997).
118. Barford, D., Flint, A. J. & Tonks, N. K. Crystal structure of human protein tyrosine phosphatase 1B. *Science* **263**, 1397–1404 (1994).
119. Jia, Z., Barford, D., Flint, A. J. & Tonks, N. K. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science* **268**, 1754–1758 (1995).
120. Puius, Y. A. *et al.* Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: a paradigm for inhibitor design. *Proc. Natl. Acad. Sci.* **94**, 13420–13425 (1997).

121. Szczepankiewicz, B. G. *et al.* Discovery of a Potent, Selective Protein Tyrosine Phosphatase 1B Inhibitor Using a Linked-Fragment Strategy. *J. Am. Chem. Soc.* **125**, 4087–4096 (2003).
122. Liu, G. *et al.* Selective Protein Tyrosine Phosphatase 1B Inhibitors: Targeting the Second Phosphotyrosine Binding Site with Non-Carboxylic Acid-Containing Ligands. *J. Med. Chem.* **46**, 3437–3440 (2003).
123. Liu, G. *et al.* Fragment screening and assembly: a highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor. *J. Med. Chem.* **46**, 4232–4235 (2003).