

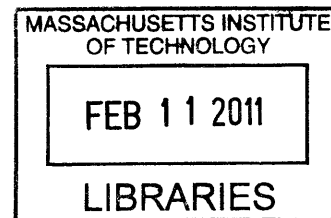
# Mapping the Synapse: Synthesis and Validation of Chemical Probes for Interrogating PDZ Domains

by

Wendy S. Iskenderian-Epps

B.A., Chemistry and Music (2004)

Pomona College



Submitted to the Department of Chemistry  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Chemistry

**ARCHIVES**

at the

Massachusetts Institute of Technology

January 2011

[February 2011]

© 2011 Massachusetts Institute of Technology

All rights reserved

Signature of Author: \_\_\_\_\_

\_\_\_\_\_  
Department of Chemistry  
January 20<sup>th</sup>, 2011

Certified by: \_\_\_\_\_

\_\_\_\_\_  
Barbara Imperiali  
Class of 1922 Professor of Chemistry and Professor of Biology  
Thesis Supervisor

Accepted by: \_\_\_\_\_

\_\_\_\_\_  
Robert W. Field  
Haslam and Dewey Professor of Chemistry  
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:



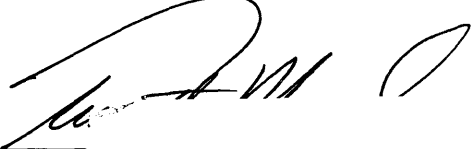
---

Timothy F. Jamison  
Chairman  
Professor of Chemistry  
Massachusetts Institute of Technology



---

Barbara Imperiali  
Thesis Supervisor  
Class of 1922 Professor of Chemistry and Professor of Biology  
Massachusetts Institute of Technology



---

Timothy M. Swager  
John D. MacArthur Professor of Chemistry  
Massachusetts Institute of Technology

**Mapping the Synapse:  
Synthesis and Validation of Chemical Probes  
for Interrogating PDZ Domains**

by

Wendy S. Iskenderian-Epps

Submitted to the Department of Chemistry  
on January 20<sup>th</sup>, 2011 in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy in  
Organic Chemistry

**ABSTRACT**

Macromolecular protein complexes at neuronal synapses are critical for establishing synaptic plasticity, which is the basis of information storage in the brain. These complexes consist of many PDZ domain-containing proteins. PDZ domains bind selectively to peptide sequences at the C-termini of partner proteins, and they play an essential role in the regulation of synaptic macromolecular complexes. While many PDZ domain-containing proteins are well characterized, much remains to be learned about their binding dynamics. The goal of the research presented herein is to gain a quantitative understanding of the complex PDZ-domain binding dynamics through the design, synthesis, and application of sensitive and selective biophysical probes. The peptidyl probes are based on three critical elements: solvatochromic fluorophores, molecular caging, and multivalency.

This thesis presents the development of a systematic approach to screen for selective solvatochromic fluorophore-based probes for class I PDZ domains, and the biophysical characterization of probes for the PSD-95 and Shank3 PDZ domains. These probes have been utilized to examine the effects of Shank3 PDZ domain dimerization on ligand binding *in vitro*, and the quantitative results presented here implicate PDZ domain dimerization as a potential modular control mechanism for ligand-binding in the biological context. Further development of the probes via the application of a novel C-terminal caging strategy is also discussed. Validation *in vitro* shows the utility of a C-terminal 1-(2-nitrophenyl)ethyl cage for blocking the interaction between probe and cognate PDZ domain, and demonstrates the release of viable probe upon photoactivation of the molecular cage.

This thesis also presents the *in vivo* application of PDZ domain probes, using the simple yet powerful eukaryotic *C. elegans* model system. Progress toward solvatochromic fluorophore-based probes for *C. elegans* Lin-10 PDZ domains is presented. Additionally, the design and synthesis of a bivalent peptidyl inhibitor, based on the *C. elegans* C-terminal sequence of STG-2 (a protein of the GLR-1 receptor complex), is discussed. This STG-2 sequence is a putative PDZ domain ligand, and its ortholog in higher eukaryotes functions in receptor-mediated multivalent interactions with partner PDZ domains. Electrophysiological data presented in this thesis suggest the effectiveness of the bivalent STG-2 peptide in the inhibition of the critical *C. elegans* GLR-1 neural receptor complex.

Thesis Supervisor: Professor Barbara Imperiali

## Acknowledgements

I would first like to thank my advisor Barbara Imperiali. Thank you for helping me to become the researcher and teacher I am today, and for all of the opportunities that you have afforded me during my years as a graduate student. I have learned much from you as a mentor, and I have been grateful for the encouraging and friendly lab environment that you always strive to cultivate. I admire your strength and creativity, and your dedication to all aspects of your work. I will never forget all that you have done to open up opportunities for me, now and in the future. And I will of course never forget the scuba diving adventure in the cenote!

A lab wouldn't be what it is without lab members, and I have been fortunate to have wonderful colleagues during my time at MIT. I have learned a lot about science and life from all of the members of the Imperiali group over the past five and a half years, and would like to thank every one of them for this. Thank you to Dr. Matthieu Sainlos, my fellow "neuro" lab member. I would especially like to thank Meredith Hartley, Brenda Goguen, and Angelyn Larkin – together we make up the "four Imperiali first years" from 2005, which is how I will always think of us. Thank you for your support and friendship.

Thank you to all of the MIT professors with whom I have had the pleasure to interact during my graduate school experience. I would like to thank my thesis committee members, Professors Tim Jamison and Tim Swager. A special thank you to the thesis chair Tim Jamison, who has always been interested in my work and my career goals, and in offering helpful guidance.

I would be amiss if I did not acknowledge the exceptional teachers who peaked my interest in chemistry before my days at MIT. To Mrs. Connie Grosse, my El Dorado high school chemistry teacher, I thank you for introducing me to chemistry with your excitement and zest for the subject. Your dedication to students always amazed me, and I am lucky to have been a part of your general chemistry and AP chemistry classrooms. To Professor Daniel O'Leary, my Pomona College chemistry professor and research advisor, I cannot thank you enough. You helped spark my interest in organic chemistry, and later encouraged me into the bioorganic field. Thank you for all of the time you spent with me during your office hours, when I wanted to delve deeper into the subjects we were learning in class. I will always remember the moment when you proudly shook my hand after I told you that I had been accepted to MIT for grad school. I am thankful to have had you as a mentor during my college years, and hope to have the kind of impact on students' lives through my teaching that you had on mine.

Last but certainly not least, I would like to thank my entire family (by blood and marriage) and all of my friends. I thank you all for who you are, and for your unending support. Mom and Dad, this thesis is dedicated to you. Mom, you always supported whatever I wanted to try in life, sometimes making sacrifices of your own on my behalf, and for all of this I thank you. You have been my greatest teacher, and I hope to become at least half the teacher you are to me. Thank you for teaching me to be a strong woman, and teaching me to follow my heart. Dad, you have forever been my cheerleader, and your voice has always reminded me that *I can* (and also reminded me, "Don't worry, be happy"). I thank you for all the paper editing you did for me as I was growing up, even after you said I had "surpassed you" in certain subjects. I thank you for being the only



parent to ask questions (of my peers) during my college thesis defense symposium, even though at the time I was mortified that you might embarrass me. Of course you didn't embarrass me, and I know that your questions always signify genuine interest. I thank you for your interest in who I am and what I do, and for your unconditional belief in me from the very beginning.

Tam(i) and Ryan, thank you for being such wonderful siblings. Tam, you and I have a sisterly bond that cannot be explained. We may have chosen very different career paths, but you have always really understood me like no one else. Our phone chats have helped me more than you will ever know in getting me to this point in my life. I thank you for sharing your wisdom about life with me, and I thank you for believing in me. Ryan, you always know how to bring a smile to my face, which is more important than you may know. I will never forget the message you left for me, in which you played the song "cause you had a bad day..." I love the types of conversations we sometimes have now, when you tell me about the macroscopic human body, and I tell you about the microscopic chemical systems at play. I look forward to the kinds of conversations we will have in the future. Thank you for being like a big brother to me.

Brenden, thank you for all that you are to me. You believed in me every minute of the way, and I truly don't know how I would have done this without you. We have celebrated (and commiserated) together all along, and I couldn't have chosen a better partner. Thank you for your curiosity in what I do. Thank you for encouraging me to pursue the things in life that bring me happiness and fulfillment. Thank you for savoring all of life's moments with me, from small to grand. I am so excited to now be Dr. and Dr. with you.

# Table of Contents

Abstract .....	3
Acknowledgements.....	4
Table of Contents.....	6
List of Figures and Schemes .....	9
List of Tables.....	12
List of Abbreviations .....	14
<b>Chapter 1. Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules</b>	
1.1. Introduction to Modular Domains.....	19
1.2. C-terminal Blocking Peptides for PDZ Domain Inhibition.....	25
1.3. Nonnative Peptides as PDZ Domain Ligands .....	27
1.4. Small Molecule Inhibitors for PDZ Domains.....	31
1.5. Probing Multivalency and Oligomerization of PDZ Domains .....	33
1.6. Perspectives and Dissertation Objectives.....	35
References.....	36
<b>Chapter 2. Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules</b>	
Introduction .....	40
Results and Discussion	
2.1. Design and Synthesis of a Library of Probes .....	42
2.2. Design and Expression of Protein Constructs .....	47
2.3. Overall Screening Strategy and Validation of Probes.....	48
Conclusions .....	57
Experimental Methods.....	58
Acknowledgements.....	63
References .....	64

### **Chapter 3. Development and Validation of a Novel C-terminal Caging Strategy for PDZ Domain Ligands**

Introduction .....	66
Results and Discussion	
3.1. C-terminal Caging Strategy .....	67
3.2. Design and Synthesis of Caged C-terminal Ligands .....	69
3.3. Validation of Caging Strategy <i>in vitro</i> .....	72
Conclusions .....	76
Experimental Methods .....	78
Acknowledgements.....	82
References .....	82

### **Chapter 4. Application of Chemical Probes to the Study of Shank3 PDZ Domain Dimerization**

Introduction .....	83
Results and Discussion	
4.1. Molecular Biology and Expression of all Shank3 PDZ Domain Constructs .....	86
4.2. Probing the Binding Properties of the Obligate Shank3 PDZ Domain Monomer and Dimer .....	89
4.3. Probing Shank3 PDZ Domain Monomer and Dimer via External Regulator .....	91
4.4. Effects of Dimerization on a Biologically-Relevant Ligand .....	93
Conclusions .....	95
Experimental Methods .....	96
Acknowledgements.....	102
References .....	102

### **Chapter 5. Development of Chemical Probes for Investigation of the *C. elegans* Lin-10 PDZ Domains**

Introduction .....	104
Results and Discussion	
5.1. Initial Ligand Screening for Lin-10 Probes.....	105
5.2. Cell-Permeant Second Generation Probes.....	110
5.3. Initial Probe Testing in <i>C. elegans</i> .....	111
5.4. Efforts Toward Third Generation Probes .....	113
Conclusions .....	115
Experimental Methods .....	116
Acknowledgements.....	119
References .....	120

## **Chapter 6. Development and Evaluation of Dimeric Chemical Probes for the *C. elegans* Glutamate Receptor Complex**

Introduction .....	121
Results and Discussion	
6.1. Design of Dimeric Competitor for STG-2 Protein .....	123
6.2. Synthesis of TAT-STG-2 Dimer and TAT-Control Dimer.....	128
6.3. Testing of Probes <i>in vivo</i> in <i>C. elegans</i> .....	133
Conclusions.....	135
Experimental Methods.....	136
Acknowledgements.....	139
References.....	139

## **Appendix 1. Development of a Sox-Based Sensor for Calmodulin Kinase II (CaMKII) Activity**

Introduction .....	141
Results and Discussion	
A1.1. Sensor Development and Design .....	141
A1.2. Sensor Validation .....	143
A1.3. Cross Reactivity with Other Kinases.....	144
Conclusions.....	145
Experimental Methods.....	146
Acknowledgements.....	149
References.....	149

## List of Figures and Schemes

### Chapter 1. Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules

Figure 1.1. Ribbon diagrams of representative PDZ, SH3, SH2, and WW domain crystal structures .....	20
Figure 1.2. PDZ domain ligand numbering convention .....	21
Figure 1.3. PDZ domains, cognate ligands, and inhibitors discussed in this chapter.....	24
Figure 1.4. Representative examples of diverse inhibitors and probes for PDZ domains.....	25
Figure 1.5. Schematic of two different multivalent PDZ domain interactions .....	33

### Chapter 2. Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules

Figure 2.1. Probe design and function .....	41
Figure 2.2. Three 4-DMAP-derived residues incorporated into the PDZ domain probes..	45
Figure 2.3. Recombinant PDZ domains.....	47
Figure 2.4. Definition of relative fluorescence increase (RFI) .....	50
Figure 2.5. PSD-95, SAP-97, and SAP-102 screen results.....	51
Figure 2.6. Relative fluorescence increases (RFI) with the Shank3 PDZ domain .....	52
Figure 2.7. Relative fluorescence increases (RFI) with the PICK1 and GRIP45 PDZ domains.....	52
Figure 2.8. Representative competition titration .....	54
Figure 2.9. ITC thermogram .....	55
Figure 2.10. Structural modeling of PSD-95 PDZ3 electrostatic potential map and fluorogenic peptide.....	56
Scheme 2.1. Synthesis of general fluorescent probe 4 .....	44
Scheme 2.2. Synthesis of 4-( <i>N,N</i> -dimethylamino)-phthalic anhydride 2 .....	44

### Chapter 3. Development and Validation of a Novel C-terminal Caging Strategy for PDZ Domain Ligands

Figure 3.1. Schematic of control of PDZ domain binding by caging of the ligand's critical C-terminal acid .....	68
Figure 3.2. Qualitative validation of caged PDZ domain ligands .....	74
Figure 3.3. HPLC timecourse of Stargazin probe uncaging .....	75
Scheme 3.1. Synthesis of free acid fluorogenic peptide 2 for caging.....	70
Scheme 3.2. Synthesis of NPE-caged probe 7 .....	71

## **Chapter 4. Application of Chemical Probes to the Study of Shank3 PDZ Domain Dimerization**

Figure 4.1. Probing the Shank PDZ domain .....	84
Figure 4.2. Shank3 PDZ domain constructs.....	87
Figure 4.3. Size exclusion chromatography results for protein constructs .....	88
Figure 4.4. Size exclusion chromatography results for Shank3 PDZ domain alone .....	89
Figure 4.5. Ligand binding to obligate monomer and dimer of Shank3 PDZ domain.....	90
Figure 4.6. Effects of Shank3 PDZ domain dimerization on ligand binding.....	92
Figure 4.7. Data fits of competitive titrations of unlabeled GKAP versus fluorescent probe / Shank3PDZ complexes.....	94

## **Chapter 5. Development of Chemical Probes for Investigation of the *C. elegans* Lin-10 PDZ Domains**

Figure 5.1. Sequence alignment of PDZ domains from <i>C. elegans</i> Lin-10 and <i>R. norvegicus</i> PSD-95.....	106
Figure 5.2. Fluorescence increase results for first generation probes 6-9 and second generation probes 12-15 with Lin-10 constructs.....	108
Figure 5.3. Effect of peptides on inhibition of GLR1-GFP synaptic punctae in a single pair of AVA neurons .....	112
Figure 5.4. Fluorescence increase results for probes 19-22 with Lin-10 constructs .....	114

## **Chapter 6. Development and Evaluation of Dimeric Chemical Probes for the *C. elegans* Glutamate Receptor Complex**

Figure 6.1. Proteins involved in the function of the ionotropic glutamate receptor GLR-1 complex in <i>C. elegans</i> .....	122
Figure 6.2. Schematic of dimer versus tandem PDZ domains .....	123
Figure 6.3. Alignment of <i>R. norvegicus</i> Stargazin, <i>C. elegans</i> STG-1, and <i>C. elegans</i> STG-2 terminal 15 residues .....	123
Figure 6.4. Schematic of the proposed STG-2 dimeric ligand function.....	125
Figure 6.5. Schematic of the importance of linker length for the function of dimeric probes.....	126
Figure 6.6. Modular TAT-modified dimeric ligand design .....	127
Figure 6.7. Modified STG-2 sequence used for monomeric units of the TAT-STG-2 dimer, and control sequence used for monomeric units of TAT-control dimer.....	127
Figure 6.8. Dimer HPLC traces.....	132
Figure 6.9. Current responses to glutamate, from voltage-clamped body wall muscle cells.....	134
Scheme 6.1. Synthetic Route A, two variations.....	129
Scheme 6.2. Synthetic Route B .....	130

**Appendix 1. Development of a Sox-Based Sensor for Calmodulin Kinase II (CaMKII) Activity**

Figure A1.1. Sox-derived amino acid building block for peptide synthesis..... 142  
Figure A1.2. Normalized cross reactivity of PKC isoforms with sensor 3..... 144

## List of Tables

### **Chapter 2. Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules**

Table 2.1. PDZ domains and ligands used in these studies .....	43
Table 2.2. Peptides synthesized and characterized, numbered 1 through 20.....	46
Table 2.3. Fluorescence increases of the C-terminal Stargazin-derived peptide library with PSD-95 and Shank3 PDZ domains.....	50
Table 2.4. Binding constants and fluorescence increases for peptides with the optimally positioned 4-DMAP fluorophore .....	53
Table 2.5. Summary of ITC control results and comparison with literature values.....	55

### **Chapter 3. Development and Validation of a Novel C-terminal Caging Strategy for PDZ Domain Ligands**

Table 2.1. Characterization of caged fluorogenic probes and their parent noncaged peptides .....	72
---	----

### **Chapter 4. Application of Chemical Probes to the Study of Shank3 PDZ Domain Dimerization**

Table 4.1. Dissociation constants from fluorescence titrations and competition experiments .....	93
---	----

### **Chapter 5. Development of Chemical Probes for Investigation of the *C. elegans* Lin-10 PDZ Domains**

Table 5.1. Probes used in the initial screening of Lin-10 PDZ domain constructs .....	107
Table 5.2. Dissociation constants and overall fluorescence increases from fluorescence titrations of probes 6-9 with Lin-10 PDZ-12.....	108
Table 5.3. Second generation probes 12-15 and their characterization .....	109
Table 5.4. Dissociation constants and overall fluorescence increases from fluorescence titrations of probes 13 and 14 with Lin-10 PDZ-12.....	109
Table 5.5. Dissociation constants and overall fluorescence increases from fluorescence titrations of TAT probes 16 and 17 and control 18 with Lin-10 PDZ-12 .....	110
Table 5.6. Probes based on Mint PDZ domain ligands, to screen with Lin-10 PDZ domain constructs.....	113
Table 5.7. Dissociation constants and RFI of probes 19, 23, and 24 .....	114



## **Chapter 6. Development and Evaluation of Dimeric Chemical Probes for the *C. elegans* Glutamate Receptor Complex**

Table 6.1. Optimization of the ratio of 8:9 used in the coupling step of the dimer synthesis.....	131
Table 6.2. Characterization of dimers synthesized.....	133

## **Appendix 1. Development of a Sox-Based Sensor for Calmodulin Kinase II (CaMKII) Activity**

Table A1.1. Sensor sequences and their fluorescence increases at 485 nm.....	142
Table A1.2. Recombinant kinase competency tests .....	143
Table A1.3. Kinetic assays of substrates.....	144

## List of Abbreviations

### Standard one- and three-letter codes used for the amino acids:

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid (Aspartate)
E	Glu	Glutamic acid (Glutamate)
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

### Standard one-letter codes used for the nucleotides:

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

## Other abbreviations:

$\epsilon$	Molar extinction coefficient (at a certain wavelength)
$\phi$	Any hydrophobic amino acid (in PDZ domain ligands)
$\lambda_{em}$	Emission wavelength
$\lambda_{ex}$	Excitation wavelength
4-DMAP	4- <i>N,N</i> , -dimethylaminophthalimide
4-DMN	4- <i>N,N</i> , -dimethylamino-1,8-naphthalimide
6-DMN	6- <i>N,N</i> , -dimethylamino-2,3-naphthalimide
Ac	Acetyl
Ac <sub>2</sub> O	Acetic anhydride
AcOH	Acetic Acid
AKT (PKB)	Protein Kinase B
AMPA	$\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor
AN2	Cell surface protein
AP	Dimerizer AP20187 for Fv domain dimerization
ATP	Adenosine triphosphate
AVA	A pair of symmetric interneurons in <i>C. elegans</i>
BAPTA	(1,2-bis(o-aminophenoxy)ethane- <i>N,N,N',N'</i> -tetraacetic acid)
Boc	<i>tert</i> -Butyloxycarbonyl
BSA	Bovine Serum Albumin protein
<i>C. elegans</i>	<i>Caenorhabditis elegans</i> (worm)
CaMKII	Calmodulin Kinase II
cDNA	Complementary DNA
CID	Chemical Inducer of Dimerization
CRIP1	Cysteine-rich PDZ-binding protein
C-terminus	The carboxy-terminal end of a peptide
Da	Daltons
Dab	L-2,4-Diaminobutyric acid
Dap	L-2,3-Diaminopropionic acid
DHB	2,5-dihydroxybenzoic acid
DIPEA	<i>N,N</i> -Diisopropylethyl amine; Hünig's base
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
Dvl	Dishevelled protein
<i>E. coli</i>	<i>Escherichia coli</i> (bacteria)
EDTA	Ethylenediaminetetraacetic acid
Eq., equiv.	equivalents
ErbB2	Epidermal growth factor receptor
ERK	Extracellular signal-regulated kinase
ESI-MS	ElectroSpray Ionization Mass Spectroscopy
EtOH	Ethanol
FKBP	FK506 binding protein
FLAG	Octapeptide protein purification tag

Fmoc	9-Fluorenylmethoxycarbonyl
FP	Fluorescent peptide
Fv	FKBP Phe36Val domain
GB1	IgG-binding domain B1 of <i>Streptococcal</i> protein G
GFP	Green Fluorescent Protein
GKAP	Guanylate Kinase Associated Protein
GLR-1	Glutamate Receptor 1 ( <i>C. elegans</i> )
GluR1	Glutamate receptor 1 (AMPA subunit)
GluR2	Glutamate receptor 2 (AMPA subunit)
GluR6	Glutamate receptor 6
GRIP1	Glutamate Receptor Interacting Protein 1
GST	Glutathione-S-Transferase (protein)
hv	Light
HATU	O-(7-azabenzotriazole-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate
HBTU	2-(1- <i>H</i> -benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	<i>N</i> -(2-hydroxyethyl)piperazine- <i>N'</i> -ethanesulfonic acid
His <sub>6</sub>	Polyhistidine protein purification tag
HOBt	1-Hydroxy-benzotriazole
HPLC	High-Performance Liquid Chromatography
<i>In cellulo</i>	In the context of the cellular environment; Latin “in cells”
<i>In vacuo</i>	under vacuum; Latin “in a vacuum”
<i>In vitro</i>	Performed outside the context of the cell; Latin: “in glass”
<i>In vivo</i>	Performed in the context of living cells; Latin: “in life”
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
K <sub>D</sub>	Dissociation constant
K <sub>i</sub>	Inhibitor constant
K <sub>m</sub>	Michaelis constant for enzyme-substrate binding
Kv1.4	Potassium channel
LTD	Long-Term Depression
LTP	Long-Term Potentiation
MAGI-3	Membrane Associated Guanylate kinase-related protein 3
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight
MES	2-( <i>N</i> -Morpholino)ethanesulfonic acid
MESNA	2-mercaptoethanesulfonic acid
MHC	Major Histocompatibility Complex
Min	Minutes
MK2	MAPK-activated protein kinase-2
MT	Melatonin receptor
MUPP1	Multi PDZ domain protein 1
MW	Molecular Weight
NaCl	Sodium chloride
nF	<i>p</i> -nitrophenylalanine
NHERF-1	Na/H regulatory factor-1 protein
Ni-NTA	Nickel-NTA agarose resin
Nle	L-Norleucine

NMDAR	N-methyl D-aspartate receptor
NMP	N-Methylpyrrolidinone
nNOS	Nitric Oxide Synthase
NPE	1-(2-nitrophenyl)ethyl cage
NR2A	NMDA receptor subunit A
NR2B	NMDA receptor subunit B
N-terminus	The amino-terminal end of a peptide
Orn	L-Ornithine
pA	Picoamps
PAL-PEG	Peptide Amide Linker Polyethyleneglycol (Peptide synthesis resin)
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDZ	PSD-95/SAP-90, Discs-large, and ZO-1 domain
PDZ12	Tandem of PDZ domains one and two of a protein
PDZ3	Third PDZ domain of a protein
PG	Protecting Groups intact
PICK1	Protein kinase C $\alpha$ binding protein
PKA	Protein Kinase A
PKC	Protein Kinase C
Pre	Presenilin protein
PSD	Postsynaptic Density
PSD-95	Postsynaptic Density protein 95
pSer	Phosphoserine
PTEN	Tumor suppressor protein
<i>R. norvegicus</i>	<i>Rattus norvegicus</i>
RFI	Relative Fluorescence Increase
RhoGEF	Rho Guanine nucleotide Exchange Factor protein
S, sec	Seconds
<i>S. frugiperda</i>	<i>Spodoptera frugiperda</i>
SAM	Sterile Alpha Motif
SAP-102	Synapse-Associated Protein 102
SAP-97	Synapse-Associated Protein 97
SDS-PAGE	Sodium Dodecyl-Sulfate PolyAcrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SH2	Src Homology 2 domain
SH3	Src Homology 3 domain
SOL-1	Transmembrane protein in <i>C. elegans</i>
Sox	Sulfonamido-oxine chromophore
SPPS	Solid Phase Peptide Synthesis
STG-1	Stargazin-like protein 1 ( <i>C. elegans</i> )
STG-2	Stargazin-like protein 2 ( <i>C. elegans</i> )
TARPs	Tetraspanning transmembrane proteins
TAT	Truncated portion of HIV-1 Tat protein basic domain
TEV	Tobacco Etch Virus
TFA	Trifluoroacetic acid

TGT	Resin for peptide synthesis
THF	Tetrahydrofuran
TIPS	Triisopropyl silane
TNBS	Trinitrobenzene sulfonic acid
$t_R$	HPLC retention time
Trt	Tryl protecting group
UV	Ultraviolet
$V_c$	Column volume
$V_e$	Elution volume
$V_{max}$	Maximum (enzymatic) turnover velocity
$V_o$	Void volume
WW	Domain named after defining tryptophan residues
Xaa	Used to denote any amino acid
ZO	Zolula Occludens protein

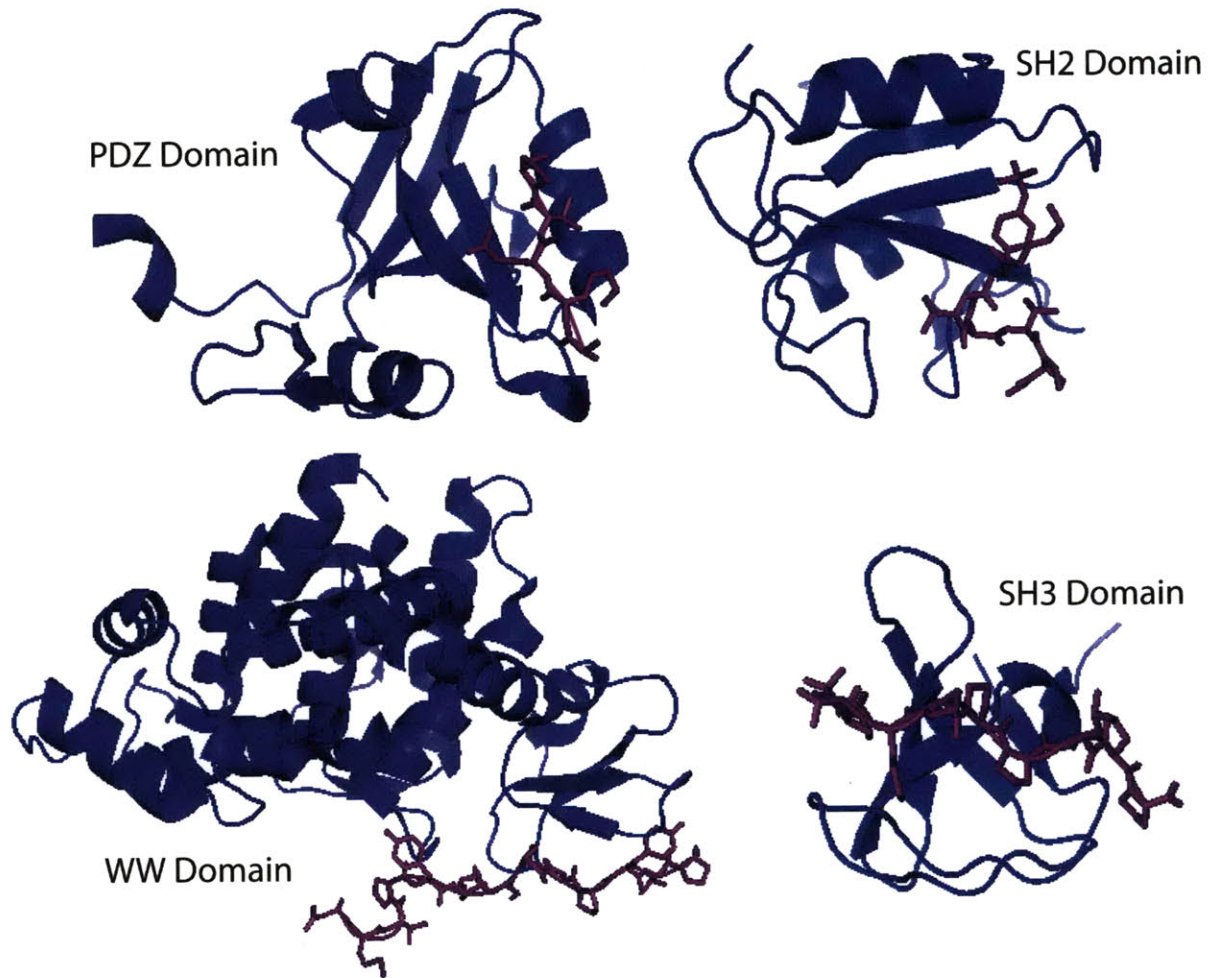
# Chapter 1

## Chemical Biology Approaches to the Study of Modular PDZ Domains via Small Molecules

### *1.1. Introduction to Modular Domains*

Modular domains that mediate protein-protein interactions involved in cellular and receptor-mediated signaling are prevalent in biology. These domains include the PDZ, SH3, SH2, and WW domains, all of which dynamically direct information through signaling networks, and are often regulated via intramolecular interactions and conformational reorganization.<sup>1</sup> From a biophysical and biochemical perspective, protein network signaling is advantageous over signaling through distinct reactions in freely diffusible environments; the advantages of the network approach include enhanced signaling efficiency, specificity provided by spatial segregation, and increased sensitivity resulting from the oligomerization of signaling complexes.<sup>2</sup>

There is structural conservation amongst each protein-interaction module family, and, even in isolation, PDZ, SH3, SH2, and WW domains are highly stable alone, due to their small size and well-defined protein architecture (see Figure 1.1). Additionally, the recognition events mediated by these conserved domains are relatively transferable, allowing for the study of the native behavior of these modular domains alone or in synthetic recombination with other protein domains.<sup>3</sup> These favorable physical properties, taken together with the critical roles the domains play in protein-protein interaction networks, point to these modules as optimal targets for chemical biology. Here a brief overview of each of these families of modular domains will be presented, followed by descriptions of the approaches that chemical biologists have taken to study PDZ domains in particular.

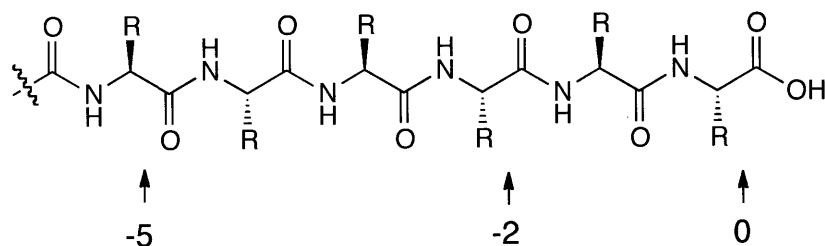


**Figure 1.1.** Ribbon diagrams of representative PDZ, SH3, SH2, and WW domain crystal structures. Clockwise from top left: PSD-95 PDZ3 complexed with KKETPV peptide, PDB entry 1TP3; Grb2 SH2 domain complexed with pYVNVP peptide, PDB entry 3N8M; Abl kinase SH3 domain complexed with Pro-rich peptide, PDB entry 1ABO; Dystrophin WW domain complexed with  $\beta$ -dystroglycan, PDB entry 1EG4.

The PDZ domain is named after the first representative domains of its family that were identified: PSD-95/SAP90, Discs-large, and ZO-1. It is a small,  $\sim$ 100-residue domain that typically binds to ligands represented by the C-terminus of partner proteins through a well-defined binding pocket. These types of domains are abundant in various species, and are among the most common protein domains represented in sequenced genomes.<sup>4</sup> PDZ domains are modular, and consist of 5 or 6  $\beta$ -strands and 2 or 3  $\alpha$ -helices.<sup>5</sup> The binding specificity of



PDZ domains is determined by the interaction of the domain with the side chain of the -2 ligand residue (where, by convention, position 0 is the C-terminal residue and the residue numbers decrease from the C-terminus in the N-terminal direction of the sequence, as shown in Figure 1.2).



**Figure 1.2.** PDZ domain ligand numbering convention.

PDZ classification is based upon the identity of critical residues in the binding ligand: class I PDZ domains bind ligands with a Ser or Thr at the (-)2 position, class II a hydrophobic residue, and class III an Asp or Glu residue. The ligand C-terminal carboxylate is critical for binding in most cases, as are some more N-terminal residues that contribute to binding specificity.<sup>4</sup> PDZ domain modules are prevalent in multi-domain scaffolding proteins, especially in the postsynaptic density of neuronal excitatory synapses. One major example of these PDZ-domain containing scaffolding proteins is the postsynaptic density protein 95 (PSD-95), which is involved in organization of glutamate receptors and in the dynamic trafficking of synaptic proteins. The PDZ domain modules of PSD-95 aid in the assembly of macromolecular protein complexes at the synapse, allowing for this protein to regulate synapse strength.<sup>6</sup>

Another modular protein-protein interaction domain is the Src homology 3 (SH3) domain, one of the best-characterized protein-interaction modules, which is a ~60-residue domain that is found across many protein families. It exhibits a characteristic  $\beta$ -barrel fold, and the domain surface bears a relatively flat, hydrophobic ligand-binding surface. The SH3 domain typically binds ligands with a ProXaaXaaPro tetrapeptide motif, which adopts a specific helical conformation, although there are examples of SH3 ligands without this consensus motif. The binding affinity and selectivity is quite low between SH3 domains and ProXaaXaaPro motif-containing peptides or proteins, implying that other forces lead to better

affinity and selectivity in these systems.<sup>7</sup> One of the most prominent examples of an SH3-mediated interaction comes from the Src tyrosine kinase itself. The interaction of the ProXaaXaaPro motif helical linker region of Src with its SH3 domain constrains this kinase into an inactive conformation. In this case of intramolecular binding, the effective local concentration of ligand and SH3 domain is high, favoring the SH3-bound inactive form even though the binding affinity is modest.<sup>8-10</sup>

One of the smallest yet most versatile protein-interaction modules, the WW domain, interacts with proline-containing ligands, and is named after the two defining tryptophan (W) residues it contains spaced 20-22 residues apart. The interaction of short proline-containing motifs with the WW domain relies upon an interaction between the ligand proline and two WW domain stacked aromatic residues, one of which is a conserved tryptophan.<sup>11</sup> One prominent example of this small domain in cell biology is the role it plays in regulating the connections between the extracellular matrix and the cytoskeleton. The cytoplasmic tail of the membrane-associated dystroglycan binds to the WW domain of the cytoskeletal protein dystrophin, which in turn interacts with the actin cytoskeleton.<sup>12, 13</sup>

Other modular domains are evolved to recognize phosphorylated substrates, which play critical roles in cell signaling. The Src homology 2 (SH2) domain is a ~100-residue domain found in a variety of protein families, from kinases to adaptor proteins to transcription factors. The SH2 domain fold comprises a central anti-parallel  $\beta$ -sheet sandwiched between two  $\alpha$ -helices, and the positively charged pocket that is created by this fold binds to the phosphotyrosine residue of the ligands to which it binds. The specificity of SH2 domains for their respective ligands come from specific interactions between the domain and about 3-6 residues C-terminal to the phosphorylated tyrosine.<sup>14</sup> SH2 domains often recognize autophosphorylated tyrosine residues of receptor tyrosine kinases such as Src, sometimes in an intramolecular fashion. In the case of the autoinhibited Src-family enzymes, the Src SH2 domain binds to a C-terminal phosphotyrosine site within the kinase, which conformationally allows for the SH3 domain to bind its intramolecular ligand. The SH3 and SH2 domains of Src work in concert to achieve this autoinhibition.<sup>8</sup>

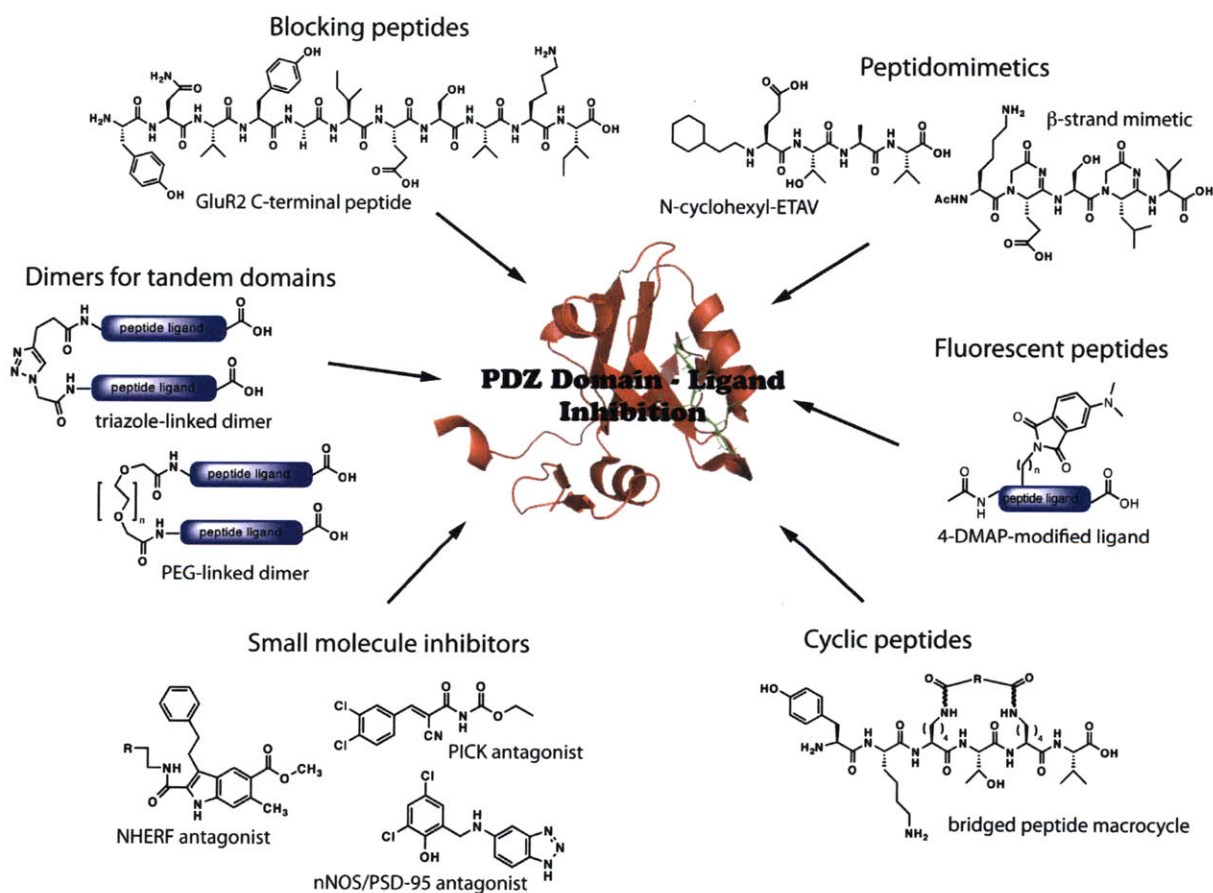
While much is known about the function and binding properties of the modular protein-interaction domains mentioned here, there remain many questions to be answered and important applications to be explored. One focus of particular importance is the development of therapeutics and chemical probes that either inhibit or report on specific protein-protein interactions involving these well-defined protein modules.<sup>15</sup> The field of chemical biology is well equipped for this type of challenge, and much headway has been made on small molecule binders of PDZ, SH3, WW, and SH2 domains. Here the advances that have been made in the development of small molecule probes and inhibitors of PDZ domains in particular will be discussed. Peptide-based probes and inhibitors, including blocking peptides, peptidomimetics, fluorescent peptides, and cyclic peptides, as well as non-peptidyl small molecule antagonists, have been developed; the uses of these molecules to address unanswered biological questions and to move toward therapeutic applications will be reviewed here. Figure 1.3 summarizes the PDZ domain-containing proteins and their ligands discussed in this chapter.

<b>PDZ Domain(s)</b>	<b>Ligand(s)</b>	<b>PDZ Domain Class</b>	<b>Function</b>	<b>Inhibitor(s)</b>
<b>PSD-95</b> (postsynaptic density protein 95)	AMPA interacting <b>stargazin</b> , NMDAR subunits <b>NR2A/NR2B</b> , <b>GluR6</b> kainate receptor, neuronal nitric oxide synthase <b>nNOS</b>	<b>I</b>	glutamate receptor organization; trafficking of synaptic proteins; target for chronic pain treatment	NR2B-based blocking peptides; NR2B-based peptidomimetics; fluorescent probes; cyclic peptides; small molecules; NR2A/NR2B and stargazin-based dimers
<b>PICK1</b> (protein kinase C $\alpha$ binding protein)	AMPA subunit <b>GluR2</b>	<b>II</b>	trafficking of AMPARs	GluR1-based blocking peptide; small molecules
<b>GRIP1</b> (glutamate receptor interacting protein)	AMPA subunit <b>GluR2</b>	<b>II</b>	trafficking and recycling of AMPARs	GluR1-based blocking peptide
<b>MUPP1</b> (multi PDZ domain protein 1)	melatonin receptor <b>MT(1)</b>	<b>III</b>	melatonin receptor signaling	nonnative peptide
<b>Erbin</b> (ErbB2 interacting protein)	epidermal growth factor receptor <b>ErbB2 (HER2)</b>	<b>II</b>	ErbB2 function and localization	nonnative peptide
<b>Shank3</b> (SH3 and multiple ankyrin repeat domains 3 protein)	AMPA subunit <b>GluR1</b> , guanylate kinase associated protein <b>GKAP</b>	<b>I</b>	structural framework of the postsynaptic density; bridge between structural proteins and surface receptors	fluorescent probes
<b>Syntrophin</b> (dystrophin associated protein)	neuronal nitric oxide synthase <b>nNOS</b>	<b>n/a<sup>a</sup></b>	scaffold for proteins involved in cell signaling	cyclic peptides
<b>MAGI-3</b> (membrane associated guanylate kinase-related protein)	tumor suppressor <b>PTEN</b>	<b>I</b>	regulation of PTEN suppressor functions	indole-based small molecules
<b>NHERF-1</b> (Na/H regulatory factor-1 protein)	tumor suppressor <b>PTEN</b> , <b><math>\beta</math>-catenin</b>	<b>I</b>	G protein-coupled trafficking	indole-based small molecules
<b>Dvl</b> (Dishevelled protein)	<b>Frizzled</b> receptors	<b>n/a<sup>a</sup></b>	transduction of Wnt pathway signals	small molecules
<b>RhoGEF</b> (Rho guanine nucleotide exchange factor protein)	receptor <b>plexin B</b>	<b>I</b>	guanine nucleotide exchange factor in Rho-mediated signaling	plexin B-based dimers

**Figure 1.3.** PDZ domains, cognate ligands, and inhibitors discussed in this chapter. References are given in the text. <sup>a</sup> n/a = PDZ domains that bind cognate ligands in non-canonical manners, and are not classified in any of three canonical PDZ domain classes.

## 1.2. C-terminal blocking peptides for PDZ domain inhibition

In principle, if a protein-protein interaction relies upon a distinct interaction site, such as the binding groove of a PDZ domain, this interaction should be modulated or inhibited by “blocking peptides.” In the case of the PDZ domains, these peptides mimic the C-terminus of a PDZ domain ligand, and in unaltered form are simply the terminal ~10 amino acids of the ligand sequence, sometimes with a cell permeabilizing sequence.<sup>16</sup> Since the interaction of the PDZ domain and ligand most often relies exclusively on these terminal residues, blocking peptides should bind with close to the same affinity as their parent proteins from which they are derived. Many successful examples of unmodified blocking peptides have been reported, and have been used to study the effects of competition with specific PDZ domain – ligand interactions *in vitro* and in the native cellular environment.



**Figure 1.4.** Representative examples of diverse inhibitors and probes for PDZ domains. PDZ domain shown is PDZ3 of PSD-95 protein complexed with KKETPV peptide ligand, PDB entry 1TP3. References are given in the text.

C-terminal peptides based on the PDZ binding motif of the glutamate receptor subunit GluR2 have been used to investigate the roles of postsynaptic protein kinase C alpha binding protein (PICK1) and glutamate receptor interacting protein (GRIP) on  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors.<sup>17-19</sup> Specifically, a peptide corresponding to the wild-type GluR2 C-terminus blocks binding of both PICK1 and GRIP to the GluR2 subunit. However, a mutated Ser880 phosphorylated GluR2 C-terminus (Ser880Glu) only blocks PICK1 from interaction with GluR2. Patch-clamp studies, which allow for the electrophysiological investigation of single or multiple ion channels in cells, have shown that the wild-type GluR2 C-terminal peptide (see Figure 1.4) causes an increase in AMPA receptor currents, while identical studies with the Ser880Glu peptide have shown no such effect. This indicates that blocking of both the PICK1 and GRIP interactions with GluR2 are necessary for inhibition of long-term depression (LTD), a lasting decrease in synaptic effectiveness that follows some types of electrical stimulation in the brain.<sup>18</sup>

In addition to the study of AMPA receptors, the study of N-methyl D-aspartate (NMDA) receptors has been greatly aided by the use of blocking peptides that mimic PDZ domain-binding ligands. NMDA receptors are known to mediate ischemic brain injury,<sup>20</sup> and while inhibiting NMDA receptors directly has proven to be a failed strategy for the treatment of specific brain disorders, an alternative is to block downstream excitotoxic signaling without completely blocking NMDARs and the corresponding neurotransmission. A blocking peptide consisting of the C-terminal residues of the NR2B receptor subunit fused to the cell-membrane transduction domain of the HIV-1 TAT protein,<sup>21, 22</sup> which allows for cell permeability, has been shown to successfully reduce the susceptibility of neurons to excitotoxicity and ischemia. Even though there are many possible binding partners for this blocking peptide, the authors determined biochemically that it bound selectively to PSD-95 at low concentrations, and confirmed that only neurons lacking PSD-95 (but not other PSD-95 family members) exhibited reduced vulnerability to NMDA toxicity.<sup>23</sup> This demonstrates that while blocking peptides based on the C-terminus of PDZ domain ligands may have multiple binding partners, general biochemical analyses and cellular investigations can define a maximum threshold concentration under which the blocking peptides are selective for the inhibition of particular interactions.

Additional approaches can provide blocking peptides with higher affinity and selectivity, especially in the context of promiscuous PDZ domain ligands that often bind to multiple partner proteins. Most blocking peptides are based on the wild-type C-terminus of a PDZ domain ligand, and high-throughput screens based on support-bound PDZ domain peptides,<sup>24</sup> as well as microarrays of PDZ domains analyzed in the presence of peptide ligands,<sup>25</sup> can reproduce these native interactions. Variants of ligands that bind to specific PDZ domains can be rationally engineered and screened via the same technologies, to yield peptides that are optimized for both specificity and affinity.<sup>24</sup> Rational (focused) and random libraries of peptides can be subjected to iterative selection-enrichment cycles against particular PDZ domains in C-terminal peptide phage display, allowing for the selection of unnatural peptide ligands. This particular approach does not require chemical synthesis of the peptides until the later validation stages of the process, and has been employed to successfully identify a ligand for the tenth PDZ domain of the scaffolding multi-PDZ domain protein 1 (MUPP1) with a two-fold affinity improvement over the native C-terminal peptide ligand,<sup>26</sup> as well as a 0.02  $\mu$ M ligand for the ErbB2 interacting protein (Erbin) PDZ domain, one of the highest affinity peptide ligands reported for a PDZ domain.<sup>27</sup> Overall, there are many available strategies for screening and optimizing blocking peptides, which have been utilized to address critical questions regarding biological protein-protein interaction systems such as those involving PDZ domains.

### ***1.3. Nonnative Peptides as PDZ Domain Ligands***

While C-terminal blocking peptides have been used successfully to investigate the roles of many PDZ domains in the biological setting, many nonnative PDZ domain ligands have also been designed and pursued, and some of these address shortcomings of native blocking peptides. An overview of the main classes of these nonnative PDZ domain ligands is presented here, including the methods by which they were developed and the information regarding PDZ domain function that has been garnered with their application.

## *Peptidomimetics*

A peptidomimetic approach to develop modified peptides has been used successfully in several studies. For example, it was used to develop inhibitors for the scaffolding protein  $\alpha$ 1-syntrophin PDZ domain that are more potent than the corresponding native peptides. These modified peptides, one of which is shown in Figure 1.4, are protease-resistant peptidomimetic structures incorporating conformationally constrained amino acid surrogates, which reproduce the hydrogen-bonding pattern and side-chain functionality of a  $\beta$ -strand.<sup>28</sup>

The peptidomimetic approach has also proven efficacious for the development of small molecules to inhibit the interaction between NMDA receptors and the PDZ domains of PSD-95, a critical target interaction for the treatment of ischemic brain disorders. Previously-discussed blocking peptide studies identified a TAT-NR2B inhibitory peptide, and this template was used for the development of smaller potential drug targets. Specifically, truncation to a tetrapeptide, N-terminal modifications, and the optimization of side chains with both natural and unnatural functionalities (see Figure 1.4) all led to the most potent inhibitor of the PSD-95 / NMDA receptor interaction, with a 20-fold tighter affinity than the parent TAT-NR2B peptide.<sup>29</sup>

In another approach, involving the rational modification of hexapeptide ligands for PSD-95, researchers utilized structural information from an x-ray crystal structure of ligand bound to PDZ domain 3 of PSD-95 to extend the side chains of specific residues into available space in the PDZ domain-binding pocket. A small library of modified ligands was made via the coupling of a panel of organic acids onto acyl acceptors in the peptide sequence, leading to multiple ligands with submicromolar affinities for this modular domain.<sup>29</sup> These molecules have been validated using *in vitro* binding studies, and the strong affinities they exhibit for the third PDZ domain of PSD-95 point to their potential as cellular probes and drug targets, most likely with stability toward cellular proteases.

## *Fluorescent Peptides*

A different approach to the development of cellular PDZ domain probes relies on fluorescence as a method for detection of binding to and inhibition of target PDZ domains. Fluorescence polarization has been widely utilized to determine binding constants of PDZ peptide ligands identified from larger peptide screens, often via the attachment of the



fluorophore to the N-termini of the peptides.<sup>25</sup> This strategy allows for the quantification of binding affinity and, in addition to isothermal titration calorimetry, is one of the more widely used techniques to quantify peptide–protein dissociation constants. However, not all fluorimeters are equipped for fluorescence polarization measurements, and the low fluorescence signal often afforded by this technique can lead to issues with data interpretation, as the signal to noise ratio can be rather low. Additionally, fluorescence polarization measures the presence of any complex, not just the targeted protein-ligand complex; since the binding site is not necessarily known or targeted, false positive results can be obtained from this method.

A parallel fluorescence-based approach that avoids this concern is the incorporation of small solvatochromic fluorophores into the PDZ ligand as part of the design, effectively creating a fluorescence reporter for binding to PDZ domains and a potential PDZ domain inhibitor (see Figure 1.4). This strategy has been successfully applied to the development of probes for class I PDZ domains including those of PSD-95 and SH3 and multiple ankyrin repeat domains 3 (Shank3) synaptic proteins.<sup>30</sup> The incorporation of dimethylaminophthalimide chromophores in the probes aided in the screening and optimization of the ligands, and functioned as reporters of binding. In this case, incorporation of the solvatochromic moiety at appropriate positions actually increased binding affinity of the probes to partner PDZ domains by an order of magnitude, as compared to the native parent peptides. Overall, this strategy yielded fluorescent peptide probes with enhanced ability to inhibit targeted PDZ domain–ligand interactions,<sup>30</sup> and will be discussed in detail in Chapter 2.

### *Cyclic peptides*

In addition to fluorescent peptides and peptidomimetics, cyclic peptides have shown promise as PDZ domain binders and inhibitors. Cyclization enhances stability to proteases, and therefore increases the duration over which the peptides can inhibit PDZ domain–ligand interactions *in vivo*. The PDZ domains of PSD-95 have been targeted with blocking peptides, peptidomimetics, and fluorescent peptides, and multiple cyclic peptide designs have proven to be effective binders and inhibitors of these domains.<sup>31-34</sup>

In order to introduce the macrocycles onto native peptide scaffolds, researchers have replaced the less critical (-)1 and (-)3 residues of known PSD-95 PDZ domain ligands with residues that can be tethered via bridging elements. Cyclic scaffolds containing lactam bridges of varying lengths have been synthesized via the reaction of  $\alpha$ -,  $\beta$ -, or  $\gamma$ -amino acids with amino- and carboxy-functionalized side chains of residues at the (-)1 and (-)3 positions in PDZ domain ligands.<sup>31, 32</sup> Organic biscarboxylic acids have also been incorporated as bridges, via the formation of amide bonds with two amine-presenting side chains at the (-)1 and (-)3 positions of the peptide.<sup>33</sup> An example of this type of macrocycle is given in Figure 1.4. Each of these strategies allows for the systematic variation of ring size and rigidity of the resulting macrocycles. In one particular case, a lactam bridge-based cyclic analog for the first PDZ domain of PSD-95 produced a 10-fold enhancement in binding affinity over its parent linear peptide ligand.<sup>31</sup> Studies showed that this cyclic scaffold disrupted the interaction of the GluR6 kainate receptor subunit with PSD-95 five- to ten-fold more effectively than the parent linear peptide. The cyclic peptide was also found to disrupt the clustering of kainate receptors for longer periods of time than the linear peptide, pointing to enhanced proteolytic stability.<sup>31</sup>

A cell-permeable version of this cyclic peptide, referred to as CN2097, was utilized to study the effects of PSD-95 PDZ domain inhibition on neuropathologic conditions.<sup>34</sup> PSD-95 is thought to be a target for managing many neurological conditions including chronic pain, and therefore injection of this peptide into rats was predicted to interfere with physiologic phenomena related to pain. Specifically, this cyclic compound was reported to mitigate several phenomena known to promote central sensitization, which is an undesired enhanced responsiveness of certain neurons in the central nervous system to normal input. Application of this inhibitor showed signs of halting chronic pain, through attenuation of neuronal firing in affected rats.<sup>34</sup> Overall, bridged cyclic peptides show promise biochemically and *in vivo* as PSD-95 PDZ domain inhibitors with heightened resistance to protease degradation.

A different approach to the design of cyclic peptide inhibitors veers further from the traditional definition of a PDZ domain ligand, as it yields ligands that lack a free C-terminal carboxylate. Specifically, phage-displayed combinatorial libraries have yielded cyclic peptides without free C-termini that bind the syntrophin PDZ domain.<sup>35</sup> In this case, it was found that peptides with two Cys residues forming an intramolecular disulfide bond could act

as C-terminal peptide mimetics and bind to syntrophin, even while fused to the N-terminus of a bacteriophage via their C-termini. The intrachain disulfide bonds of these peptides were necessary for binding to syntrophin, and like cyclic ligands for other protein classes derived from phage-displayed libraries, these peptides appear to act as structural mimics of natural ligands. While the peptides identified for the syntrophin domain display weaker affinities than the natural ligands,<sup>35</sup> the pursuit of structural ligand mimics instead of primary sequence ligand mimics may prove useful for the generation of inhibitors for this and other PDZ domains.

#### ***1.4. Small Molecule Inhibitors for PDZ Domains***

While peptide-based inhibitors are both useful for probing PDZ domains and as potential drug targets, small molecule antagonists tend to be more efficacious as drugs. The majority of current pharmacotherapeutics target transmembrane receptor proteins, which are critical components of disease-modulating signaling pathways. However, an alternative strategy is the development of drugs for specific protein-protein interaction modules within those pathways. This approach should lead to less undesired side effects, since only specific signaling pathways mediated by each receptor would be affected.<sup>36, 37</sup> Since PDZ domains are implicated in many neurological-based disorders, they have become an attractive target for drug development.

##### *Rational inhibitor design*

Small molecule reversible inhibitors of class I PDZ domains have been successfully developed through rational design.<sup>38-40</sup> Researchers designed a generalizable indole scaffold for PDZ-domain binding, which could be substituted with functional groups mimicking critical peptide ligand side chains. Through molecular modeling analysis, they were able to define the geometry of the side chain functionalities with respect to each other, and incorporate them on the indole scaffold to achieve a similar spatial arrangement.<sup>38</sup> This elegant strategy yielded a reversible inhibitor for the membrane associated guanylate kinase-related (MAGI-3) PDZ domain, yet the molecule also bound to other class I PDZ domains. This is not surprising given the promiscuity of PDZ domain ligands among different classes, and given the inhibitor design, which relies upon critical 0 and (-)2 residue side chains that

are shared among each class of ligands. Due to this lack of selectivity, small changes in the substitution of the indole inhibitor scaffold yielded an inhibitor of the G protein-coupled trafficker Na/H regulatory factor-1 (NHERF-1) PDZ domain with similar promiscuity.<sup>40</sup> Interestingly, a derivative of this inhibitor containing an indole-3-methanol functionality yielded an electrophilic probe that could covalently label the conserved His residue of class I PDZ domains such as MAGI-3 and NHERF-1 (see structure in Figure 1.4).<sup>40</sup> Although not selective among class I PDZ domains, this type of probe could be useful for proteomics studies involving PDZ domains, and the general design could be applied to create probes for other classes of PDZ domains. Optimization would be necessary to apply this type of rational design strategy to the development of specific PDZ domain inhibitors, although this research marks an important step toward the rational design of small molecule inhibitors of protein-protein interactions.

### *Screening for small molecule inhibitors*

NMR spectroscopy-based screening,<sup>41</sup> often in combination with computational structure-based ligand screening,<sup>42, 43</sup> has proven useful in the identification of small molecule PDZ domain inhibitors. Small molecule inhibitors for the Dishevelled (Dvl) PDZ domain have been identified in this manner.<sup>42, 43</sup> Dvl is involved in the Wnt signaling pathway, and helps to transduce Wnt signals from the receptor Frizzled to downstream components. Low micromolar binders of Dvl identified through these computational and NMR-based screens were shown to effectively block Wnt signaling *in vivo*.<sup>43</sup> Indeed, targeting of the PDZ domain of Dvl with small molecule inhibitors effectively inhibits Wnt signaling, both offering a tool for dissecting the Wnt pathway components, as well as a potential drug target.

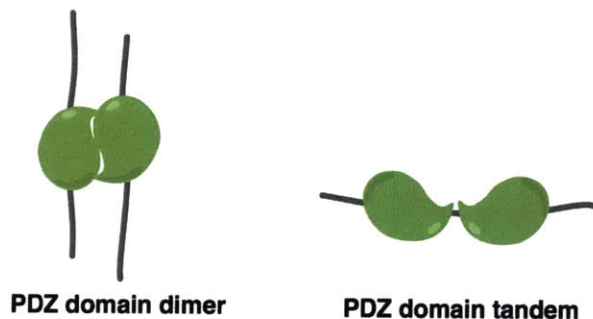
A select number of small molecule inhibitors of PDZ domain interactions that also show promising effects *in vivo* have been identified via chemical library screens.<sup>44, 45</sup> These small molecules exhibit promise as therapeutic targets for the management of neuropathic pain. One of these, a small molecule inhibitor of the PICK1 PDZ domain (see Figure 1.4), displays a low micromolar affinity akin to that observed for native PICK1 ligands, and is specific for this particular PDZ domain. PICK1 binds to the C-terminus of AMPA receptors, and is involved in the trafficking of the GluR2 subunits of these receptors. Pretreatment of cultured hippocampal neurons with the identified small molecule inhibited

coimmunoprecipitation of GluR2 and PICK1, and accelerated the recycling of GluR2. In addition, this small molecule blocked long-term depression and long-term potentiation in neurons (the decrease and enhancement in neuronal signal transmission, respectively), demonstrating its ability to inhibit synaptic plasticity.<sup>44</sup>

Another promising small molecule for the treatment of chronic pain inhibits the interaction between neuronal nitric oxide synthase (nNOS) and PSD-95 PDZ domains (structure given in Figure 1.4).<sup>45</sup> This small molecule was shown to disrupt *in vitro* binding between these two proteins, and was efficacious in cell-based and *in vivo* pain models. Additionally, there is approximately a 100-fold difference between the estimated IC<sub>50</sub> values of this compound for inhibiting thermal hyperalgesia (a measure of neuropathic pain) and the undesired inhibition of motor coordination, unlike the best 3- to 6-fold difference found for a reported NMDA receptor antagonist.<sup>45</sup> These types of small molecule drugs show promise in terms of potency as well as specificity for the particular signaling pathways of interest, since they target protein-protein interactions further downstream from receptors.

### 1.5. Probing Multivalency and Oligomerization of PDZ Domains

Thus far, progress in the development of binders for single PDZ domains has been discussed. However, PDZ domain – ligand interactions are complex phenomena that are regulated by many factors, including presentation of multiple PDZ domains along single polypeptide chains, as well as the ability to homo- and hetero-oligomerize. In order to best enhance the ligand-binding affinity of PDZ domains, properties such as oligomerization have been exploited. Figure 1.5 depicts these variations of multivalent interactions.



**Figure 1.5.** Schematic of two different multivalent PDZ domain interactions. Left: dimer (made up of two PDZ domains from different protein chains). Right: tandem domains (interacting domains from the same protein chain).

Monovalent fluorescent probes, which were introduced earlier, have been used to quantify the effects of Shank3 PDZ domain dimerization on ligand-binding affinity, as will be discussed in Chapter 4. The dissociation constant of monovalent ligands was shown to decrease 8- to 30-fold for dimeric Shank3 PDZ domain (in comparison to monomeric domain).<sup>46</sup> This quantitative information reveals the critical role of oligomeric interactions in the binding of peptidyl probes, which can be efficacious in the design of higher affinity PDZ domain binders.

Toward that aim, there have been reports of bivalent ligands designed to bind to either dimeric PDZ domain targets<sup>47</sup> or tandem PDZ domains of the same protein.<sup>48, 49</sup> It has been suggested that hetero- and homo- oligomerization control the activity of the Rho guanine nucleotide exchange factor (RhoGEF) PDZ domain. Additionally, clustering of plexin B, a ligand for the RhoGEF PDZ domain, is of functional importance in Rho-mediated signaling. For these reasons, the interaction between a dimeric form of RhoGEF PDZ domain and a bivalent peptide mimicking the dimer of the plexin B receptor ligand for RhoGEF was investigated. Peptide dimerization was achieved via cross-linking of the N-terminal ends of peptides with PEG (polyethylene glycol) spacers, keeping the critical C-termini available for binding (see Figure 1.4). These bivalent ligands bound to the dimeric PDZ domain with an up to 20-fold increase in affinity compared to the simple monomeric protein and ligand system.<sup>50</sup>

Tandem PDZ domains typically function together to bind ligands in a cooperative manner. The first two PDZ domains of PSD-95 and its family of proteins are linked by a conserved peptide linker of less than five amino acids, and the corresponding C-terminal peptide-binding grooves are structurally aligned in a manner for preferential binding to multimeric membrane PDZ domain ligands.<sup>51</sup> As such, these tandem domains present validated targets for bivalent peptide ligands. Bivalent PEG-linked peptides, based on the C-terminal residues of the NMDA receptor subunit ligands for PSD-95, bind the PSD-95 tandem via a two-step process, and show increased affinity relative to the parent NR2A/NR2B monomeric ligand.<sup>48</sup> Bivalent triazole-linked peptides (shown in Figure 1.4) synthesized via click chemistry, based on the C-terminus of the membrane-bound Stargazin ligand, have also been developed and tested.<sup>49</sup> Studies *in vitro* and in a cellular FRET-based model show enhanced affinity of the dimer for the PSD-95 tandem as compared to the parent monovalent

peptide. These bivalent ligands were also successfully used as PSD-95 PDZ domain inhibitors in rat hippocampal neurons, clearly showing effective disruption of AMPA receptor complexes.<sup>49</sup> Overall, these various bivalent ligands represent a critical step toward higher affinity probes and inhibitors of PDZ domains, and demonstrate the utility and necessity of biological contextualization in the design of inhibitors.

### ***1.6. Perspectives and Dissertation Objectives***

Protein-protein interactions have become of interest as targets for inhibitors and therapeutics due to their prevalence in biology and the specific roles they play in normal and disease states. Protein-protein interactions involving the PDZ domain are well defined, typically depending on the interaction between the PDZ binding groove and the C-terminal peptide sequence of the partner protein ligand. In this thesis, the design, development, validation, and application of tools to study these important domains will be presented. Specifically, fluorescent peptidyl probes for PDZ domains have been developed, which both detect cognate binding and out-compete natural PDZ domain-ligand interactions. These peptide inhibitors have been further developed as caged probes (using a novel caging approach), which only bind after uncaging with ultraviolet light and release of the “active” ligand, and are therefore temporally and spatially controllable. The fluorescent probes have been utilized *in vitro* to study the behavior of a specific PDZ domain and the effects of domain dimerization on C-terminal peptide binding, and the fluorescent probes have also been applied to *in vivo* studies in *C. elegans*. Lastly, bivalent peptides have been developed and utilized for the investigation of *C. elegans* AMPA receptor complexes and the interactions between their component proteins. These efforts, as well as ongoing research and future directions of study, will be discussed in detail in the forthcoming chapters.

## References

1. Pawson, T. "Dynamic control of signaling by modular adaptor proteins." *Curr. Opin. Cell. Biol.* **2007**, *19*, 112-116.
2. Zhang, M.; Wang, W. "Organization of signaling complexes by PDZ-domain scaffold proteins." *Accounts Chem. Res.* **2003**, *36*, 530-538.
3. Bhattacharyya, R. P.; Remenyi, A.; Yeh, B. J.; Lim, W. A. "Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits." *Annu. Rev. Biochem.* **2006**, *75*, 655-680.
4. Hung, A. Y.; Sheng, M. "PDZ domains: structural modules for protein complex assembly." *J. Biol. Chem.* **2002**, *277*, 5699-5702.
5. Lee, H. J.; Zheng, J. J. "PDZ domains and their binding partners: structure, specificity, and modification." *Cell Commun. Signal.* **2010**, *8*, 8-26.
6. Kim, E.; Sheng, M. "PDZ domain proteins of synapses." *Nat. Rev. Neurosci.* **2004**, *5*, 771-781.
7. Mayer, B. J. "SH3 domains: complexity in moderation." *J. Cell Sci.* **2001**, *114*, 1253-1263.
8. Sicheri, F.; Moarefi, I.; Kuriyan, J. "Crystal structure of the Src family tyrosine kinase Hck." *Nature* **1997**, *385*, 602-609.
9. Williams, J. C.; Weijland, A.; Gonfloni, S.; Thompson, A.; Courtneidge, S. A.; Superti-Furga, G.; Wierenga, R. K. "The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions." *J. Mol. Biol.* **1997**, *274*, 757-775.
10. Xu, W.; Harrison, S. C.; Eck, M. J. "Three-dimensional structure of the tyrosine kinase c-Src." *Nature* **1997**, *385*, 595-602.
11. Ilsley, J. L.; Sudol, M.; Winder, S. J. "The WW domain: linking cell signalling to the membrane cytoskeleton." *Cell Signal.* **2002**, *14*, 183-189.
12. Bork, P.; Sudol, M. "The WW domain: a signalling site in dystrophin?" *Trends Biochem. Sci.* **1994**, *19*, 531-533.
13. Andre, B.; Springael, J. Y. "WWP, a new amino acid motif present in single or multiple copies in various proteins including dystrophin and the SH3-binding Yes-associated protein YAP65." *Biochem. Biophys. Res. Commun.* **1994**, *205*, 1201-1205.
14. Pawson, T.; Gish, G. D.; Nash, P. "SH2 domains, interaction modules and cellular wiring." *Trends Cell Biol.* **2001**, *11*, 504-511.
15. Blazer, L. L.; Neubig, R. R. "Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles." *Neuropsychopharmacol.* **2009**, *34*, 126-141.
16. Dev, K. K. "Making protein interactions druggable: targeting PDZ domains." *Nat. Rev. Drug Discov.* **2004**, *3*, 1047-1056.
17. Hirbec, H.; Francis, J. C.; Lauri, S. E.; Braithwaite, S. P.; Coussen, F.; Mulle, C.; Dev, K. K.; Coutinho, V.; Meyer, G.; Isaac, J. T.; Collingridge, G. L.; Henley, J. M. "Rapid and differential regulation of AMPA and kainate receptors at hippocampal mossy fibre synapses by PICK1 and GRIP." *Neuron* **2003**, *37*, 625-638.
18. Daw, M. I.; Chittajallu, R.; Bortolotto, Z. A.; Dev, K. K.; Duprat, F.; Henley, J. M.; Collingridge, G. L.; Isaac, J. T. "PDZ proteins interacting with C-terminal GluR2/3 are involved in a PKC-dependent regulation of AMPA receptors at hippocampal synapses." *Neuron* **2000**, *28*, 873-886.



19. Nishimune, A.; Isaac, J. T.; Molnar, E.; Noel, J.; Nash, S. R.; Tagaya, M.; Collingridge, G. L.; Nakanishi, S.; Henley, J. M. "NSF binding to GluR2 regulates synaptic transmission." *Neuron* **1998**, *21*, 87-97.
20. Simon, R. P.; Swan, J. H.; Griffiths, T.; Meldrum, B. S. "Blockade of N-methyl-D-aspartate receptors may protect against ischemic damage in the brain." *Science* **1984**, *226*, 850-852.
21. Fawell, S.; Seery, J.; Daikh, Y.; Moore, C.; Chen, L. L.; Pepinsky, B.; Barsoum, J. "Tat-mediated delivery of heterologous proteins into cells." *Proc. Natl. Acad. Sci.* **1994**, *91*, 664-668.
22. Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F. "In vivo protein transduction: delivery of a biologically active protein into the mouse." *Science* **1999**, *285*, 1569-1572.
23. Cui, H.; Hayashi, A.; Sun, H. S.; Belmares, M. P.; Cobey, C.; Phan, T.; Schweizer, J.; Salter, M. W.; Wang, Y. T.; Tasker, R. A.; Garman, D.; Rabinowitz, J.; Lu, P. S.; Tymianski, M. "PDZ protein interactions underlying NMDA receptor-mediated excitotoxicity and neuroprotection by PSD-95 inhibitors." *J. Neurosci.* **2007**, *27*, 9901-9915.
24. Wiedemann, U.; Boisguerin, P.; Leben, R.; Leitner, D.; Krause, G.; Moelling, K.; Volkmer-Engert, R.; Oschkinat, H. "Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides." *J. Mol. Biol.* **2004**, *343*, 703-718.
25. Stiffler, M. A.; Grantcharova, V. P.; Sevecka, M.; MacBeath, G. "Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays." *J. Am. Chem. Soc.* **2006**, *128*, 5913-5922.
26. Sharma, S. C.; Memic, A.; Rupasinghe, C. N.; Duc, A. C.; Spaller, M. R. "T7 phage display as a method of peptide ligand discovery for PDZ domain proteins." *Biopolymers* **2009**, *92*, 183-193.
27. Skelton, N. J.; Koehler, M. F.; Zobel, K.; Wong, W. L.; Yeh, S.; Pisabarro, M. T.; Yin, J. P.; Lasky, L. A.; Sidhu, S. S. "Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin PDZ domain." *J. Biol. Chem.* **2003**, *278*, 7645-7654.
28. Hammond, M. C.; Harris, B. Z.; Lim, W. A.; Bartlett, P. A. "Beta strand peptidomimetics as potent PDZ domain ligands." *Chem. Biol.* **2006**, *13*, 1247-1251.
29. Bach, A.; Chi, C. N.; Olsen, T. B.; Pedersen, S. W.; Roder, M. U.; Pang, G. F.; Clausen, R. P.; Jemth, P.; Stromgaard, K. "Modified peptides as potent inhibitors of the postsynaptic density-95/N-methyl-D-aspartate receptor interaction." *J. Med. Chem.* **2008**, *51*, 6450-6459.
30. Sainlos, M.; Iskenderian, W. S.; Imperiali, B. "A general screening strategy for peptide-based fluorogenic ligands: Probes for dynamic studies of PDZ domain-mediated interactions." *J. Am. Chem. Soc.* **2009**, *131*, 6680-6682.
31. Piserchio, A.; Salinas, G. D.; Li, T.; Marshall, J.; Spaller, M. R.; Mierke, D. F. "Targeting specific PDZ domains of PSD-95; structural basis for enhanced affinity and enzymatic stability of a cyclic peptide." *Chem. Biol.* **2004**, *11*, 469-473.
32. Li, T.; Saro, D.; Spaller, M. R. "Thermodynamic profiling of conformationally constrained cyclic ligands for the PDZ domain." *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1385-1388.

33. Udugamasooriya, G.; Saro, D.; Spaller, M. R. "Bridged peptide macrocycles as ligands for PDZ domain proteins." *Org. Lett.* **2005**, *7*, 1203-1206.
34. LeBlanc, B. W.; Iwata, M.; Mallon, A. P.; Rupasinghe, C. N.; Goebel, D. J.; Marshall, J.; Spaller, M. R.; Saab, C. Y. "A cyclic peptide targeted against PSD-95 blocks central sensitization and attenuates thermal hyperalgesia." *Neuroscience* **2010**, *167*, 490-500.
35. Gee, S. H.; Sekely, S. A.; Lombardo, C.; Kurakin, A.; Froehner, S. C.; Kay, B. K. "Cyclic peptides as non-carboxyl-terminal ligands of syntrophin PDZ domains." *J. Biol. Chem.* **1998**, *273*, 21980-21987.
36. Arkin, M. R.; Wells, J. A. "Small-molecule inhibitors of protein-protein interactions: progressing towards the dream." *Nat. Rev. Drug Discov.* **2004**, *3*, 301-317.
37. Wells, J. A.; McClendon, C. L. "Reaching for high-hanging fruit in drug discovery at protein-protein interfaces." *Nature* **2007**, *450*, 1001-1009.
38. Fujii, N.; Haresco, J. J.; Novak, K. A.; Gage, R. M.; Pedemonte, N.; Stokoe, D.; Kuntz, I. D.; Guy, R. K. "Rational design of a nonpeptide general chemical scaffold for reversible inhibition of PDZ domain interactions." *Bioorg. Med. Chem. Lett.* **2007**, *17*, 549-552.
39. Fujii, N.; Shelat, A.; Hall, R. A.; Guy, R. K. "Design of a selective chemical probe for class I PDZ domains." *Bioorg. Med. Chem. Lett.* **2007**, *17*, 546-548.
40. Mayasundari, A.; Ferreira, A. M.; He, L.; Mahindroo, N.; Bashford, D.; Fujii, N. "Rational design of the first small-molecule antagonists of NHERF1/EBP50 PDZ domains." *Bioorg. Med. Chem. Lett.* **2008**, *18*, 942-945.
41. Joshi, M.; Vargas, C.; Boisguerin, P.; Diehl, A.; Krause, G.; Schmieder, P.; Moelling, K.; Hagen, V.; Schade, M.; Oschkinat, H. "Discovery of low-molecular-weight ligands for the AF6 PDZ domain." *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 3790-3795.
42. Shan, J.; Shi, D. L.; Wang, J.; Zheng, J. "Identification of a specific inhibitor of the dishevelled PDZ domain." *Biochemistry* **2005**, *44*, 15495-15503.
43. Grandy, D.; Shan, J.; Zhang, X.; Rao, S.; Akunuru, S.; Li, H.; Zhang, Y.; Alpatov, I.; Zhang, X. A.; Lang, R. A.; Shi, D. L.; Zheng, J. J. "Discovery and characterization of a small molecule inhibitor of the PDZ domain of dishevelled." *J. Biol. Chem.* **2009**, *284*, 16256-16263.
44. Thorsen, T. S.; Madsen, K. L.; Rebola, N.; Rathje, M.; Anggono, V.; Bach, A.; Moreira, I. S.; Stuhr-Hansen, N.; Dyhring, T.; Peters, D.; Beuming, T.; Haganir, R.; Weinstein, H.; Mülle, C.; Stromgaard, K.; Ronn, L. C.; Gether, U. "Identification of a small-molecule inhibitor of the PICK1 PDZ domain that inhibits hippocampal LTP and LTD." *Proc. Natl. Acad. Sci.* **2010**, *107*, 413-418.
45. Florio, S. K.; Loh, C.; Huang, S. M.; Iwamaye, A. E.; Kitto, K. F.; Fowler, K. W.; Treiberg, J. A.; Hayflick, J. S.; Walker, J. M.; Fairbanks, C. A.; Lai, Y. "Disruption of nNOS-PSD95 protein-protein interaction inhibits acute thermal hyperalgesia and chronic mechanical allodynia in rodents." *Br. J. Pharmacol.* **2009**, *158*, 494-506.
46. Iskenderian-Epps, W. S.; Imperiali, B. "Modulation of Shank3 PDZ domain ligand-binding affinity by dimerization." *ChemBioChem* **2010**, *11*, 1979-1984.
47. Paduch, M.; Biernat, M.; Stefanowicz, P.; Derewenda, Z. S.; Szweczuk, Z.; Otlewski, J. "Bivalent peptides as models for multimeric targets of PDZ domains." *ChemBioChem* **2007**, *8*, 443-452.

48. Chi, C. N.; Bach, A.; Gottschalk, M.; Kristensen, A. S.; Stromgaard, K.; Jemth, P. "Deciphering the kinetic binding mechanism of dimeric ligands using a potent plasma-stable dimeric inhibitor of postsynaptic density protein-95 as an example." *J Biol Chem* **285**, 28252-60.
49. Sainlos, M.; Tigaret, C.; Poujol, C.; Olivier, N. B.; Bard, L.; Breillat, C.; Thiolon, K.; Choquet, D.; Imperiali, B. "Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization." *Nat. Chem. Biol.* **2011**, *in press*.
50. Paduch, M.; Biernat, M.; Stefanowicz, P.; Derewenda, Z. S.; Szewczuk, Z.; Otlewski, J. "Bivalent peptides as models for multimeric targets of PDZ domains." *Chembiochem* **2007**, **8**, 443-52.
51. Long, J.-F.; Tochio, H.; Wang, P.; Fan, J.-S.; Sala, C.; Niethammer, M.; Sheng, M.; Zhang, M. "Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95." *J. Mol. Biol.* **2003**, **327**, 203-214.

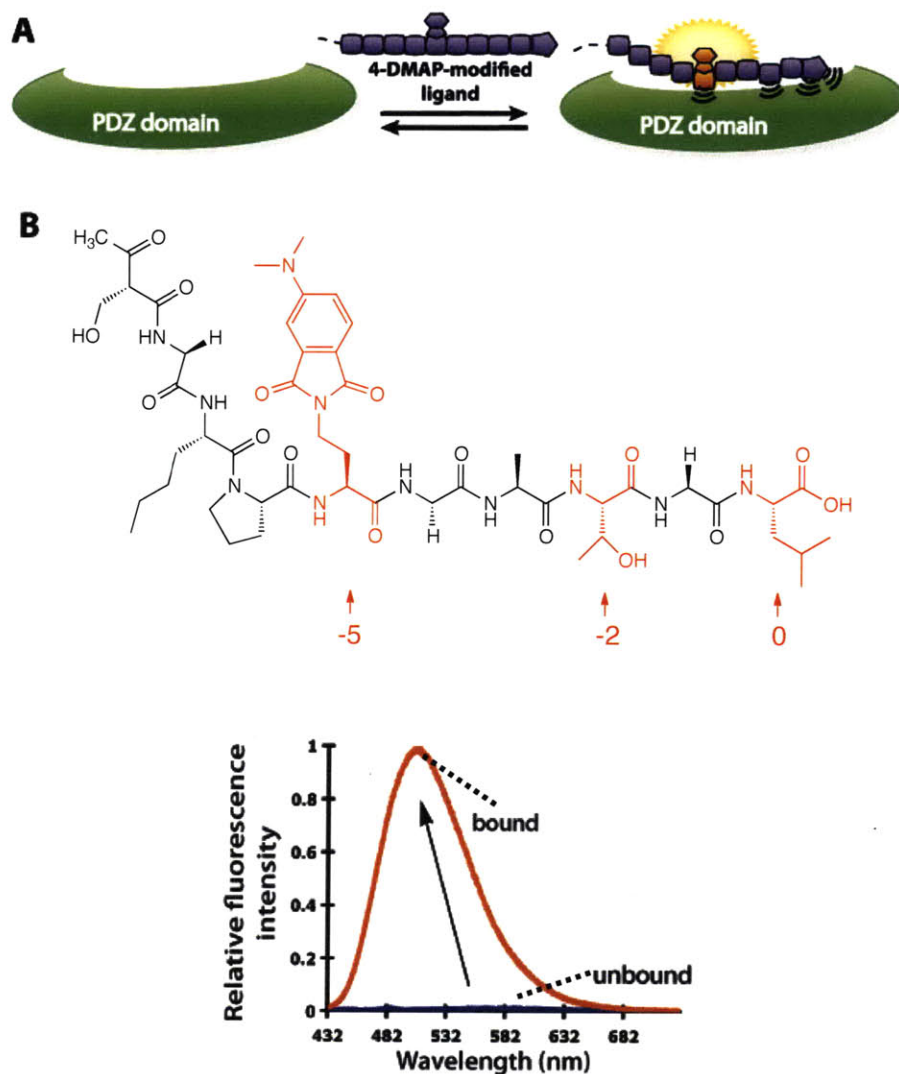
## Chapter 2

# Peptide-Based Fluorogenic Probes for Dynamic Studies of Class I PDZ Domain-Mediated Interactions

Portions of this chapter have been published in the *Journal of the American Chemical Society*<sup>1</sup> as noted in the text. Copyright © 2009, American Chemical Society.

### Introduction

The development of a general strategy for the design, synthesis and evaluation of fluorogenic probes for PDZ domains is presented here. These probes are based in part on natural ligand sequences and incorporate an environment-sensitive fluorophore to report specific binding events. This strategy necessitates appropriate placement of the solvatochromic group within the ligand sequence, such that changes in the local environment due to interactions with cognate domains result in significant modifications of the fluorescence properties of the fluorophore, as depicted in Figure 2.1. Chromophores of the dimethylaminophthalimide family (4-DMAP,<sup>2</sup> 6-DMN,<sup>3</sup> and 4-DMN<sup>4</sup>) exemplify these properties and have previously been applied to the development of sensors for small protein domains such as SH2 (phosphotyrosine-binding) domains.<sup>3</sup> In this case, placement of the fluorophore adjacent to the conserved binding determinants has afforded sensors with modest (~10-fold) signal enhancement. Alternatively, solvatochromic fluorophore-based probes have been developed for pSer/pThr-binding domains (14-3-3 proteins),<sup>2, 5</sup> calmodulin,<sup>4</sup> and class II MHC proteins,<sup>6</sup> by replacement of a conserved hydrophobic aromatic ligand residue which is known to interact with a defined site on the cognate binding protein. However, not all protein interaction domains allow for such a replacement approach to be utilized, as the domains may either not rely upon defined hydrophobic interactions or, alternatively, the native hydrophobic residues may constitute critical specificity determinants that cannot be modified. PDZ domains typically fall into both of these categories and therefore necessitate the development of a different, more general strategy to generate optimized fluorescent probes.



**Figure 2.1.** Probe design and function. A) Optimized fluorescent ligands report binding to partner PDZ domains by dramatic changes in emission intensity and maximum  $\lambda_{em}$ . B) Representative structure of fluorogenic probes with 4-DMAP fluorophore (on highlighted residue), numbering convention for PDZ domain ligands (0 = C-terminal residue, residue numbers decrease from there; critical residues 0 and (-)2 are highlighted in red). Fluorescence emission spectra of probe **12** in the unbound (purple) and Shank3-bound (orange) states.

There is significant interest in probing PDZ domain-mediated interactions of postsynaptic density scaffold proteins in neurons. More specifically, the target proteins in this study are PSD-95 and Shank3, representative members of the MAGUK/SAP<sup>7</sup> and Shank/ProSAP<sup>8</sup> families respectively. Both proteins contain one or several class I PDZ

domains. In contrast to previous studies, where a defined hydrophobic pocket in the cognate domains immediately suggested an ideal position to insert the environment-sensitive fluorophores,<sup>9-12</sup> PDZ domains accommodate a wide variety of ligands conforming to specific consensus sequences defined by residues at positions 0 and (-)2, and do not provide conserved residues to substitute without affecting critical specificity determinants. (A schematic of the PDZ domain ligand numbering convention is shown in Figure 2.1B.) Furthermore, apart from the hydrophobic pocket, which accommodates the C-terminal aliphatic ligand residue, PDZ domains lack any obvious non-polar site flanking the binding groove. Hence a systematic approach was utilized to identify probes that efficiently report binding. The design strategy included: (1) development of a peptide library to screen for the optimal fluorophore placement, and (2) screening of probes with the fixed fluorophore position for improved affinity and specificity.

## **Results and Discussion**

### ***2-1. Design and Synthesis of a Library of Probes<sup>1</sup>***

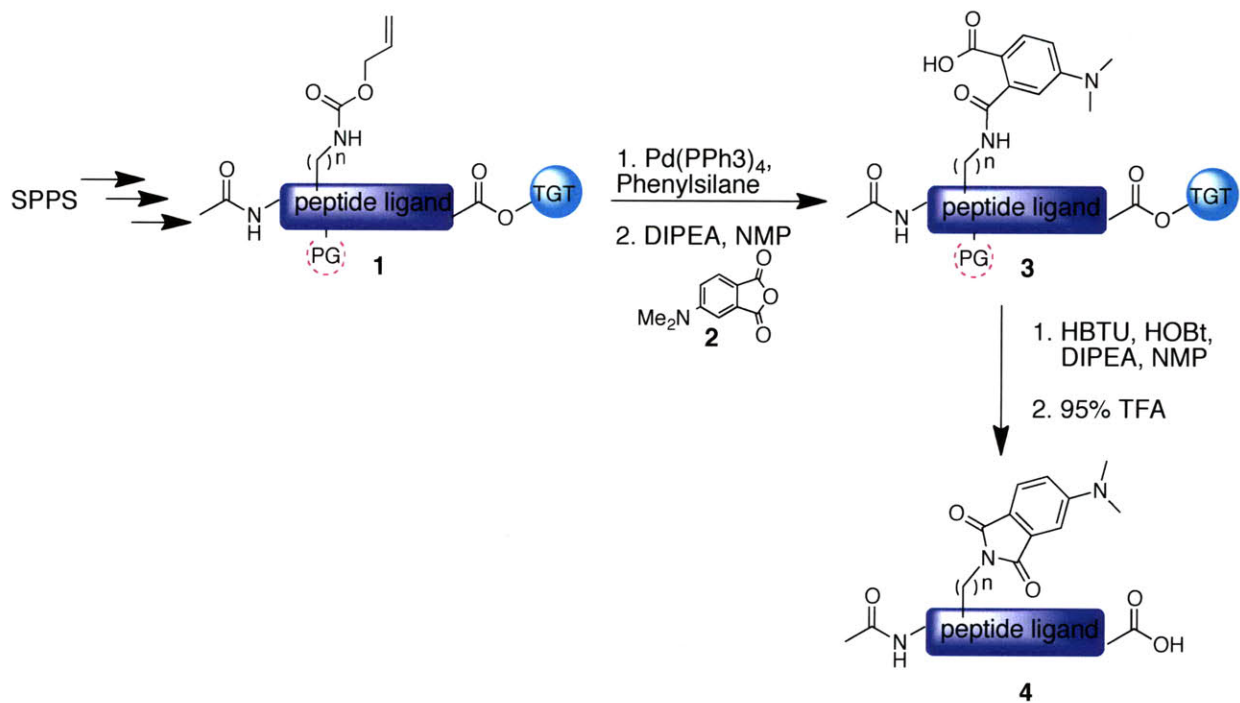
The design of the probes for PDZ domains is based in part on natural PDZ domain ligand sequences, and incorporates a 4-DMAP environment-sensitive fluorophore to report specific binding events, as depicted in Figure 2.1. Use of natural ligand sequences affords native sequence specificity. Although structural studies indicate that key interactions with PDZ domains primarily involve the last four ligand residues, some data suggest that additional residues may also be important,<sup>13</sup> which is why studies focused on the terminal 10 residues of the native ligands in the targeted probes. The protein PDZ domains and their partner ligands (from which C-terminal peptides were fashioned) utilized in these studies are summarized in Table 2.1.

Entry	PDZ Domain(s)	Partner Ligand(s)
1	PSD-95 PDZ12	Stargazin, NR2A, CRIPT
2	PSD-95 PDZ3	Stargazin
3	SAP-97 PDZ12	Stargazin, NR2A, CRIPT
4	SAP-97 PDZ3	Stargazin
5	SAP-102 PDZ12	Stargazin, NR2A, CRIPT
6	SAP-102 PDZ3	Stargazin
7	Shank3 PDZ	GluR1
8	PICK1 PDZ	
9	GRIP PDZ45	
10		mGluR7a

**Table 2.1.** PDZ domain and ligands used in these studies. PDZ domains in black are all class I. Class II PDZ domains used for control testing are shown in red, as is a control ligand.

Fluorescent peptides based on the terminal 10 residues of PDZ domain ligands were synthesized, with incorporation of the fluorophore at varying sites along the peptide, as delineated in Scheme 2.1. Specifically, the peptides were synthesized by standard 9-fluorenylmethoxycarbonyl (Fmoc)-based solid phase peptide synthesis (SPPS) on the highly acid labile TGT resin, with incorporation of the commercially available, allyloxycarbonyl (Alloc) orthogonally protected amino acids Fmoc-Dap-(Alloc), Fmoc-Dab-(Alloc), and Fmoc-Orn-(Alloc). Following Alloc deprotection, 4-(*N,N*-dimethylamino)-phthalic anhydride **2**, synthesized according to the literature and Scheme 2.2,<sup>2</sup> was coupled onto the peptide to yield **3**. The resulting peptide was treated with HBTU/HOBt and base for ring closure, and released from the TGT resin with a free C-terminal carboxylic acid. Subsequent HPLC purification yielded the desired probe **4**, confirmed by analytical HPLC and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) analysis.

**Scheme 2.1.** Synthesis of general fluorescent probe **4** ( $n = 1$ , Dap;  $n = 2$ , Dab;  $n = 3$ , Orn).



**Scheme 2.2.** Synthesis of 4-(*N,N*-dimethylamino)-phthalic anhydride **2**.

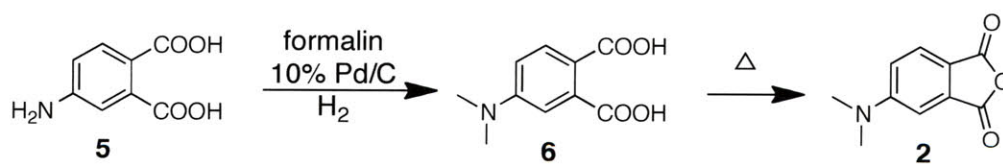
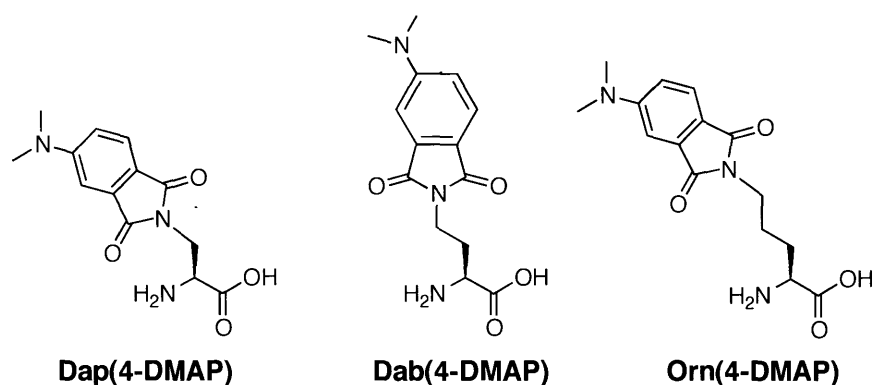




Figure 2.2 shows the Dap, Dab, and Orn residues with the appended 4-DMAP chromophore. All of these residues were incorporated into the probes for screening, since tuning of the linker length between the peptide backbone and the environment-sensitive fluorophore is critical to probe function. Results from the screens, presented later in this chapter, will demonstrate the necessity of linker length tuning.



**Figure 2.2.** Three 4-DMAP-derived residues incorporated into the PDZ domain probes. Each differs solely by the number of methylene units in the linker between the backbone of the peptide and the fluorophore. Dap = diaminopropionic acid, Dab = diaminobutyric acid, Orn = ornithine.

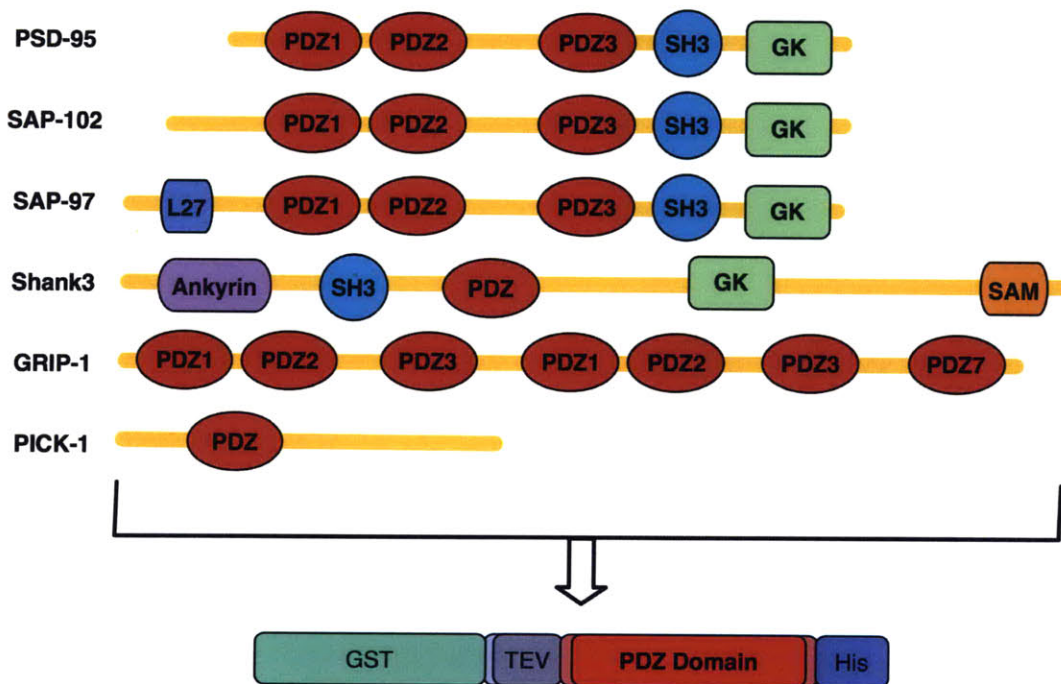
All of the peptides that were synthesized and analyzed are compiled in Table 2.2, along with their MALDI-TOF mass characterization and analytical HPLC retention times. Purity (>95%) of the fluorescent peptides was judged by analytical HPLC. These peptide sequences and the choices of fluorophore position will be discussed in section 2.3.

#	Origin	Peptide Sequence <sup>a</sup>	Formula	Mass expected	Mass <sup>b</sup> obtained [MH] <sup>+</sup>	t <sub>R</sub> <sup>c</sup>
1	Stargazin	Ac-NTANRRRT $\alpha$ V-COOH	C <sub>55</sub> H <sub>83</sub> N <sub>20</sub> O <sub>19</sub>	1332.7	1333.5	21.8
2	Stargazin	Ac-NTANRR $\alpha$ TPV-COOH	C <sub>56</sub> H <sub>88</sub> N <sub>20</sub> O <sub>18</sub>	1328.7	1330.5	21.7
3	Stargazin	Ac-NTANR $\alpha$ TPV-COOH	C <sub>54</sub> H <sub>83</sub> N <sub>17</sub> O <sub>19</sub>	1273.6	1247.6	22.1
4	Stargazin	Ac-NTANR $\beta$ TPV-COOH	C <sub>55</sub> H <sub>83</sub> N <sub>17</sub> O <sub>19</sub>	1287.6	1288.8	22.4
5	Stargazin	Ac-NTAN $\alpha$ RTPV-COOH	C <sub>54</sub> H <sub>83</sub> N <sub>17</sub> O <sub>19</sub>	1273.6	1275.6	21.8
6	Stargazin	Ac-NTAN $\beta$ RTPV-COOH	C <sub>55</sub> H <sub>83</sub> N <sub>17</sub> O <sub>19</sub>	1287.6	1288.6	22.6
7	Stargazin	Ac-NTAN $\gamma$ RTPV-COOH	C <sub>56</sub> H <sub>87</sub> N <sub>17</sub> O <sub>19</sub>	1301.6	1302.7	22.6
8	Stargazin	Ac-NTA $\beta$ RRTPV-COOH	C <sub>57</sub> H <sub>91</sub> N <sub>19</sub> O <sub>18</sub>	1329.7	1330.7	23.0
9	Stargazin	Ac-NT $\beta$ NRRTTPV-COOH	C <sub>58</sub> H <sub>92</sub> N <sub>20</sub> O <sub>19</sub>	1372.7	1373.6	21.7
10	Stargazin	Ac-N $\beta$ ANRRTPV-COOH	C <sub>57</sub> H <sub>90</sub> N <sub>20</sub> O <sub>18</sub>	1342.7	1343.5	21.8
11	Stargazin	Ac- $\beta$ TANRRTPV-COOH	C <sub>57</sub> H <sub>91</sub> N <sub>19</sub> O <sub>18</sub>	1329.7	1331.0	22.7
12	Stargazin	Ac-NRRNTAN $\beta$ RTTPV-COOH	C <sub>71</sub> H <sub>115</sub> N <sub>27</sub> O <sub>23</sub>	1713.9	1714.4	22.7
13	NR2a	Ac-NRRKK $\alpha$ P $\beta$ IESDV-COOH	C <sub>78</sub> H <sub>128</sub> N <sub>24</sub> O <sub>23</sub>	1769.0	1769.9	23.7
14	CRIP1	Ac-DTKN $\beta$ KQTSV-COOH	C <sub>57</sub> H <sub>90</sub> N <sub>16</sub> O <sub>21</sub>	1334.7	1335.7	20.5
15	GluR1	Ac-NRRSG $\alpha$ P $\beta$ GATGL-COOH	C <sub>65</sub> H <sub>104</sub> N <sub>22</sub> O <sub>20</sub>	1512.8	1513.7	24.1
16	Stargazin	Ac-NRRNTAN $\beta$ RTRPR-NR2	C <sub>74</sub> H <sub>123</sub> N <sub>33</sub> O <sub>22</sub>	1826.0	1826.5	19.5
17	mGluR7a	Ac-AKKKYVS $\beta$ NNLVI-COOH	C <sub>79</sub> H <sub>126</sub> N <sub>20</sub> O <sub>21</sub>	1690.9	1692.6	24.0
18	Stargazin	Ac- $\phi$ NTANRRTPV-COOH	C <sub>56</sub> H <sub>90</sub> N <sub>20</sub> O <sub>20</sub>	1362.7	1363.4	21.1
19	GluR1	Ac- $\phi$ NRRSG $\alpha$ PLGATGL-COOH	C <sub>66</sub> H <sub>108</sub> N <sub>22</sub> O <sub>21</sub>	1545.2	1545.2	26.1
20	CRIP1	Ac- $\phi$ DTKNYKQTSV-COOH	C <sub>61</sub> H <sub>91</sub> N <sub>16</sub> O <sub>23</sub>	1416.7	1417.7	22.2

**Table 2.2.** Peptides synthesized and characterized, numbered 1 through 20. <sup>a</sup> Native methionine residues were replaced by norleucine (Nle), nF = *p*-nitrophenylalanine,  $\alpha$  = Dap(4-DMAP),  $\beta$  = Dab(4-DMAP),  $\gamma$  = Orn(4-DMAP), the *NRR* sequence was added on the N-terminus of peptides to improve solubility at high concentrations, as low solubility of the probe can lead to high background fluorescence signal due to aggregation. <sup>b</sup> Masses were determined by MALDI-TOF spectroscopy. <sup>c</sup> Retention time from analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 min followed by 5-95% acetonitrile containing 0.1% TFA over 35 min in water containing 0.1% TFA at a flow rate of 1 mL/min).

## 2-2. Design and Expression of Protein Constructs<sup>1</sup>

The target proteins of the first study are PSD-95 and Shank3, which contain class I PDZ domains. In order to facilitate the development of probes for these PDZ domains, each was cloned as a recombinant GST-fusion protein, shown schematically in Figure 2.3. In addition, SAP102 and SAP97 PDZ domains (which share a high sequence homology with PSD95 PDZ domains and can accommodate the same ligands)<sup>14</sup> and class II PDZ domains from PICK1 and GRIP1 were cloned as recombinant GST-fusion constructs. Residues flanking the PDZ domains were incorporated into these constructs, in order to ensure proper folding of the isolated PDZ domains.



**Figure 2.3.** Recombinant PDZ domains. PDZ domain-containing proteins used in this study and common GST-fusion construct for the expressed domains.

Studies have shown that tandem PDZ domains (closely grouped pairs of PDZ domains) structurally and functionally affect one another. For example, both GRIP1 PDZ1 and PDZ5 require the presence of neighboring domains (PDZ2 and PDZ4, respectively) to fold properly and bind to target ligands.<sup>15-17</sup> Interactions have been observed, to a lesser extent, between

PDZ1 and PDZ2 of PSD-95.<sup>18</sup> For this reason, tandem domains were cloned together in order to preserve their native structure and avoid artifacts due to improper folding of artificially isolated domains.

### ***2-3. Overall Screening Strategy and Validation of Probes<sup>1</sup>***

A two-pronged screening approach to identify optimal fluorescent probes for class I PSD-95 and Shank3 PDZ domains was developed. In the first phase, a library of peptides derived from the C-terminal sequence of Stargazin and incorporating the 4-DMAP fluorophore (Table 2.2) was screened. Stargazin, an auxiliary sub-unit of  $\alpha\pm$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, was chosen for its capacity to interact with the three domains of PSD-95.<sup>19</sup> Furthermore, given the general promiscuity of PDZ domains, it was anticipated that a library incorporating the consensus motif T-X-V would also enable screening for other class I PDZ domains such as Shank3. The terminal ten residues of Stargazin were chosen in order to best preserve native specificity. Optimal fluorophore positioning was assessed in a step-wise process, where native residues were first replaced with 4-DMAP-derived amino acids (except for positions 0 and (-)2, which are essential for canonical binding) and evaluated for fluorescence increases. In the second iteration, the fluorescence increase was further optimized by adjusting the length of the diamino acid linker bearing 4-DMAP in the most sensitive positions (Dap, n=1; Dab, n=2 or Orn, n=3). Of note, the 4-DMAP fluorophore was initially chosen for its minimal size and the ease of synthesis of a peptide library by on-resin derivatization; hence, at this stage other environment-sensitive fluorophores, 6-DMN and 4-DMN, were also tested at the optimal position with similar and shorter linkers to compensate for ring extension, albeit with lower relative fluorescence increases (data not shown). This may be due to the fact that these fluorophores are too large to be accommodated, even though they show some enhanced properties relative to the 4-DMAP chromophore including maximum fluorescence increase and stability. These results with the various environment-sensitive fluorophores highlight the need for the screening approach developed here.

The peptide library was screened for ability to report binding, which was evaluated by ratiometric analysis of 4-DMAP fluorescence emission intensities in the absence and presence of the target PDZ domains (see equation (1) and Figure 2.4 for clarification). PSD95-12/3

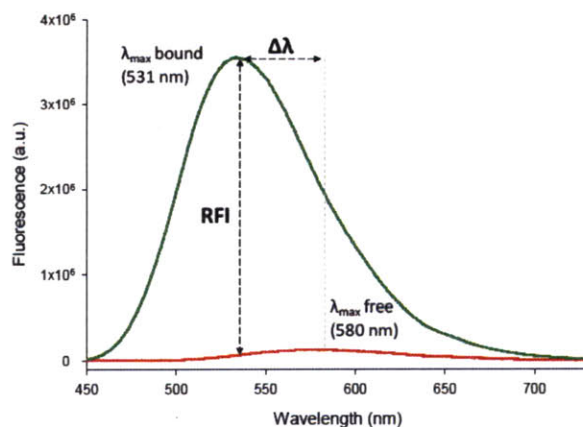
and Shank3 PDZ domain were tested, and Table 2.3 shows the results of the screen for each of these proteins.

$$RFI = \frac{I_{bound}(\lambda_{\max \text{ bound}})}{I_{free}(\lambda_{\max \text{ bound}})} \quad (1)$$

The maximum relative fluorescence increase for all PDZ constructs was obtained for peptide **6**; thus, the optimal position and linker length are -5 and Dab (n=2), respectively. This probe was particularly sensitive to PSD95-3; more than an eighty-fold increase in emission intensity was observed. Interestingly, similar overall results to PSD-95 were obtained with other MAGUK PDZ domains, which share a high sequence homology with PSD95-12/3 (SAP102-12/3 and SAP97-12/3, Figure 2.5) and can accommodate the same ligands.<sup>7, 14</sup> This suggests that MAGUK PDZ domains share common structural features that can be sensed by the probes. Peptide **6** was also sensitive to the Shank3 PDZ domain, another class I PDZ domain, as shown in Figure 2.6. A comparison of Figure 2.6 with the top portion of Figure 2.5 demonstrates that the fluorescence increases for the single PDZ domain of Shank3 are lower in general than those of the single PDZ3 of the MAGUK proteins, most likely because Stargazin is not the optimal ligand sequence for this particular class I domain. This factor will be taken into account in the second step of the screening process.

Control experiments were employed to further validate the screening approach. As expected, no significant fluorescence increase was observed for the Stargazin peptides with class II PDZ domains from PICK1 and GRIP1 (Figure 2.7). This confirmed that the probe binding is driven by specific PDZ domain-cognate ligand interactions. Next, the relative fluorescence increase was measured for full-length PSD-95 and compared to its constituent PDZ domains, PSD95-12 and PSD95-3 using peptide **6** (data not shown). The similarity of the fluorescence output for the native protein and the domain constructs indicated the relevance of the results obtained with isolated domains and validated the screening approach.

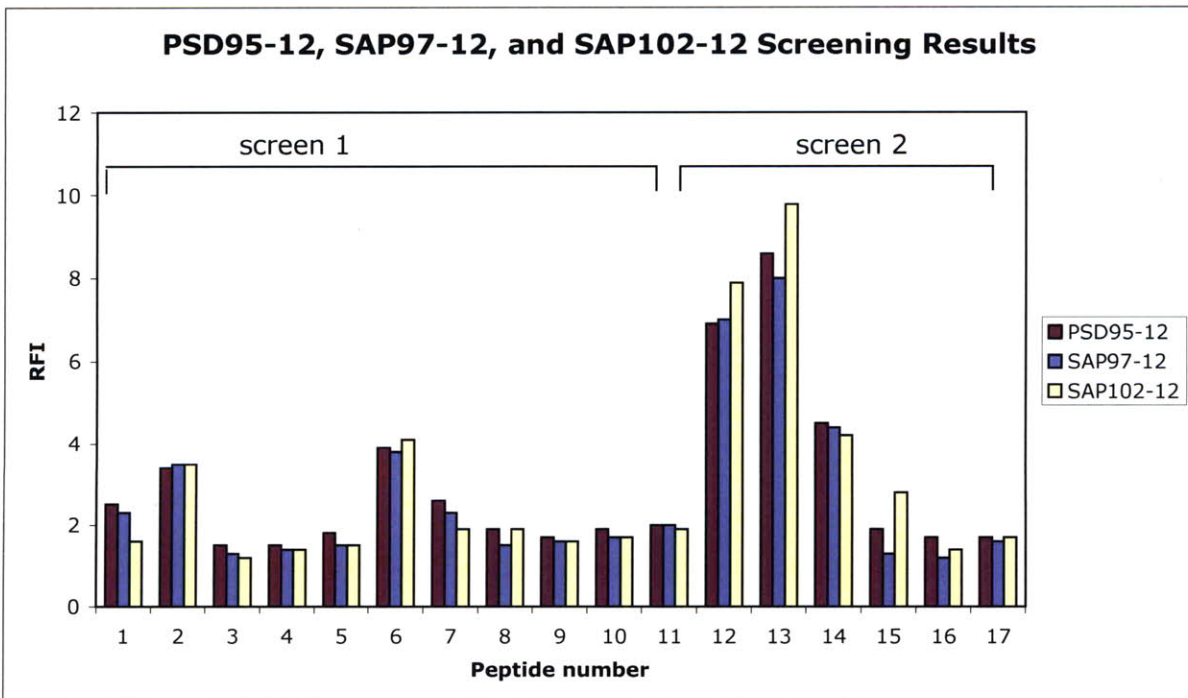
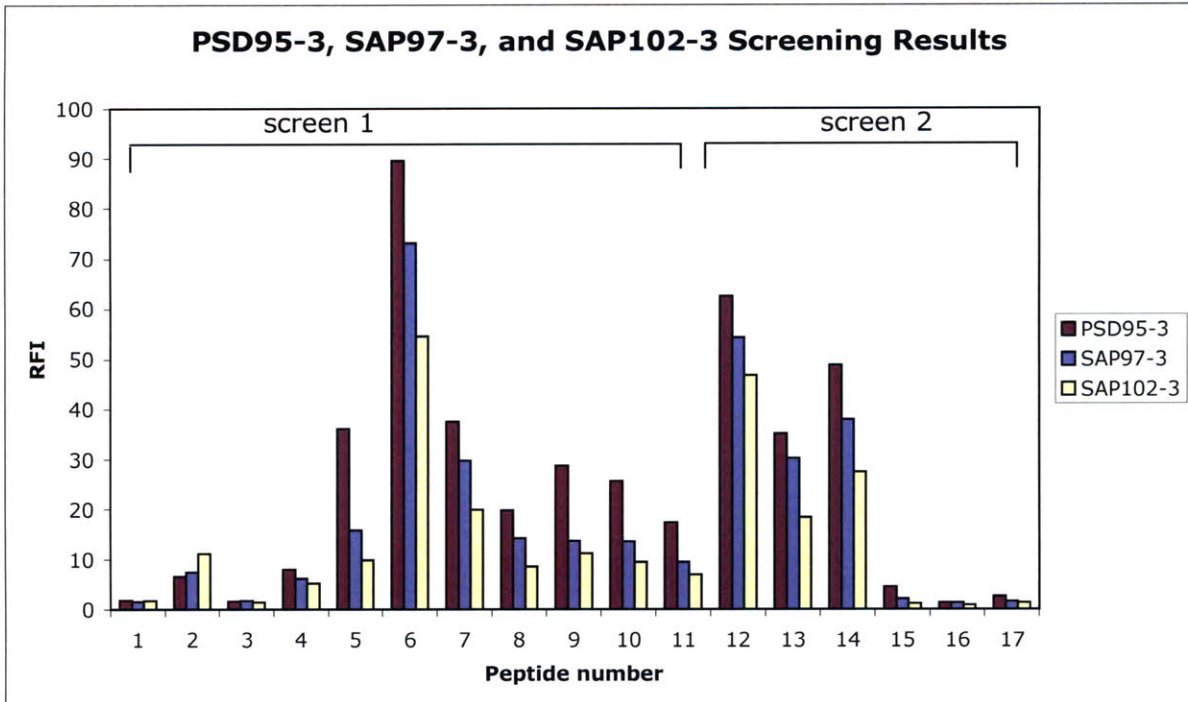




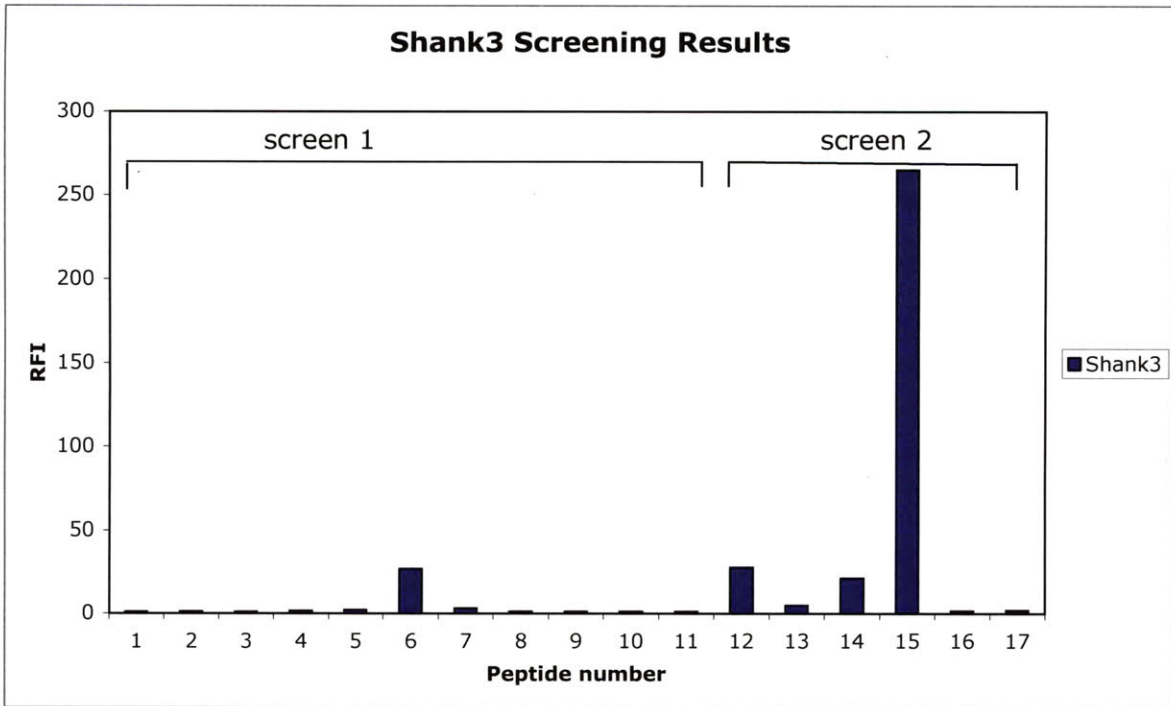
**Figure 2.4.** Definition of relative fluorescence increase (RFI). (A) Example of fluorescence increase obtained with peptide 12. Red line: peptide 12 in PBS (2  $\mu$ M); green line: peptide 12 in PBS (2  $\mu$ M) + PSD95-3 (20  $\mu$ M). The relative fluorescence increase is defined by the ratio given in equation (1), where the fluorescence intensities  $I^{\text{bound}}$  and  $I^{\text{free}}$  are obtained by averaging at least three independent measurements and are calculated over a range of 5 nm centered around  $\lambda_{\text{max, bound}}$ .

#	Peptide Sequence	PSD95-3 RFI	PSD95-12 RFI	Shank3 RFI
1	Ac-NTANRRRTT $\alpha$ V-COOH	1.9	2.5	1.0
2	Ac-NTANRR $\alpha$ TPV-COOH	6.6	3.4	1.1
3	Ac-NTANR $\alpha$ TPV-COOH	1.7	1.5	1.0
4	Ac-NTANR $\beta$ TPV-COOH	8.0	1.5	1.5
5	Ac-NTAN $\alpha$ RTPV-COOH	36.1	1.8	2.0
6	Ac-NTAN $\beta$ RTPV-COOH	<b>89.5</b>	<b>3.9</b>	<b>26.7</b>
7	Ac-NTAN $\gamma$ RTPV-COOH	37.6	2.6	3.0
8	Ac-NTA $\beta$ RRTPV-COOH	19.8	1.9	1.2
9	Ac-NT $\beta$ NRRTTPV-COOH	28.7	1.7	1.3
10	Ac-N $\beta$ ANRRTPV-COOH	25.6	1.9	1.3
11	Ac- $\beta$ TANRRTPV-COOH	17.4	2.0	1.3

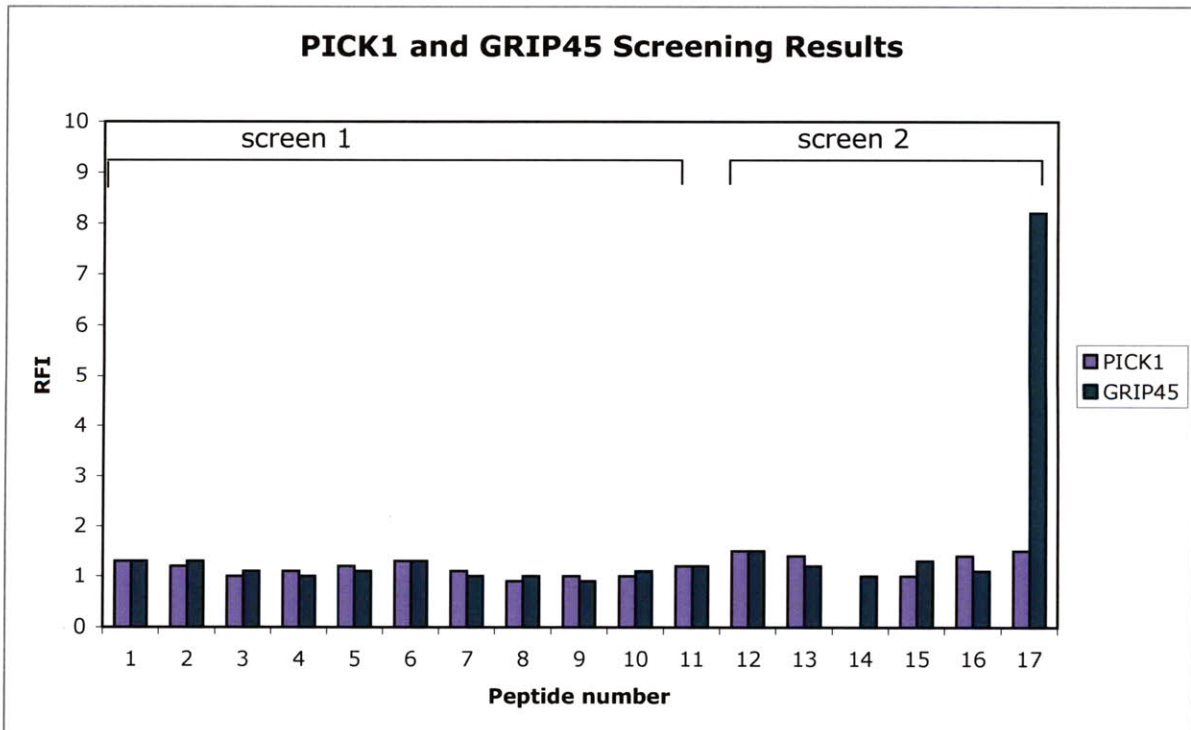
**Table 2.3.** Fluorescence increases of the C-terminal Stargazin-derived peptide library with PSD-95 and Shank3 PDZ domains.  $\alpha$  = Dap(4-DMAP),  $\beta$  = Dab(4-DMAP),  $\gamma$  = Orn(4-DMAP).



**Figure 2.5.** PSD-95, SAP-97, and SAP-102 screen results. Top: relative fluorescence increases (RFI) with the third PDZ domain of PSD-95 and related SAP-97 and SAP-102 proteins. Bottom: relative fluorescence increases (RFI) with the first and second PDZ domains of PSD-95 and related SAP-97 and SAP-102 proteins.



**Figure 2.6.** Relative fluorescence increases (RFI) with the Shank3 PDZ domain.



**Figure 2.7.** Relative fluorescence increases (RFI) with the PICK1 and GRIP45 PDZ domains.



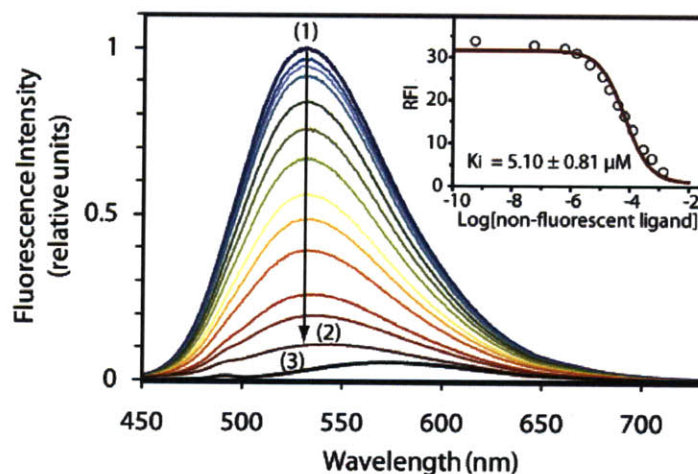
In the second step of the design process, the goal was to generate probes with improved signaling properties by addressing the affinity and specificity of the ligands for each domain. Therefore, C-terminal sequences of CRIPT,<sup>14, 20</sup> NR2a,<sup>14</sup> and GluR1,<sup>21</sup> known binders for PSD95-3, PSD95-12, and Shank, respectively, were selected as ligands for the second series of probes. The series also included negative controls such as a class II ligand, derived from mGluR7a,<sup>22</sup> and a Stargazin sequence in which critical residues were replaced by Arg (residues 0 and (-)2). Applying the results of the library screen, the 4-DMAP fluorophore was appended to a Dab residue at the (-)5 position, and probes were tested with each construct. The Asn-Arg-Arg (NRR) sequence was appended to each of the peptides used in the second screen, in order to aid in overall solubility. The affinities of the new sensors were obtained by fluorescence-based titrations and are reported in Table 2.4, together with relative fluorescence increases. The studies reveal that **12** and **14** are the best probes for PSD95-3, **13** for PSD95-12, and **15** for Shank3. Remarkably, probe **15** exhibited a 265-fold fluorescence increase with the Shank3 PDZ domain. Control sequences, **16** and **17**, exhibited negligible fluorescence increases (and poor affinities when measured) for the targeted domains, thus confirming class specificity of the probes. Figures 2.5, 2.6, and 2.7 summarize the results from both the first and second screening processes.

#	Origin	Peptide Sequence	PSD95-3		PSD95-12		Shank3	
			K <sub>D</sub> (μM)	RFI	K <sub>D</sub> (μM)	RFI	K <sub>D</sub> (μM)	RFI
12	Stargazin	Ac-NRRNTANβRTTPV-COOH	0.57±0.23	62.6	5.99±0.76	6.9	10.00±2.44	27.6
13	NR2a	Ac-NRRKKλPβIESDV-COOH	1.66±0.18	35.1	1.50±0.26	8.6	23.10±0.68	4.9
14	CRIPT	Ac-DTKNβKQTSV-COOH	1.21±0.20	48.8	9.63±1.34	4.5	12.40±4.30	21.0
15	GluR1	Ac-NRRSGλPβGATGL-COOH	89.5±20.6	4.6	20.45±3.17	2.0	0.17±0.02	265.0
16	Control	Ac-NRRNTANβRTRPR-COOH	>>50	1.4	n.d.	1.7	n.d.	1.5
17	mGluR7a	Ac-AKKKYVSβNNLVI-COOH	n.d.	2.7	n.d.	1.7	n.d.	1.9

**Table 2.4.** Binding constants and fluorescence increases for peptides with the optimally positioned 4-DMAP fluorophore. β = Dab(4-DMAP), λ = Norleucine (replaces Methionine in native sequence), NRR- added to improve solubility, n.d. = not determined.

Together, these results indicate that the selectivity of the native sequences was preserved after insertion of the fluorophore. Furthermore, for each domain, K<sub>D</sub> values of ligands identically derived with 4-DMAP (both position and linker) correlated with relative

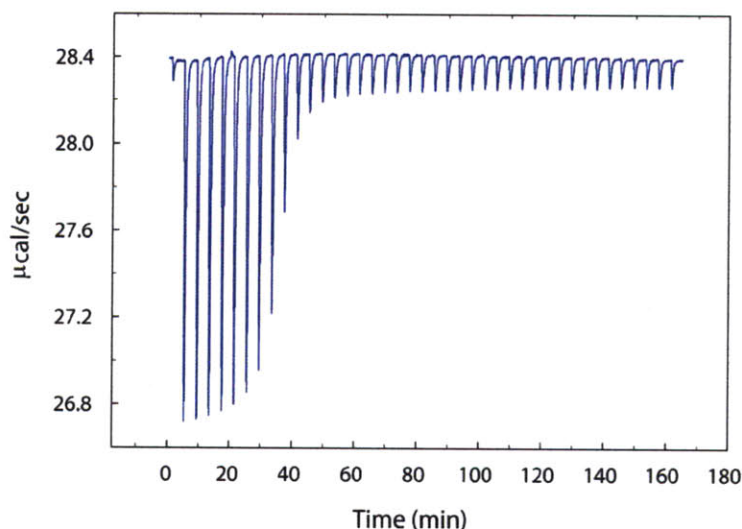
fluorescence increases when screened under the same conditions. Next, competition experiments with non-fluorescent peptides were carried out on probes **12** and **15** with their respective targeted PDZ domains. As shown in Figure 2.8, gradual displacement of probe **12** from PSD95-3 by a non-fluorescent peptide derived from Stargazin can be readily observed by decrease of the fluorescence signal (similar results were obtained with the Shank3 probe peptide **15**). Affinity constants of the unlabeled peptides could be calculated from these competitive titrations, showing a stronger global affinity of the labeled peptides. While further demonstrating the reversibility and specificity of the fluorescent probe-PDZ domain interactions, these experiments also illustrate the potential of these tools for the screening of high affinity ligands for target PDZ domains.



**Figure 2.8.** Representative competition titration. Unlabeled Stargazin peptide vs. probe **12** / PSD95-3 complex. Main graph: Fluorescence emission spectra of a 5  $\mu$ M solution of peptide **12** and PSD95-3 after sequential additions of 0 (1) to 200 eq. (2) of the non-fluorescent peptide; (3) 5  $\mu$ M solution of peptide **12**. Inset: Relative fluorescence increases at corresponding concentrations of unlabeled competing peptide (data, fit, and  $K_i$ ).

Finally, the binding affinities were also independently evaluated by using a non-fluorescence-based method, isothermal titration calorimetry (ITC, see Figure 2.9 for representative data). In this case, the aims were to validate the fluorescence-based results and also to determine the potential influence of the fluorophore on the probe affinity. The binding constant of peptide **12** for PSD95-3, determined by ITC, was similar to that obtained by fluorescence titration, while the dissociation constants of non-fluorescent peptides derived

from CRIPT and Stargazin were in agreement with reported data (Table 2.5). Overall, these results confirmed the method and revealed an improved affinity of peptide **12** in comparison to the native peptide sequence.



**Figure 2.9.** ITC thermogram. Peptide **12** versus PSD95-3.

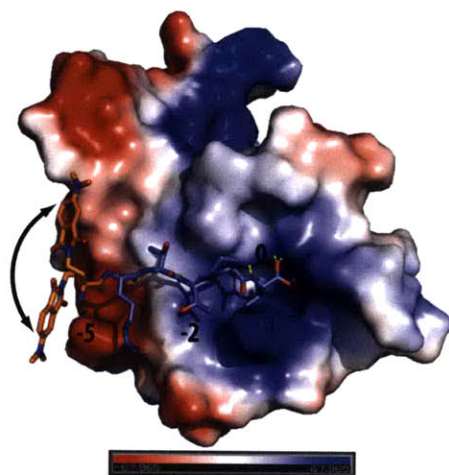
#	Origin	PSD95-3 RFI	$K_D$ ( $\mu\text{M}$ ) by ITC	$K_D$ ( $\mu\text{M}$ ) from lit <sup>a</sup>
12	Stargazin	Ac-NRRNTAN $\beta$ RTTPV-COOH	0.59 $\pm$ 0.60	-
18	Stargazin	Ac- $\phi$ NTANRRRTPV-COOH	7.49 $\pm$ 1.07	7-15
19	CRIPT	Ac- $\phi$ DTKNYKQTSV-COOH	3.44 $\pm$ 0.50	3.39

**Table 2.5.** Summary of ITC control results and comparison with literature values. <sup>a</sup>Fluorescence polarization values for Stargazin<sup>19</sup>- and CRIPT<sup>14</sup>-derived peptides.

The affinity improvement was attributed to interactions of the fluorophore with a non-polar patch of the domains that, when combined with the other ligand interactions, translates into high fluorescence increases and stronger affinities. However, attempts at correlating these results with structural information did not allow identification of obvious shared elements for the three sets of domains that would account for the observed response from the 4-DMAP chromophore at the (-)5 position. By considering the X-ray crystallographic structure of



PSD95-3 complexed with a ligand (PDB entry 1TP5) as a starting point and modeling the C-terminal 6 residues of probe **12**, the regions accessible by 4-DMAP on the PDZ domain surface could be defined, as illustrated in Figure 2.10.



**Figure 2.10.** Structural modeling of PSD-95 PDZ3 electrostatic potential map and fluorogenic peptide. Red = negative, white = hydrophobic, blue = positive.

Surprisingly, the fluorophore is located at the edge of the PDZ domain binding groove and does not seem to be in the vicinity of any clearly defined hydrophobic patch. In this context, it is worth noting that another PSD95-3 structure complexed with a 9-residue long CRIPT peptide<sup>23</sup> (PDB entry 1BE9) failed to show sufficient electron density to resolve the location of the ligand (-)5 Tyr residue, thus stressing the lack of a predefined strong interaction of the PDZ domain with an aromatic ligand residue at this position. Overall, these structural considerations underscore the importance and efficiency of the general screening approach that is presented in the design of environment-sensitive probes for interaction domains such as PDZ domains that do not allow for a rational ligand residue replacement strategy.

## Conclusions

A systematic library approach to the design of fluorogenic peptide-based probes for PDZ domains has been developed and successfully applied to the generation of probes. This method can be applied to other classes of domains and can be readily implemented for the screening of sensors as well as high affinity competing ligands. The new probes exhibit large fluorescence increases while conserving native sequence specificity. Thus, the probes constitute valuable tools for *in vitro* studies of PDZ domain-mediated interactions by enabling a direct readout of binding. These probes have been further evolved and extended to the study of other PDZ domains both *in vitro* and *in vivo*, which will be discussed in later chapters of this thesis.

## Experimental Methods

**Peptide Synthesis (SPPS), Fluorophore Synthesis, and Fluorophore Insertion.** Peptides were obtained either manually or with an automated synthesizer (Advanced ChemTech automated synthesizer model 396  $\Omega$ ) by using standard Fmoc-based solid phase peptide synthesis procedures. Typically manual synthesis was performed on a 0.02 to 0.04 mmol scale with Fmoc-Val/Leu/Ile-NovaSyn® TGT resin (0.2 mmol/g) or Fmoc-PAL-PEG-PS resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents (6 equivalents each), and DIPEA (12 equivalents) in DMF or NMP. Coupling steps were conducted with a 50 mM solution of amino acid (~1 mL per 100 mg of resin) for 1 hour. Fmoc group removal was performed with a 20% solution of 4-methylpiperidine in DMF (vol/vol) for 3 x 5 minutes. After removal of the N-terminal Fmoc group, the resulting free amine was capped with an acetyl group by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

Peptides incorporating the 4-DMAP environment-sensitive fluorophore were synthesized using the on-resin derivatization approach previously reported.<sup>24</sup> An Alloc-protected diaminoacid (either Fmoc-Dap(Alloc)-OH, Fmoc-Dab(Alloc)-OH or Fmoc-Orn(Alloc)-OH from AnaSpec) was initially inserted during the SPPS steps. After capping of the N-terminal amino group of the peptide sequence, the Alloc-protecting group was removed. The resin was resuspended in dry dichloromethane (20 mM in peptide) with a stream of N<sub>2</sub> bubbling through the solution for 5 minutes. Pd(PPh<sub>3</sub>)<sub>4</sub> (0.8 eq.) and phenylsilane (25 eq.) were then added to the solution, which was maintained under N<sub>2</sub> bubbling for another 15 minutes. The resin was then washed with dichloromethane and DMF, and the degassing/deprotection cycle was repeated two additional times. The resin containing the free amine was reacted with a solution of the anhydride form of 4-DMAP (4-dimethylaminophthalenedicarboxylic anhydride, 2 equiv., 50 mM) and DIPEA (4 equiv.) in NMP. The mixture was allowed to stir overnight. The resin was then washed with DMF and dichloromethane and ring closure was performed by using a solution of HBTU/HOBt (6 equiv., 50 mM) in NMP or DMF with DIPEA (12 eq.) for 2 hours. The resin was then washed with DMF and dichloromethane. The coupling/washing cycle was repeated two additional times in order to achieve full ring closure.

Peptides were deprotected and cleaved from the resin with a TFA/H<sub>2</sub>O/TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours. TFA was evaporated until the peptides precipitated. The peptides were then triturated and precipitated in cold ether before purification by reverse phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250x20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 350 nm. Peptides were stored lyophilized at -80 °C until use.

**Peptide Characterization and Quantification.** Peptide identity was confirmed by MALDI-TOF mass spectroscopy (MALDI-TOF, PerSeptive Biosystems Voyager) using DHB as a matrix. Purity was assessed by analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 min followed by 5-95% acetonitrile containing 0.1% TFA over 35 min in water containing 0.1% TFA at a flow rate of 1 mL/min). The final peptides were quantified using molar extinction coefficients of either 4-DMAP ( $\epsilon_{421\text{nm}} = 6480 \text{ M}^{-1}\text{cm}^{-1}$  in water) or *p*-nitrophenylalanine ( $\epsilon_{280\text{nm}} = 12500 \text{ M}^{-1}\text{cm}^{-1}$ ).

**PDZ Domain Cloning.** The cDNA encoding for the genes used for PDZ domain subcloning were provided by Daniel Choquet (Laboratoire Physiologie Cellulaire de la Synapse; PSD-95, SAP-102, GRIP-1, PICK-1 and Shank-3) and Morgan Sheng (Picower Institute; SAP-97). The PDZ domains of interest were cloned into the pGEX-4T-2 vector (GE Healthcare) using BamHI and XhoI restriction sites. Primers were designed to subclone the PDZ domain(s) as defined by the ExPASy Proteomics Server (UniProtKB/Swiss-Prot database) plus 5 to 10 extra flanking amino acids on either terminus. The primers also incorporate a TEV cleavage site between GST and the domain(s) as well as a FLAG tag at the C-terminus. The PCR-amplified inserts were digested with either BamHI/XhoI or BamHI/EcoRI and ligated into a BamHI/XhoI- or BamHI/EcoRI-digested pGEX-4T-2 vector.

The first two PDZ domains of PSD-95, PSD95-12 (residues 61 to 249, from UniProtKB/Swiss-Prot entry P31016), were PCR-amplified using the primers 5'-CGG GAT CCG AGA ATT TGT ATT TTC AGG GCA TGG AGT ATG AGG AGA TCA CAT TGG and 5'-CCG CTC GAG TTA CTT ATC GTC ATC GTC TTT GTA GTC GGC ATT GCT GGG CTT GGC CAC CTT T.

The third PDZ domain of PSD-95, PSD95-3 (residues 302 to 402, from UniProtKB/Swiss-Prot entry P31016), was PCR-amplified using the primers 5'-CGG GAT CCG AGA ATT TGT ATT TTC AGG GCC TGG GGG AGG AAG ACA TTC CCC GGG and 5'-CCG CTC GAG TTA CTT ATC GTC ATC GTC TTT GTA GTC GGC CTC GAA TCG ACT ATA CTC TTC T.

The first two PDZ domains of SAP-102, SAP102-12 (residues 139 to 338, from UniProtKB/Swiss-Prot entry Q62936), were PCR-amplified using the primers 5'-CGG GAT CCG AGA ACC TGT ACT TCC AGG GCA ATG GCA GTG ATG GCA TGT TCA AGT and 5'-CCG GAA TTC TTA CTT GTC GTC ATC GTC CTT GTA GTC CAT GTC GTT GAG GTG GAG ACT GCC.

The third PDZ domain of SAP-102, SAP102-3 (residues 393 to 493, from UniProtKB/Swiss-Prot entry Q62936), was PCR-amplified using the primers 5'-CGG GAT CCG AGA ACC TGT ACT TCC AGG GC CTG GCT GAG GAA GAC TTT ACC A and 5'-CCG GAA TTC TTA CTT GTC GTC ATC GTC CTT GTA GTC GGA TTC AAA GCG ACT GTA CTC.

The 4<sup>th</sup> and 5<sup>th</sup> PDZ domains of GRIP-1, GRIP1-45 (residues 430 to 663, from UniProtKB/Swiss-Prot entry P97879), were PCR-amplified using the primers 5'-CGG GAT CCG AGA ATT TGT ATT TTC AGG GCT CTA CTA GTC CAC GAG GAA CCA TGA and 5'-CCG CTC GAG TTA CTT ATC GTC ATC GTC TTT GTA GTC CTC TTG CTC ATC TGA GTT ATC TTC A.

The PDZ domain of Shank-3, Shank3 (residues 635 to 749, from UniProtKB/Swiss-Prot entry Q9JLU4), was PCR-amplified using the primers 5'-CGG GAT CCG AGA ACC TGT ACT TCC AGG GCT CAC ACA GTG ATT ATG TCA TTG ATG and 5'-CCG GAA TTC TTA CTT ATC GTC ATC GTC CTT GTA GTC TGC GCC GAG CAC TAT CCT CCT CTG G.

For the cloning of the full PSD-95 protein (UniProtKB/Swiss-Prot entry P31016), PCR-amplified insert was digested with EcoRI/XhoI and ligated into an EcoRI/XhoI-digested pET24a vector (Novagen), which afforded a C-terminal His<sub>6</sub>-tag. PCR amplification was done with the primers 5'-CGG AAT TCA TGG ACT GTC TCT GTA TAG T and 5'-CCG CTC GAG GAG TCT CTC TCG GGC TGG GAC.



**Expression and Purification of PDZ Domain Constructs.** *E. Coli* BL21 codon plus™ (DE3)-RIL or RP cells were transformed with each of the GST fusion protein expression plasmids. Cells were amplified in 0.5 L of an auto-inducing medium<sup>25</sup> (ZYM-5052, see reference for detailed composition) first at 37 °C for 4 hours followed by 12 hours at 16 °C. Cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (100 mM EDTA, 10% glycerol, 1% TritonX-100, 1 mg/mL lysozyme in PBS, pH 7.4) containing protease inhibitor cocktail III (Calbiochem). Cells were lysed by ultrasonication (Branson Sonifier 450 at 50% power with a 40% duty cycle for 4 min at 4 °C). After addition of 1 mL of a 1 mM solution of DTT in water, the lysate was cleared by centrifugation (15,334 g, 40 minutes, 4 °C). The GST-fusion proteins were purified with 5 mL of Glutathione Sepharose™ 4 Fast Flow (GE Healthcare) by batch-binding. The bound proteins were washed with 20 volumes of PBS (pH 7.4) and eluted with 2 volumes of a 10 mM Glutathione in 50 mM Tris (pH 8) solution. Fractions were analyzed by 12% SDS-PAGE followed by Coomassie staining and Western blotting (anti-FLAG). Fractions containing the protein were pooled and dialyzed against PBS (pH 7.4). Protein concentrations were measured using either the BCA assay (Pierce) with BSA as the reference standard or by determining the absorption at 280 nm in 6 M guanidinium chloride. Purified PDZ domains were aliquoted, flash-frozen and stored at -80 °C in PBS (pH 7.4) until use.

**Fluorescence Studies.** Fluorescence spectra were recorded on a Fluoromax 3 instrument (Horiba Jobin Yvon) in 1 cm path length quartz cells (100 mL nominal volume from Starna Cells). All measurements were performed at a constant temperature of 25 °C. Slit widths were 3 nm for excitation and 6 nm for emission. The 4-DMAP fluorophore was excited at 421 nm and the spectra were recorded between 432 nm and 730 nm (0.5 nm increments and 0.1 s integration time).

**Relative Fluorescence Increases.** The fluorescence increase measurements were performed at 25 °C in PBS buffer (pH 7.4) comparing solutions of (1) 2 μM of 4-DMAP-containing peptide alone in PBS and (2) 2 μM of 4-DMAP-containing peptide with 20 μM of PDZ domain construct in PBS. Blanks consisting of PBS for (1) and PBS with 20 μM of the corresponding PDZ domain construct for (2) were subtracted from the respective spectra.

Each final spectrum results from the average of 3 different runs. The fluorescence increase for each series of peptide and PDZ domain(s) was evaluated by comparing the fluorescence emission intensities of (2) and (1) at the wavelength of maximal emission of (2). Ratios were calculated over a 5 nm range centered on the wavelength of maximal emission of (2) and the values were averaged to yield the reported final ratio.

**Fluorescence Titrations.** For each titration, the peptide concentration was kept constant (around 1 to 20  $\mu\text{M}$ ) and the protein concentration was varied from values lower than or close to the anticipated  $K_D$  to saturation of the fluorescence signal increase. A 150 or 130  $\mu\text{L}$  solution in PBS (pH 7.4) [or MES (pH 6.9)] was prepared for each protein concentration. Dissociation constants were evaluated using SPECFIT/32™ Global Analysis System for Windows (version 3.0.39) after averaging at least three independent titrations and taking into account wavelengths from 432 to 730 nm.

**Competition Titrations.** Competition experiments were conducted on 450  $\mu\text{L}$  of a PBS (pH 7.4) solution containing 5  $\mu\text{M}$  fluorescent peptide and 5  $\mu\text{M}$  PDZ domain. The non-fluorescent peptide solution was added sequentially in small volumes to achieve a range of concentrations from 1 nM to 2 mM. The titrations were done at least in triplicate and averaged. The  $K_i$  were obtained by fitting the relative fluorescence increase (RFI, obtained as described above) of the averaged spectrum to the logarithm of the non-fluorescent peptide concentration ( $\text{Log}[\text{nFP}]$ , Molar) using the Competitive Binding: One Site model from GraphPad Prism 5 software with the equations (2) and (3):

$$\log(EC50) = \log\left(10^{\log(K_i)}\left(1 + \frac{[FP]}{K_D^{Fluo}}\right)\right) \quad (2)$$

$$RFI = \min_{RFI} + \frac{\max_{RFI} - \min_{RFI}}{1 + 10^{(\log[nFP] - \log(EC50))}} \quad (3)$$

where  $[FP]$  is the concentration of the fluorescent peptide,  $K_D^{Fluo}$  is the dissociation constant of the fluorescent peptide with the corresponding PDZ domain (constant, same units as  $[FP]$ ),  $\min_{RFI}$  is 1 and  $\max_{RFI}$  is the RFI in the absence of competitive ligand.

***Isothermal Titration Calorimetry.*** Titrations were conducted on a VP-ITC MicroCalorimeter (MicroCal). The PDZ domains were dialyzed against PBS buffer (pH 7.4) or MES buffer (pH 6.9). The protein and peptide samples (prepared in the same buffer) were degassed separately for 10 min under vacuum with stirring. The PDZ domain (20 to 50  $\mu$ M, 1.8 mL) was loaded into the sample cell and a 400 to 1000  $\mu$ M peptide solution (250  $\mu$ L) was drawn into the syringe. Typically 30 to 50 injections were programmed, with the first injection volume set at 0.5  $\mu$ L (1 s injection) and the remaining at 7  $\mu$ L (in 45 injections of 14 s each). The spacing between injections was 240 s. The reference power was 30  $\mu$ cal/s with an initial delay of 1 minute. Experiments were conducted at 25 °C with a stirring speed of 500 rpm. The raw data were collected and analyzed by ORIGIN software (MicroCal). Thermodynamic parameters were determined by non-linear least squares fitting using a One Set of Binding Sites model.

## **Acknowledgements**

Much of this chapter has been previously published, and I would like to thank my coauthor Dr. Matthieu Sainlos for collaboration on this project. The cloning and expression of the PICK and SAP-102 constructs were done by Dr. Matthieu Sainlos, as were the fluorescence tests on these constructs as well as SAP-97 and the full-length PSD-95. Dr. Matthieu Sainlos performed the docking studies described in this chapter. I would also like to acknowledge Dr. Daniel Choquet and Dr. Morgan Sheng for the cDNA they graciously provided for these studies.

## References

1. Sainlos, M.; Iskenderian, W. S.; Imperiali, B. "A general screening strategy for peptide-based fluorogenic ligands: Probes for dynamic studies of PDZ domain-mediated interactions." *J. Am. Chem. Soc.* **2009**, *131*, 6680-6682.
2. Vazquez, M. E.; Rothman, D. M.; Imperiali, B. "A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis." *Org. Biomol. Chem.* **2004**, *2*, 1965-1966.
3. Vazquez, M. E.; Blanco, J. B.; Imperiali, B. "Photophysics and biological applications of the environment-sensitive fluorophore 6-N,N-dimethylamino-2,3-naphthalimide." *J. Am. Chem. Soc.* **2005**, *127*, 1300-1306.
4. Loving, G.; Imperiali, B. "A versatile amino acid analogue of the solvatochromic fluorophore 4-N,N-dimethylamino-1,8-naphthalimide: a powerful tool for the study of dynamic protein interactions." *J. Am. Chem. Soc.* **2008**, *130*, 13630-13638.
5. Vazquez, M. E.; Nitz, M.; Stehn, J.; Yaffe, M. B.; Imperiali, B. "Fluorescent caged phosphoserine peptides as probes to investigate phosphorylation-dependent protein associations." *J. Am. Chem. Soc.* **2003**, *125*, 10150-10151.
6. Venkatraman, P.; Nguyen, T. T.; Sainlos, M.; Bilsel, O.; Chitta, S.; Imperiali, B.; Stern, L. J. "Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells." *Nat. Chem. Biol.* **2007**, *3*, 222-228.
7. Elias, G. M.; Nicoll, R. A. "Synaptic trafficking of glutamate receptors by MAGUK scaffolding proteins." *Trends Cell. Biol.* **2007**, *17*, 343-352.
8. Sheng, M.; Kim, E. "The Shank family of scaffold proteins." *J. Cell Sci.* **2000**, *113*, 1851-1856.
9. Vazquez, M. E.; Nitz, M.; Stehn, J.; Yaffe, M. B.; Imperiali, B. "Fluorescent caged phosphoserine peptides as probes to investigate phosphorylation-dependent protein associations." *J. Am. Chem. Soc.* **2003**, *125*, 10150-10151.
10. Vazquez, M. E.; Rothman, D. M.; Imperiali, B. "A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis." *Org. Biomol. Chem.* **2004**, *2*, 1965-1966.
11. Venkatraman, P.; Nguyen, T. T.; Sainlos, M.; Bilsel, O.; Chitta, S.; Imperiali, B.; Stern, L. J. "Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells." *Nat. Chem. Biol.* **2007**, *3*, 222-228.
12. Loving, G.; Imperiali, B. "A versatile amino acid analogue of the solvatochromic fluorophore 4-N,N-dimethylamino-1,8-naphthalimide: A powerful tool for the study of dynamic protein interactions." *J. Am. Chem. Soc.* **2008**, *130*, 13630-13638.
13. Hung, A. Y.; Sheng, M. "PDZ domains: structural modules for protein complex assembly." *J. Biol. Chem.* **2002**, *277*, 5699-5702.
14. Lim, I. A.; Hall, D. D.; Hell, J. W. "Selectivity and promiscuity of the first and second PDZ domains of PSD-95 and synapse-associated protein 102." *J. Biol. Chem.* **2002**, *277*, 21697-21711.
15. Long, J.; Wei, Z.; Feng, W.; Yu, C.; Zhao, Y. X.; Zhang, M. "Supramodular nature of GRIP1 revealed by the structure of its PDZ12 tandem in complex with the carboxyl tail of Fras1." *J. Mol. Biol.* **2008**, *375*, 1457-1468.
16. Zhang, Q.; Fan, J. S.; Zhang, M. "Interdomain chaperoning between PSD-95,Dlg, and Zo-1 (PDZ) domains of glutamate receptor-interacting proteins." *J. Biol. Chem.* **2001**, *276*, 43216-43220.

17. Feng, W.; Shi, Y.; Li, M.; Zhang, M. "Tandem PDZ repeats in glutamate receptor-interacting proteins have a novel mode of PDZ domain-mediated target binding." *Nat. Struct. Biol.* **2003**, *10*, 972-978.
18. Long, J.-F.; Tochio, H.; Wang, P.; Fan, J.-S.; Sala, C.; Niethammer, M.; Sheng, M.; Zhang, M. "Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95." *J. Mol. Biol.* **2003**, *327*, 203-214.
19. Dakoiji, S.; Tomita, S.; Karimzadegan, S.; Nicoll, R. A.; Brecht, D. S. "Interaction of transmembrane AMPA receptor regulatory proteins with multiple membrane associated guanylate kinases." *Neuropharmacology* **2003**, *45*, 849-856.
20. Niethammer, M.; Valtschanoff, J. G.; Kapoor, T. M.; Allison, D. W.; Weinberg, R. J.; Craig, A. M.; Sheng, M. "CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90." *Neuron* **1998**, *20*, 693-707.
21. Uchino, S.; Wada, H.; Honda, S.; Nakamura, Y.; Ondo, Y.; Uchiyama, T.; Tsutsumi, M.; Suzuki, E.; Hirasawa, T.; Kohsaka, S. "Direct interaction of post-synaptic density-95/Dlg/ZO-1 domain-containing synaptic molecule Shank3 with GluR1  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptor." *J. Neurochem.* **2006**, *97*, 1203-1214.
22. Hirbec, H.; Perestenko, O.; Nishimune, A.; Meyer, G.; Nakanishi, S.; Henley, J. M.; Dev, K. K. "The PDZ proteins PICK1, GRIP, and syntenin bind multiple glutamate receptor subtypes. Analysis of PDZ binding motifs." *J. Biol. Chem.* **2002**, *277*, 15221-15224.
23. Doyle, D. A.; Lee, A.; Lewis, J.; Kim, E.; Sheng, M.; MacKinnon, R. "Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ." *Cell* **1996**, *85*, 1067-1076.
24. Sainlos, M.; Imperiali, B. "Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization." *Nat. Protoc.* **2007**, *2*, 3201-3209.
25. Studier, F. W. "Protein production by auto-induction in high-density shaking cultures." *Prot. Expr. Purif.* **2005**, *41*, 207-234.

# Chapter 3

## Development and Validation of a Novel C-terminal Caging Strategy for PDZ Domain Ligands

### Introduction

Photolabile chemical caging groups are valuable tools to achieve spatiotemporal control of biological interactions. Molecules of interest can be rendered biologically inert via modification with a caging group, a photocleavable moiety that acts like a protecting group. The active species of the molecule is liberated by light-driven unmasking of the critical functional group, allowing the active species to bind its intended target in the biological pathway of interest.<sup>1</sup> The application of caged molecules allows for temporal control in the study of dynamic biological events that cannot be fully addressed with conventional genetic or immunological approaches. Caging strategies have been successfully used for many biological applications, including the masking of neurotransmitters and of key phosphorylated residues in peptide and protein sequences.<sup>2</sup>

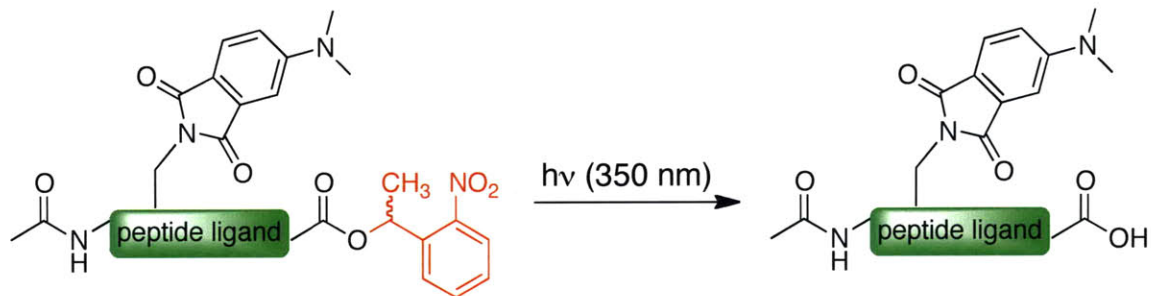
PDZ domains are involved in many critical transient interactions that contribute to the assembly and localization of macromolecular complexes in signal transduction pathways.<sup>3</sup> In chapter 2 the design and development of fluorescence-based probes to detect PDZ domain-mediated interactions was discussed. Toward the goal of gaining additional chemical approaches for studying the various modes of PDZ domain regulation *in vivo*, the utility of the probes was extended through the pursuit of functionally caged variants. A successful caging strategy would render PDZ domain ligands efficacious in probing the spatiotemporal dynamics of the complex protein-protein interactions involving these ligands. In this chapter studies toward the development and validation of caging strategies to control the function and interactions of PDZ domains *in vitro* will be presented.

## **Results and Discussion**

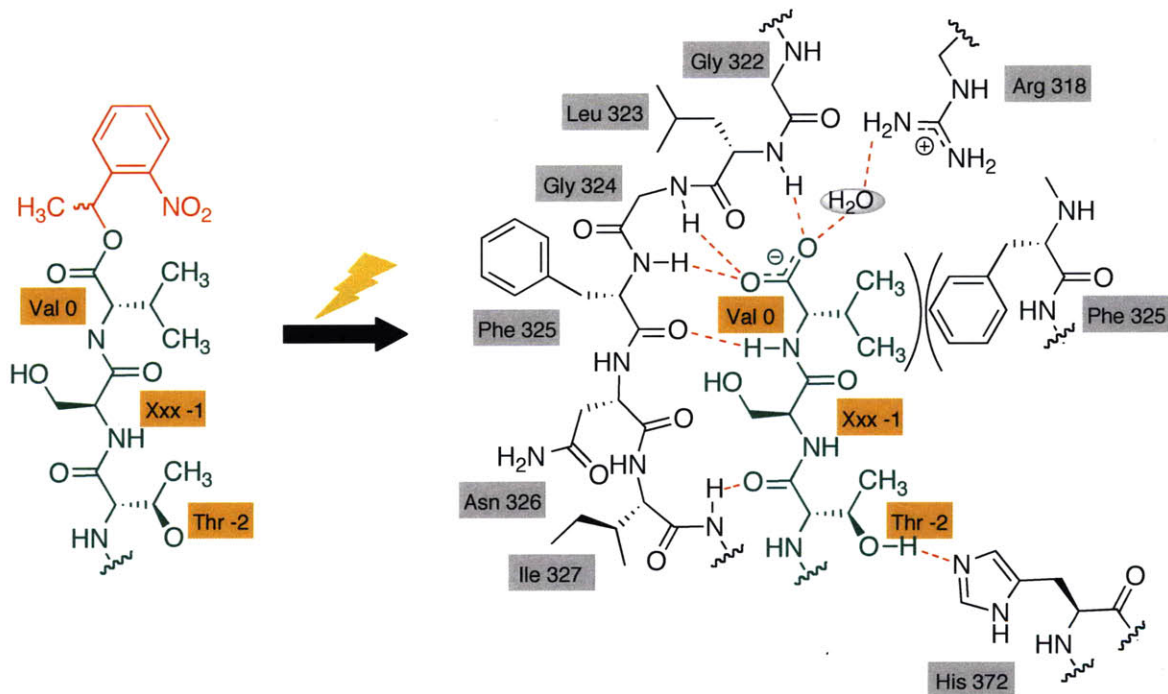
### ***3-1. C-terminal caging strategy***

We developed an approach to investigate PDZ domain-mediated interactions similar to the approach taken to study the role of phosphorylation in protein-protein interactions via caging of peptides.<sup>4, 5</sup> For example, in protein phosphorylation studies, the phosphorylated residues are critical components for binding to proteins in signaling networks, and the masking of these essential binding determinants can hinder binding. In a similar manner, the binding of PDZ domains to their cognate ligands also relies upon the presence of an essential binding determinant; in this case, it is the C-terminal carboxyl group of the PDZ domain ligand. In light of these similarities, a strategy based on caging of the critical C-terminal carboxyl of the peptide-based ligand was pursued. Caging of the C-terminal carboxyl should allow for temporal control of binding to a cognate PDZ domain, since this moiety is essential for binding. Figure 3.1 depicts the residue contacts that can be formed between a PDZ domain and its partner ligand only in the absence of a C-terminal caging group. In these studies, the advantageous properties of environment sensitive probes provided an excellent tool for studying the efficiency and utility of this strategy. Chapter 2 of this thesis discusses the utility of these tools to directly trigger release of and report binding of peptide-based ligands with cognate PDZ domains. Therefore, the caged ligands were designed as an extension of the fluorogenic probes discussed in Chapter 2.

A)



B)



**Figure 3.1.** Schematic of control of PDZ domain binding by caging of the ligand's critical C-terminal acid. A) release of caging group (red) from fluorogenic ligand. B) Example of residue contacts between PDZ ligand (green) and PDZ domain binding groove (black), which can be formed only after uncaging. Note that the C-terminal cage (red) should obliterate binding due to steric hindrance and interference with the hydrogen bond contacts that are normally made by the C-terminal ligand acid.

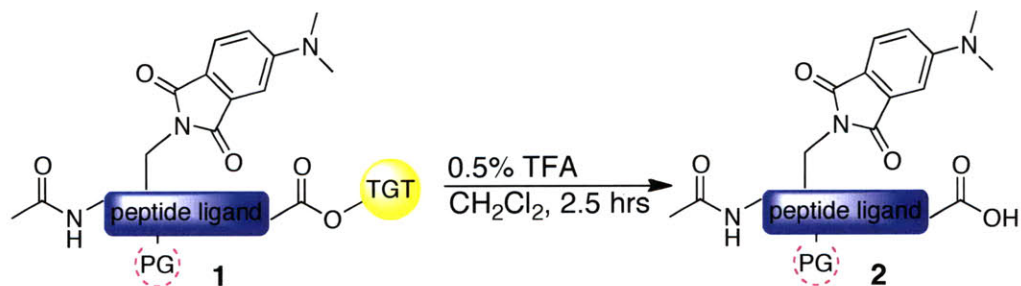


### 3-2. Design and synthesis of caged C-terminal ligands

Structural studies have shown that the canonical binding of PDZ domain ligands essentially involves C-terminal motifs of target proteins, typically through conserved interactions with the hydrophobic side chain and carboxylate group of the terminal amino acid as well as with the side chain of the residue in the (-)2 position (see Figure 3.1B). The nature of the residue at the (-)2 position differs from one domain to another, and as such is commonly used to distinguish domains into classes.<sup>6</sup> Given the universal essentiality and intrinsic chemical characteristics of the C-terminal carboxylate group, this functional group was selected for caging. Addition of a photolabile group at this position would both abolish key hydrogen-bonding interactions that are critical in binding to the domains, and introduce steric hindrance from the caging group itself. This caging strategy should therefore prevent, or significantly reduce, interactions with cognate PDZ domains.

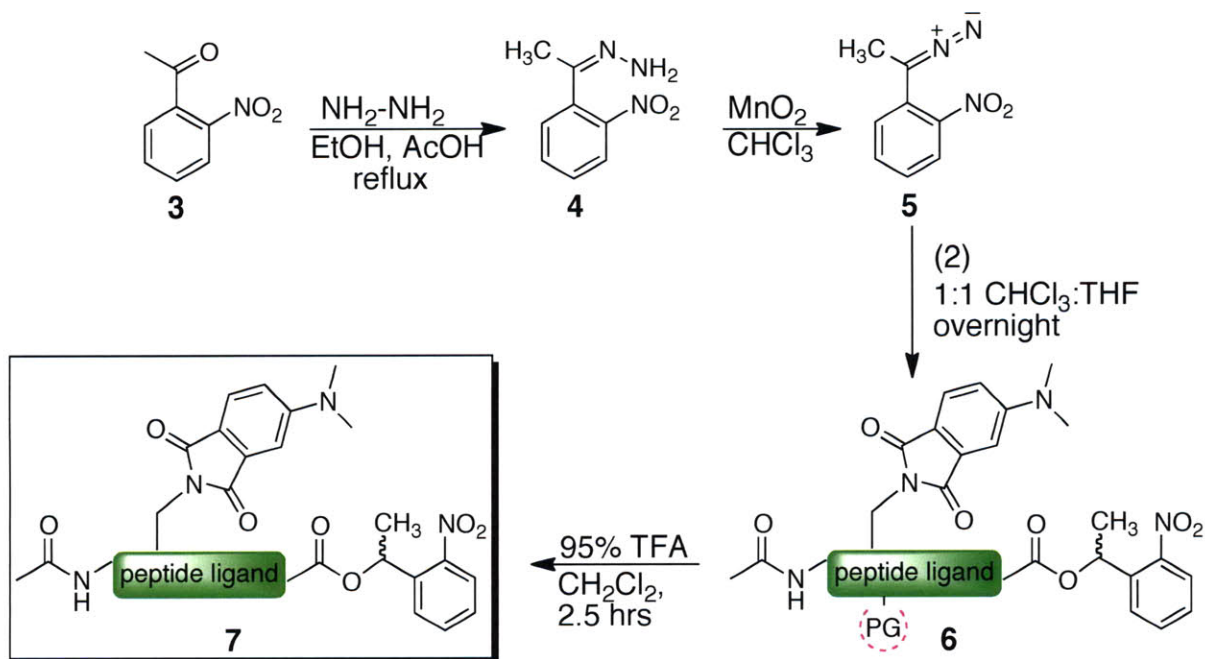
The 1-(2-nitrophenyl)ethyl (NPE) cage has been utilized successfully to cage phosphate moieties of peptides<sup>4,5</sup> and many other biologically relevant functional groups, due to the relative ease of synthesis of NPE derivatives, as well as the stability and favorable uncaging properties of this molecule.<sup>1</sup> Therefore, focus was placed on incorporating the NPE cage into the PDZ ligands. The general synthetic strategy used to obtain ligands caged at their C-terminal carboxylic acid consisted of introducing the photolabile group after solid phase peptide synthesis, by relying upon activation of the caging group rather than the C-terminal carboxylate moiety. This strategy avoids the potential for racemization of the terminal amino acid. Fluorogenic peptides were first synthesized by standard Fmoc-based SPPS methods on TGT resin, which includes a mildly acid sensitive linker, to yield peptide **1** (see Scheme 3.1), as described in Chapter 2. The resulting peptides were cleaved from the resin with 0.5% TFA, preserving the side-chain protecting groups in order to allow for selective coupling of the activated caging group. Concentration *in vacuo* afforded peptide **2** for subsequent caging.

**Scheme 3.1.** Synthesis of free acid fluorogenic peptide **2** for caging.



Activation of the NPE caging group was necessary for coupling to the C-terminal acid (see Scheme 3.2). Toward this end, 2-nitroacetophenone, **3**, was converted to the corresponding hydrazone, **4**, by reaction with hydrazine monohydrate and glacial acetic acid in ethanol. Hydrazone **4** was then oxidized with MnO<sub>2</sub> to yield 1-(2-nitrophenyl)diazoethane, **5**, which upon reaction with **2** yielded peptide **6**. The desired caged fluorogenic peptide **7** was obtained after standard TFA-based deprotection of side-chain protecting groups and HPLC purification. Products were confirmed by analytical HPLC and ESI-MS analysis.

**Scheme 3.2.** Synthesis of NPE-caged probe 7.



In this study, we chose to use the probe sequences derived from the two synaptic proteins Stargazin and GluR1, which were optimized for binding to PSD-95 and Shank3 PDZ domains, respectively, as described in the work presented in Chapter 2. For the purpose of this study, these two PDZ domains function as representative class I PDZ domains with which to investigate the new caging strategy.

The Stargazin and GluR1-5b caged probes (**8** and **10**) are shown in Table 3.1 below, along with details on the characterization by ESI-MS. It should be noted that the caging group is susceptible to cleavage during matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) laser desorption, which is why ESI-MS was used for mass characterization. Table 3.1 also shows the characterization of non-caged probes **9** and **11** (i.e. probes that were never caged).

#	Origin	Peptide Sequence <sup>a</sup>	Formula	t <sub>R</sub> <sup>b</sup>	Mass expected	Mass <sup>c</sup> obtained [MH] <sup>+</sup>
8	Stargazin	Ac-NTANβRTTPV-COONPE	C <sub>63</sub> H <sub>93</sub> N <sub>18</sub> O <sub>21</sub>	31.8	1438.5	1438.9*
9	Stargazin	Ac-NTANβRTTPV-COOH	C <sub>55</sub> H <sub>86</sub> N <sub>17</sub> O <sub>19</sub>	23.9	1289.4	1288.6
10	GluR1	Ac-NRRSGλPβGATGL-COONPE	C <sub>73</sub> H <sub>112</sub> N <sub>23</sub> O <sub>22</sub>	31.9	1663.8	832.1**
11	GluR1	Ac-NRRSGλPβGATGL-COOH	C <sub>65</sub> H <sub>105</sub> N <sub>22</sub> O <sub>20</sub>	26.9	1514.7	1513.7

**Table 3.1.** Characterization of caged fluorogenic probes and their parent noncaged peptides. <sup>a</sup> β = Dab(4-DMAP). <sup>b</sup> t<sub>R</sub> = retention time. Purity was assessed by analytical reverse phase HPLC using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 minutes followed by 5-95% acetonitrile containing 0.1% TFA over 50 minutes in water containing 0.1% TFA at a flow rate of 1 mL / min). All peptides were more than 95% pure as judged by analytical HPLC. <sup>c</sup> The peptide-based ligand identity was either confirmed by MALDI-TOF analysis or by ESI-MS. (\*) = ESI-MS, 1+ peak. (\*\*)= ESI-MS, 2+ peak.

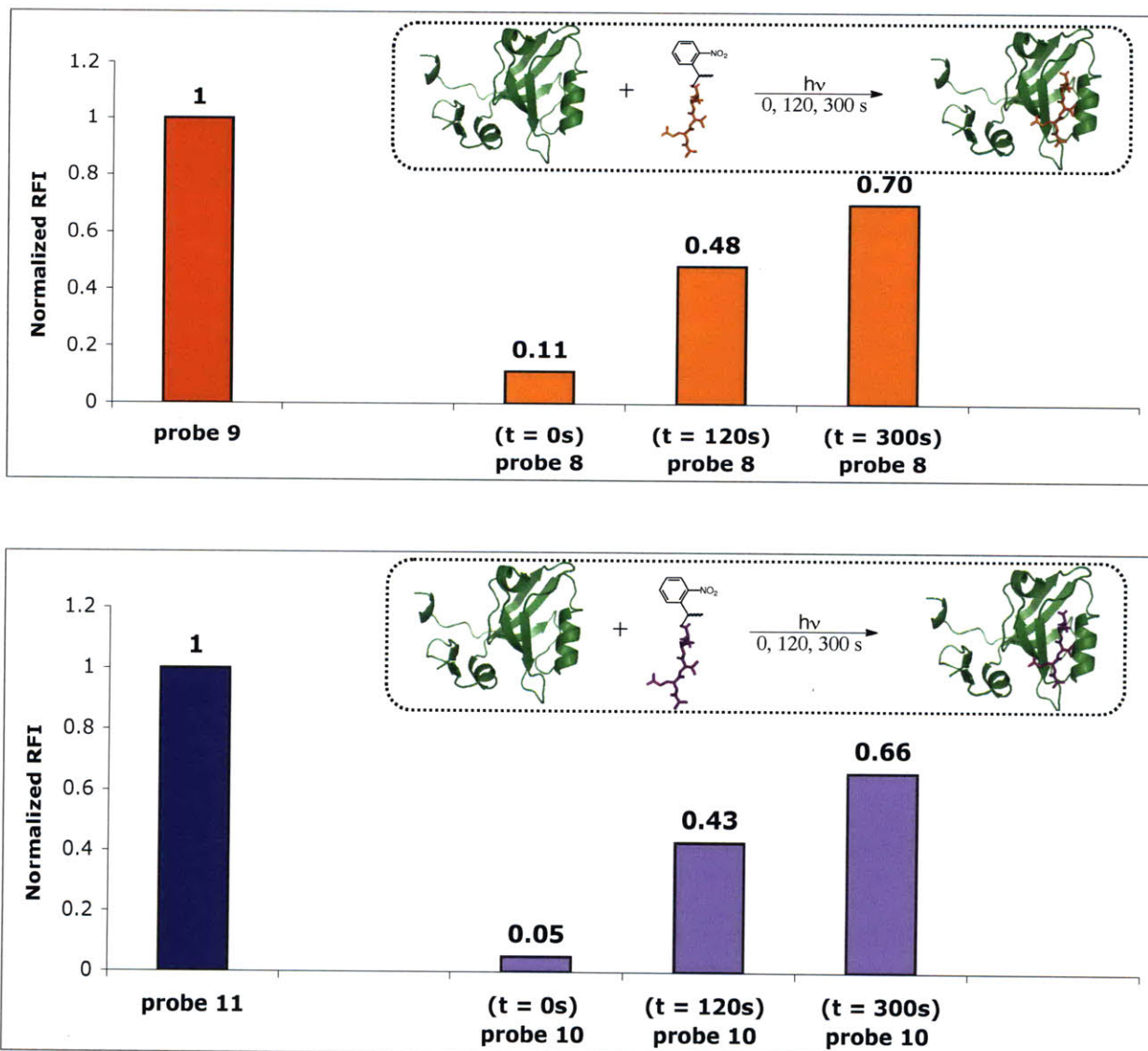
### 3-3. Validation of caging strategy in vitro

In the first phase of studies, the caging and subsequent uncaging of the C-terminal acid of probes **8** and **10** was validated qualitatively by relative fluorescence increase comparisons. The environment-sensitive fluorophore in these ligands yields a significant fluorescence increase upon binding to partner PDZ domains, which allows for the use of fluorescence spectroscopy to visualize binding after uncaging both qualitatively and quantitatively. The relative fluorescence increases of caged and uncaged peptides in the presence and absence of cognate PDZ domains were evaluated, and compared with fluorescence increases of non-caged ligands, as shown in Figure 3.2. Introduction of the NPE cage clearly resulted in significant loss of fluorescence increase upon addition of the cognate domain, yielding fluorescence increases lower than 11% of that of the non-caged peptides (Figure 3.2, t = 0 s). Caged peptide **10** displays a mere 5% of the fluorescence signal of its parent non-caged peptide. This observation also reflects a lower binding affinity of the caged ligands, as previously a correlation has been shown between fluorescence increase and affinity within a given series of ligands for the same PDZ domain. Therefore, the caged fluorescent ligands do not bind appreciably to their partner PDZ domains. Additionally, the fluorescence of each caged ligand alone was compared with the fluorescence emission of uncaged ligand alone irradiated for 300 seconds on a transilluminator, the longest duration of irradiation utilized in

the *in vitro* studies. The emission from both species is the same, demonstrating that the process of uncaging does not lead to an artificial increase in the fluorescence signal of the environment-sensitive fluorophore.

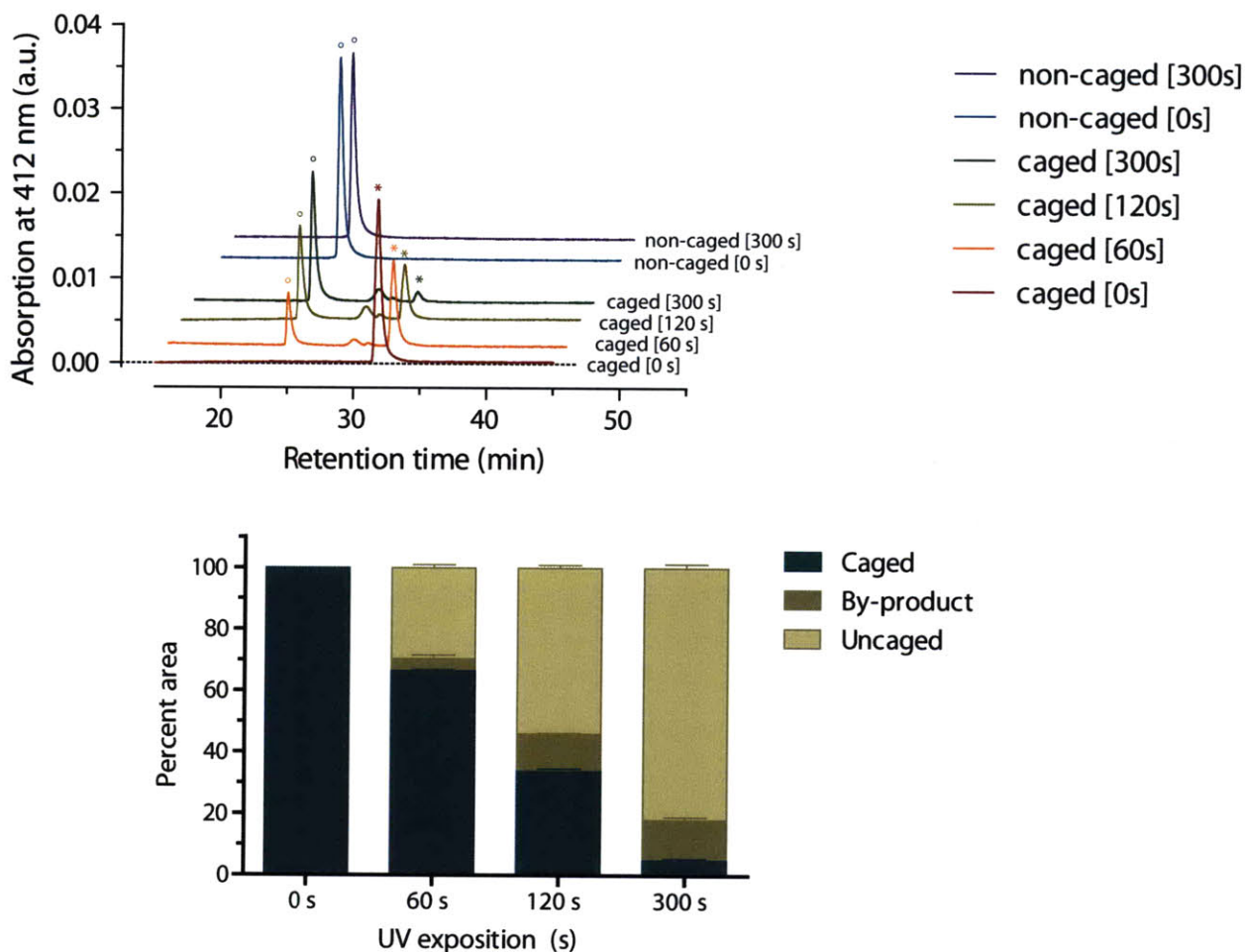
Exposure of the caged ligands to various durations of UV-irradiation (Figure 3.2,  $t = 120$  s and  $300$  s) in the presence of the scavenger 1,4-dithiothreitol (DTT) and cognate PDZ domain resulted in stepwise recovery of the fluorescence signal. Due to the release of nitroso acetophenone upon uncaging of the NPE group, the irradiation *in vitro* is carried out in the presence of DTT, which reacts irreversibly with the by-product and prevents unwanted side reactions. Application *in vivo* does not require supplementation with DTT, since the scavenger glutathione is present in ample amounts in cells. The NPE caging group is released with 350 nm light; previous *in vivo* studies have shown the efficient release of an NPE caging group from a phosphorylated amino acid after only 100 ms irradiation with an Argon ion laser (333.6-363.8 nm range).<sup>7</sup> The transilluminator is much less powerful than a laser, and therefore longer irradiation times are necessary to mimic laser-mediated release in these *in vitro* studies. This translates to the partial release of the caging group by photolysis, followed by binding of ligand to PDZ domain. At  $t = 300$  s, as seen in Figure 3.2, probe **8** yields a fluorescence increase of 70% of the maximum possible 89-fold increase in the presence of PSD-95 PDZ3. Uncaging for  $t = 300$ s of the GluR1-5b ligand yields a fluorescence increase of 66% of the maximum 265-fold (see Chapter 2 for relative fluorescence increases).





**Figure 3.2.** Qualitative validation of caged PDZ domain ligands. Probes in the presence of (top) PSD-95 PDZ3 and (bottom) Shank3PDZ. PDB entry 1BE9 used for sample PDZ domain + ligand images.

Initial investigations demonstrated that recovery of viable uncaged ligand plateaus at about  $t = 300$  s uncaging, hence this was used as the maximum uncaging time (data not shown). The application of this technology should be transferable to studies *in cellulo*, as the corresponding time of uncaging by laser is a reasonable duration of irradiation that should not damage cells. Figure 3.3 shows the HPLC timecourse of uncaging; caged peptide, uncaged peptide, and uncharacterized by-product are observed.



**Figure 3.3.** HPLC timecourse of Stargazin probe uncaging. Top: overlay of HPLC traces from different timepoints; legend indicates length in seconds of irradiation for each timepoint. Bottom: bar graph representation of caged peptide, uncaged peptide, and by-product observed at various timepoints of irradiation.

As shown in Figure 3.3, about 18% of the final product observed after irradiation for 300 s is composed of caged peptide and an unidentified by-product. Most likely, the by-product is the result of one or more modes of fluorophore degradation under the irradiation conditions. However, overall a significant recovery of viable ligand and its subsequent interaction with partner PDZ domain is observed upon uncaging, validating this caging approach for PDZ domain ligands.

To more quantitatively validate this method, the binding affinities of the caged peptides were evaluated by fluorescence titrations. Caged probe **8** binds PSD-95 with a weak affinity ( $K_D$  approximation of 34.9  $\mu\text{M}$ ), as compared with the  $K_D$  of 0.57  $\mu\text{M}$  for the parent non-caged ligand, a 60-fold difference. The binding constant between a non-fluorescent

Stargazin peptide and PSD-95 is 7-15  $\mu\text{M}$ ,<sup>8</sup> hence the caged compound would not be able to out-compete the natural binding event *in vivo*. Caged probe **10** is also a weak binder for Shank3 PDZ relative to the dissociation constant of the parent uncaged peptide. Titration studies yield a  $K_D$  of greater than 11.4  $\mu\text{M}$ . This lower bound estimate for the dissociation constant is 67-fold higher than the  $K_D$  of the parent non-caged ligand (0.17  $\mu\text{M}$ ). The large differences in caged versus non-caged probe affinities clearly demonstrate the efficiency of masking the C-terminal carboxylate residue to prevent interaction of PDZ domain ligands with cognate domains.

Overall, we have shown with two different PDZ domain ligand / PDZ domain protein pairs that NPE-caging of the C-terminal acid effectively decreases binding affinity, and that binding can be controlled by releasing the caging group with UV light. A significant recovery of viable uncaged product is observed upon uncaging. Future studies will be focused on optimizing this strategy for cellular imaging applications. In particular, various caging groups will be evaluated. Many different small-molecule caging groups have been reported, each with varying photophysical properties. One potential avenue to explore is the incorporation of a cage that is susceptible to two-photon uncaging, which is often advantageous for *in vivo* studies, since two photons of long wavelength infrared light are absorbed sequentially, rather than one photon of ultraviolet light. Two-photon uncaging offers excellent spatial control in photoactivation,<sup>9</sup> and has been particularly useful in neuronal studies.<sup>10, 11</sup>

## Conclusions

We have developed a synthetic strategy for caging the carboxyl group of the C-terminal amino acid of peptides. This caging application has not yet been reported in the literature. Since most PDZ domain ligands rely upon their C-terminal carboxyl moiety for binding to partner domains, they are perfect candidates for this caging approach. We have shown that the introduction of the NPE cage at the C-terminus of fluorogenic Stargazin and GluR1 probes successfully inhibits binding of these ligands to PSD-95 PDZ3 and Shank3 PDZ, respectively. Upon uncaging with UV light, we recover significant amounts of the uncaged ligands, whose fluorophores report binding to cognate PDZ domains. This strategy is general for PDZ domains, and should be applicable to all domains of this family.



This new strategy allows for the spatial and temporal control of binding events mediated by the peptide C-terminal carboxyl group, and could be extended to the full protein ligands. Additionally, we are particularly interested in the application of caged PDZ domain ligands in neuronal studies. Specifically, these tools could be used to perturb existing PDZ domain – ligand interactions at the synapse, allowing for observation of the downstream effects of this competitive interaction. The environment-sensitive fluorophore in these probes not only allows for validation of the caging strategy, but also increases the affinity of the ligand for the cognate PDZ domain, as previously reported. Hence, these caged fluorescent probes are appropriate for competition studies in neurons. Specifically, by caging the C-terminus the binding event is controlled, and by rendering the synthetic ligand a better binder for the PDZ domain, the ability to compete out the natural protein-protein interactions is effectively enhanced. We envision that, with optimization of imaging equipment, the environment-sensitive fluorophore within these probes could also be used to directly detect these binding events upon uncaging. Together, these approaches would allow researchers to exploit the full potential of the C-terminal caged fluorescent probes.

## Experimental Methods

### *Peptide Synthesis (SPPS), Fluorophore Synthesis, and Fluorophore Insertion.*

Peptides were obtained either manually or with an automated synthesizer (Advanced ChemTech automated synthesizer model 396  $\Omega$ ) by using standard Fmoc-based solid phase peptide synthesis procedures. Typically manual synthesis was performed on a 0.02 to 0.04 mmol scale with Fmoc-Val/Leu/Ile-NovaSyn® TGT resin (0.2 mmol/g) or Fmoc-PAL-PEG-PS resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents (6 equivalents each) and DIPEA (12 equivalents) in DMF or NMP. Coupling steps were conducted with a 50 mM solution of amino acid (~1 mL per 100 mg of resin) for 1 hour. Fmoc group removal was performed with a 20% solution of 4-methyl-piperidine in DMF (vol/vol) for 3 x 5 minutes. After removal of the N-terminal Fmoc group, the resulting free amine was capped with an acetyl group by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

Peptides incorporating the 4-DMAP environment-sensitive fluorophore were synthesized using the on-resin derivatization approach previously reported.<sup>12</sup> An Alloc-protected diaminoacid (either Fmoc-Dap(Alloc)-OH, Fmoc-Dab(Alloc)-OH or Fmoc-Orn(Alloc)-OH from AnaSpec) was initially inserted during the SPPS steps. After capping of the N-terminal amino group of the peptide sequence, the Alloc-protecting group was removed. The resin was resuspended in dry dichloromethane (20 mM in peptide) with a stream of N<sub>2</sub> bubbling through the solution for 5 minutes. Pd(PPh<sub>3</sub>)<sub>4</sub> (0.8 eq.) and phenylsilane (25 eq.) were then added to the solution, which was maintained under N<sub>2</sub> bubbling for another 15 minutes. The resin was then washed with dichloromethane and DMF, and the degassing/deprotection cycle was repeated two additional times. The resin containing the free amine was reacted with a solution of the anhydride form of 4-DMAP (4-dimethylaminophthalenedicarboxylic anhydride, 2 equiv., 50 mM) and DIPEA (4 equiv.) in NMP. The mixture was allowed to stir overnight. The resin was then washed with DMF and dichloromethane and ring closure was performed by using a solution of HBTU/HOBt (6 equiv., 50 mM) in NMP or DMF with DIPEA (12 eq.) for 2 hours. The resin was then washed with DMF and dichloromethane. The coupling/washing cycle was repeated two additional times in order to achieve full ring closure.

Cleavage from the resin by 0.5% TFA in dichloromethane (2.5 hrs) was performed, in order that the side chain protecting groups would not be affected. The reaction mixture was dissolved in the minimal volume dichloromethane, followed by addition of a large volume of hexane. Evaporation of the solvents by rotary evaporator yielded a powder, which was transferred to two Eppendorf tubes with a minimal volume dichloromethane. 10 eq. Mn(IV) oxide and 1 eq. 2-nitroacetophenone hydrazone were transferred to separate Eppendorf tubes, and they were then placed on the shaker for 30 min. The tubes were centrifuged, and then the solutions were filtered (using about 200  $\mu$ L dichloromethane) into the tubes containing peptide. These reaction mixtures were covered with foil and placed on the shaker overnight. The reaction mixture was concentrated on the SpeedVac.

Final deprotections were carried out with a TFA/H<sub>2</sub>O/TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours to afford the NPE-caged probe. The peptides were then triturated and precipitated in cold ether before purification by reverse phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250x20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 350 nm. Peptides were stored lyophilized at -80 °C until use.

***Peptide Characterization and Quantification.*** Peptide identity was confirmed by ESI-MS. Purity was assessed by analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 min followed by 5-95% acetonitrile containing 0.1% TFA over 50 min in water containing 0.1% TFA at a flow rate of 1 mL/min). The final peptides were quantified using the molar extinction coefficient of 4-DMAP ( $\epsilon_{421\text{nm}} = 6480 \text{ M}^{-1}\text{cm}^{-1}$  in water).

***Cloning of Shank3 PDZ domain constructs.*** The third PDZ domain of PSD-95, PSD95-3 (residues 302 to 402, from UniProtKB/Swiss-Prot entry P31016), was PCR-amplified using the primers 5'-CGG GAT CCG AGA ATT TGT ATT TTC AGG GCC TGG GGG AGG AAG ACA TTC CCC GGG and 5'-CCG CTC GAG TTA CTT ATC GTC ATC GTC TTT GTA GTC GGC CTC GAA TCG ACT ATA CTC TTC T, as described in Chapter 2 of this document. The PCR-amplified inserts were digested with BamHI/XhoI and ligated into a BamHI/XhoI-digested pGEX-4T-2 vector (GE Healthcare) for GST-TEV-PSD95PDZ3-FLAG.

The PDZ domain of Shank3 (residues 635 to 749, from UniProtKB/Swiss-Prot entry Q9JLU4), was PCR-amplified using the primers 5' -CGG GAT CCG AGA ACC TGT ACT TCC AGG GCT CAC ACA GTG ATT ATG TCA TTG ATG and 5' -CCG GAA TTC TTA CTT ATC GTC ATC GTC CTT GTA GTC TGC GCC GAG CAC TAT CCT CCT CTG G, as described in Chapter 2 of this document. The PCR-amplified inserts were digested with BamHI/EcoRI and ligated into a BamHI/EcoRI-digested pGEX-4T-2 vector (GE Healthcare) for GST-TEV-Shank3PDZ-FLAG.

***Expression and Purification of PDZ Domain Constructs.*** *E. Coli* BL21 codon plus™ (DE3)-RIL or RP cells were transformed with each of the GST fusion protein expression plasmids. Cells were amplified in 0.5 L of an auto-inducing medium<sup>13</sup> (ZYM-5052, see reference for detailed composition) first at 37 °C for 4 hours followed by 12 hours at 16 °C. Cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (100 mM EDTA, 10% glycerol, 1% TritonX-100, 1 mg/mL lysozyme in PBS, pH 7.4) containing protease inhibitor cocktail III (Calbiochem). Cells were lysed by ultrasonication (Branson Sonifier 450 at 50% power with a 40% duty cycle for 4 min at 4 °C). After addition of 1 mL of a 1 mM solution of DTT in water, the lysate was cleared by centrifugation (15,334 g, 40 minutes, 4 °C). The GST-fusion proteins were purified with 5 mL of Glutathione Sepharose™ 4 Fast Flow (GE Healthcare) by batch-binding. The bound proteins were washed with 20 volumes of PBS (pH 7.4) and eluted with 2 volumes of a 10 mM Glutathione in 50 mM Tris (pH 8) solution. Fractions were analyzed by 12% SDS-PAGE followed by Coomassie staining and Western blotting (anti-FLAG). Fractions containing the protein were pooled and dialyzed against PBS (pH 7.4). Protein concentrations were measured using either the BCA assay

(Pierce) with BSA as the reference standard or by determining the absorption at 280 nm in 6 M guanidinium chloride. Purified PDZ domains were aliquoted, flash-frozen and stored at -80 °C in PBS (pH 7.4) until use.

**Fluorescence Studies.** Fluorescence spectra were recorded on a Fluoromax 3 instrument (Horiba Jobin Yvon) in 1 cm path length quartz cells (100 mL nominal volume from Starna Cells). All measurements were performed at a constant temperature of 25 °C. Slit widths were 3 nm for excitation and 6 nm for emission. The 4-DMAP fluorophore was excited at 421 nm and the spectra were recorded between 432 nm and 730 nm (0.5 nm increments and 0.1 s integration time).

**Uncaging and Relative Fluorescence Increases.** Caged fluorescent ligands were uncaged in a glass apparatus with a pathlength of 1 mm on a UV transilluminator for 0, 120, and 300 seconds, in the presence of 0 or 20  $\mu\text{M}$  protein and 5 mM 1,4-dithiothreitol. The emission spectra of the resultant solutions were gathered on a fluorimeter, and the maximum emission intensities were compared with the emission intensity of the caged compound alone at that  $\lambda_{\text{max}}$ , yielding the fluorescence increase upon uncaging and binding. These values were compared with the fluorescence increase of non-caged fluorescent ligand with the partner PDZ domain.

**Fluorescence Titrations.** For each titration, the peptide concentration was kept constant (2  $\mu\text{M}$ ) and the protein concentration was varied from values lower than or close to the anticipated  $K_D$  to saturation of the fluorescence signal increase. A 150 or 130  $\mu\text{L}$  solution in PBS (pH 7.4) was prepared for each protein concentration. Dissociation constants were evaluated using GraphPad Prism 5 software after averaging at least three independent titrations and taking into account wavelengths from 432 to 730 nm.

## Acknowledgements

I would like to acknowledge Dr. Matthieu Sainlos, with whom I have collaborated on this work. He developed the initial synthesis for caged ligands, and we have worked together on the validation of this method (he performed the Stargazin / PSD-95 studies and I performed the GluR1 / Shank3 studies), which will be published in the near future.

## References

1. Ellis-Davies, G. C. "Caged compounds: photorelease technology for control of cellular chemistry and physiology." *Nat. Methods* **2007**, *4*, 619-628.
2. Lee, H. M.; Larson, D. R.; Lawrence, D. S. "Illuminating the chemistry of life: design, synthesis, and applications of "caged" and related photoresponsive compounds." *ACS Chem. Biol.* **2009**, *4*, 409-427.
3. Hung, A. Y.; Sheng, M. "PDZ domains: structural modules for protein complex assembly." *J. Biol. Chem.* **2002**, *277*, 5699-5702.
4. Rothman, D. M.; Vazquez, M. E.; Vogel, E. M.; Imperiali, B. "General method for the synthesis of caged phosphopeptides: tools for the exploration of signal transduction pathways." *Org. Lett.* **2002**, *4*, 2865-2868.
5. Nguyen, A.; Rothman, D. M.; Stehn, J.; Imperiali, B.; Yaffe, M. B. "Caged phosphopeptides reveal a temporal role for 14-3-3 in G1 arrest and S-phase checkpoint function." *Nat. Biotechnol.* **2004**, *22*, 993-1000.
6. Zhang, M.; Wang, W. "Organization of signaling complexes by PDZ-domain scaffold proteins." *Accounts Chem. Res.* **2003**, *36*, 530-538.
7. Humphrey, D.; Rajfur, Z.; Vazquez, M. E.; Scheswohl, D.; Schaller, M. D.; Jacobson, K.; Imperiali, B. "In situ photoactivation of a caged phosphotyrosine peptide derived from focal adhesion kinase temporarily halts lamellar extension of single migrating tumor cells." *J. Biol. Chem.* **2005**, *280*, 22091-22101.
8. Dakoqi, S.; Tomita, S.; Karimzadegan, S.; Nicoll, R. A.; Brecht, D. S. "Interaction of transmembrane AMPA receptor regulatory proteins with multiple membrane associated guanylate kinases." *Neuropharmacology* **2003**, *45*, 849-856.
9. Zipfel, W. R.; Williams, R. M.; Webb, W. W. "Nonlinear magic: multiphoton microscopy in the biosciences." *Nat. Biotechnol.* **2003**, *21*, 1369-1377.
10. Pettit, D. L.; Wang, S. S.; Gee, K. R.; Augustine, G. J. "Chemical two-photon uncaging: a novel approach to mapping glutamate receptors." *Neuron* **1997**, *19*, 465-471.
11. Kantevari, S.; Matsuzaki, M.; Kanemoto, Y.; Kasai, H.; Ellis-Davies, G. C. "Two-color, two-photon uncaging of glutamate and GABA." *Nat. Methods* **2010**, *7*, 123-125.
12. Sainlos, M.; Imperiali, B. "Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization." *Nat. Protoc.* **2007**, *2*, 3201-3209.
13. Studier, F. W. "Protein production by auto-induction in high-density shaking cultures." *Prot. Expr. Purif.* **2005**, *41*, 207-234.

## Chapter 4

# Application of Chemical Probes to the Study of Shank3 PDZ Domain Dimerization

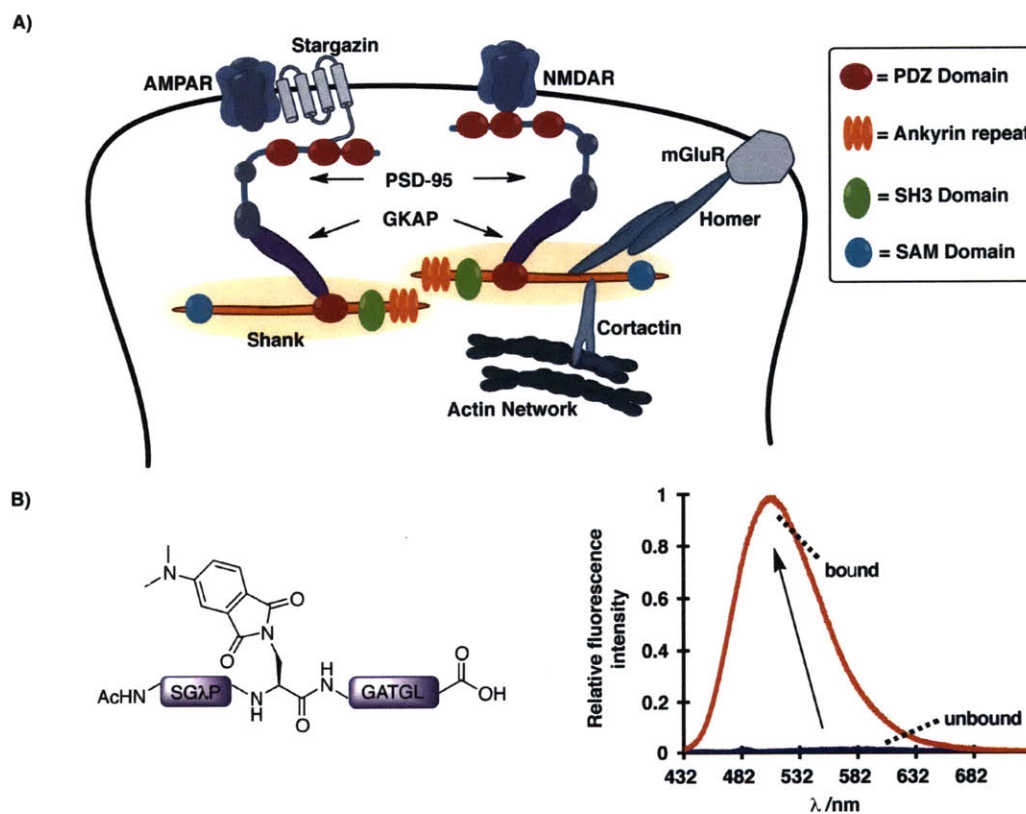
Portions of this chapter have been published in *ChemBioChem*<sup>1</sup> as noted in the text. Copyright © 2010, Wiley-VCH.

### Introduction<sup>1</sup>

The multi-domain protein Shank is a master-scaffold protein found in the postsynaptic density (PSD) of neurons.<sup>2</sup> While multiple different interactions between Shank and other proteins have been observed and delineated, specific mechanisms that modulate these interactions in the PSD remain to be elucidated. The PSD comprises a network of scaffolding, cytoskeletal and signaling proteins. These proteins form complexes that control the receptors of the postsynaptic membrane, aid in signal transduction, and make up the structural framework of the PSD, all of which are critical to learning and memory.<sup>3, 4</sup> Macromolecular assemblies of proteins in the PSD are usually assembled around scaffolding proteins, one of which is the Shank family. Shank proteins contain multiple putative interaction domains, including an ankyrin repeat domain, an SH3 domain, a SAM domain, a proline-rich domain, and a PDZ (PSD-95/DLG/ZO-1) domain.<sup>5</sup>

PDZ domains are essential protein modules involved in the assembly and localization of protein-protein complexes, and they function by binding C-terminal sequences of partner proteins. Critical to Shank are the binding partners Homer and GKAP, which bind to the proline-rich domain and the PDZ domain of Shank, respectively. Significantly, Homer, Shank and GKAP form an oligomeric complex that is thought to act as a bridge between the structural proteins of the postsynaptic cytoskeleton and the receptors at the postsynaptic membrane as shown in Figure 4.1A.<sup>6</sup> Many PDZ domain–ligand interactions are promiscuous and display modest low micromolar affinities, and it is therefore likely that other mechanisms exert control over these interactions in the PSD. Although high throughput screens have successfully identified multiple PDZ domain–ligand interactions,<sup>7</sup> these studies have not specifically taken into account higher order interactions that might provide an additional level

of affinity modulation for PDZ domain-ligand interactions. Chemical and biophysical investigations designed to probe which modes of control are indeed relevant to specific PDZ domain-ligand interactions would aid in a more complete understanding of the control, regulation, and mechanisms of these proteins and their complex networks. Experimentally controlled characterizations of the Shank3 PDZ domain in solution suggest that the formation of higher-order Shank complexes is critical to the modulation of Shank ligand-binding affinity, aiding the Shank protein in fulfilling key structural and functional roles in the neuronal PSD.



**Figure 4.1.** Probing the Shank PDZ domain. A) Schematic of macromolecular complexes that rely on the Shank protein in the PSD. B) Fluorescent probe for Shank3 PDZ domain; upon binding to the Shank3 PDZ domain, an increase in the emission intensity of the solvatochromic fluorophore is observed. The previously-reported design and screening approach for the development of PDZ domain probes yielded this GluR1-based fluorescent peptide as the optimal probe for the Shank3 PDZ domain.<sup>8</sup>  $\lambda$  = norleucine; graph shows fluorescence emission spectra of unbound (purple) and Shank3-bound (orange) states. Conventionally, PDZ ligand residue numbering begins at 0 for the C-terminal residue and continues sequentially in the negative direction toward the N-terminus.



The native Shank protein (Figure 6.2) may oligomerize through a variety of interaction mechanisms. Shank and Homer form a polymeric network that provides structural and functional integrity to dendritic spines.<sup>2</sup> Shank1 can homodimerize through interactions between its component ankyrin repeat and SH3 domains,<sup>9</sup> and Shank can potentially multimerize in a tail-to-tail manner due to the homomeric and heteromeric interactions between SAM domains of different Shank molecules. Many of these oligomerized states of Shank are believed to be necessary for maintaining overall architecture and function of synaptic proteins.<sup>10</sup> Previously, the individual PDZ domain of Shank1 has been shown to exist as both a monomer in solution and a dimer in solution and when complexed with peptide ligand in crystallization studies.<sup>11, 12</sup> The interdomain contact in the dimer observed in the Shank1 crystal structure involves an extended  $\beta$ B/ $\beta$ C loop as well as exceptionally long N- and C- terminal extended  $\beta$ -sheets. These regions of extended secondary structure exhibit high sequence identity among the Shank PDZ domains, yet vary significantly in sequence and length in other PDZ domains, including the canonical PSD-95 PDZ3 monomer (see SI Figure S1).<sup>11</sup> These structural features may account for the non-canonical dimerization behavior of the Shank1 PDZ domain, and suggest that the PDZ dimer may be a relevant species across the entire Shank family.

In light of the ample evidence for the formation of higher-order complexes involving multiple copies of Shank, synergistic approaches were designed to investigate whether Shank3 PDZ domain dimerization modulates the affinity of this domain for partner ligands. An integral component of these studies was the application of fluorescence-based probes (see Chapter 2) to quantify the functional effect of Shank3 PDZ domain dimerization on ligand binding capabilities. The design and development of these probes is described in Chapter 2 of this document.<sup>8</sup> Utilizing controlled protein dimerization in solution, in conjunction with the fluorescent probe that was developed previously for the Shank3 PDZ domain (Figure 4.1B), studies were designed to probe whether or not protein oligomerization state is a critical factor in determining the affinity of this domain for cognate ligands.

## Results and Discussion

### ***4-1. Molecular biology and expression of all Shank3 PDZ domain constructs<sup>1</sup>***

Studies to address the functional consequences of Shank3 PDZ domain dimerization on ligand binding will be presented in this chapter. In order to investigate the functional significance of Shank3 PDZ domain dimerization, the isolated PDZ domain with flanking residues was expressed as a fusion construct with multiple different fusion partner domains (see Figure 4.2). These constructs allow for the controlled study of the PDZ domain itself, and overcome the obstacles of purifying and handling the full 1741-residue Shank3 protein *in vitro*.

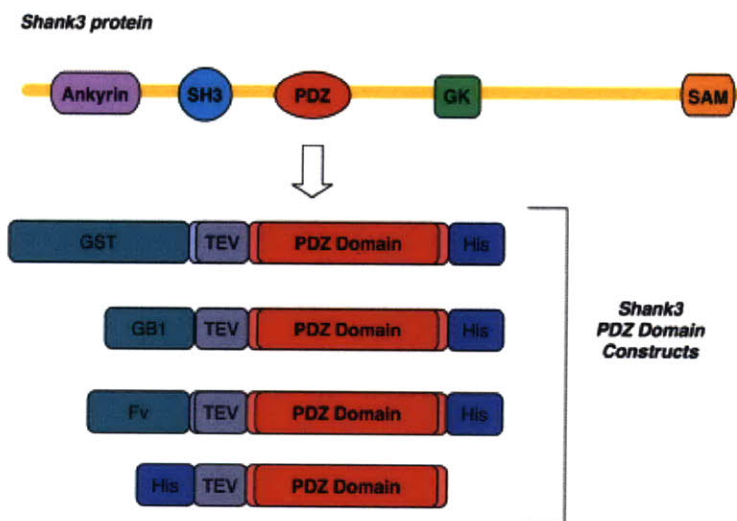
There are three different Shank family proteins, each of which shows a distinct pattern of expression in different tissues. However, they are similar in domain structure, and their PDZ domains are highly homologous (see Figure 4.2A), binding to the same C-terminal ligands. The regions highlighted in red in the Figure 4.2A alignment correspond to the  $\beta$ A and  $\beta$ F sheets of Shank, and the region in blue is the  $\beta$ B- $\beta$ C loop, all of which form the dimer interface in the Shank1 PDZ domain crystal structure. These regions display high sequence homology within the Shank family, but not with the canonical PSD-95 PDZ3.

A)

Q9WV48_Shank1_rat	SGSD <b>YI I KEKTVLLQ</b> KKDSEGFVLRGAKA <b>QTPIEEFTPTPAFPALQ</b> YLESVDEGGVAW
Q9QX74_Shank2_rat	AASD <b>CI I EDKTVVLQ</b> KKDNEGFVLRGAKAD <b>TPIEEFTPTPAFPALQ</b> YLESVDEGGVAW
Q9JLU4_Shank3_rat	SHSD <b>YV I DDKVAILQ</b> KRDHEGFVLRGAKA <b>ETPIEEFTPTPAFPALQ</b> YLESVDVEGVAW
P31016_DLG4_rat	LGEEDIPREPRRIVIHRSSTGLGFNIVGGEDGE-----GIFISFILAG----GPADL

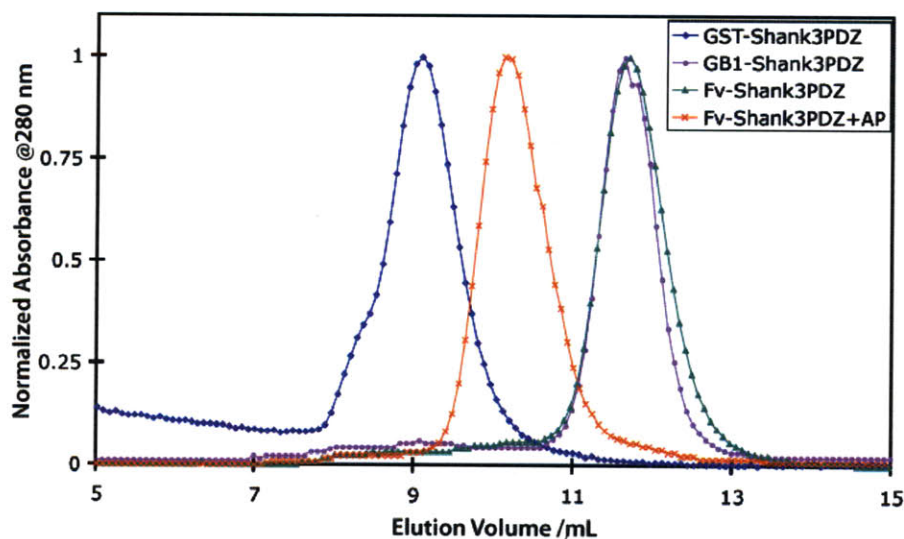
Q9WV48_Shank1_rat	RAGLRMGDFLIEVNGQNVVKGVHRQVVNMIRQGG <b>W TLMVKVVMV</b> TRHPDMDEAVH
Q9QX74_Shank2_rat	QAGLRTGDFLIEVNNENVVKGVHRQVVNMIRQGG <b>WHLVLKVVTV</b> TRNLDPDDTAR
Q9JLU4_Shank3_rat	KAGLRTGDFLIEVNGVNVVKGVHKQVVGLIRQGG <b>WRLVMKVVSV</b> TRKPEEDSARR
P31016_DLG4_rat	SGELRKGQDILSVNGVDLRNASHEQAAIALKNAGQTVTTIIAQYKPEYSRFEA--

B)



**Figure 4.2.** Shank3 PDZ domain constructs. A) Alignment of Shank family PDZ domains. The Shank3 PDZ domain residues here correspond to those used in the various protein constructs used. B) Shank3 and schematic of different Shank3PDZ protein constructs. Flanking residues from the native Shank3 sequence were incorporated around the Shank3 PDZ domain, to ensure proper folding of the domain. In every case except the His<sub>6</sub>-TEV-Shank3 construct, which was treated with TEV protease to yield the PDZ domain alone, the Shank3 PDZ domain and flanking residues are surrounded by the TEV sequence on the N-terminal end and the His<sub>6</sub> tag on the C-terminus. Before the TEV sequence in the GST and Fv constructs there are a thrombin tag and a three glycine linker, respectively. Original experiments were done using the GST-Shank3PDZ-FLAG construct. Fluorescence titrations and control experiments were repeated with a GST-Shank3PDZ-His<sub>6</sub> construct, and all were in agreement with the previously obtained results. Hence, the identity of the C-terminal tag does not affect the behavior of the GST construct. Note that the ITC results presented in the main text were obtained from the GST-Shank3PDZ-FLAG construct, and all other results presented herein are from the GST-Shank3PDZ-His<sub>6</sub> construct.

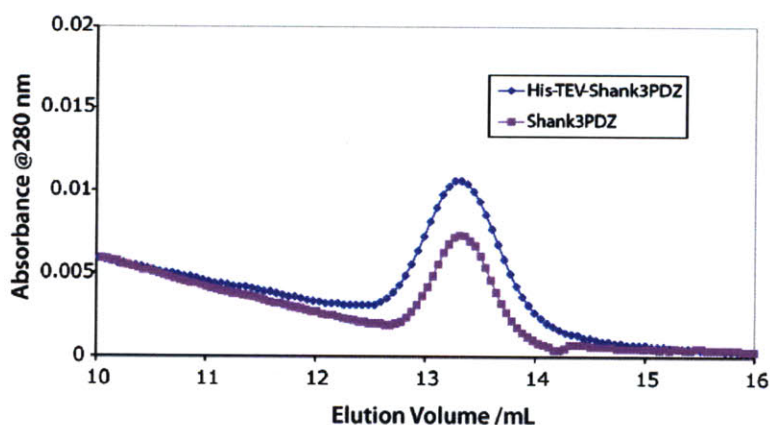
All of the constructs extend well beyond the residues necessary for PDZ domain dimerization, yet the Shank3 PDZ domain alone is not sufficient for dimerization in solution, as determined by size exclusion chromatography (Figures 4.3 and 4.4). Glutathione S-transferase (GST) and IgG-binding domain B1 of *Streptococcal* protein G (GB1) fusions afforded obligate dimeric and monomeric Shank3 PDZ domain constructs, respectively. A complementary approach utilizing the Fv (optimized FKBP) domain allowed for external regulation through a chemical inducer of dimerization (CID, AP20187);<sup>13, 14</sup> Without AP20187 (abbreviated AP) the Fv construct is a monomer, but in the presence of AP the construct is a dimer.



Monomer or Dimer?	Protein Construct	Expected MW (Da)	MW from SEC analysis (Da)
M	Fv-Shank3PDZ	27,444	26,786
D	Fv-Shank3PDZ / AP	54,888	48,532
M	GB1-Shank3PDZ	20,707	24,968
D	GST-Shank3PDZ	81,304	73,495

**Figure 4.3.** Size exclusion chromatography results for protein constructs. These results establish that GB1-Shank3PDZ and Fv-Shank3PDZ are monomeric, while GST-Shank3PDZ and Fv-Shank3PDZ+AP are dimeric. Molecular weights were calculated based on elution volume and comparison with size standards Albumin (66,000 Da), Ovalbumin (43,000 Da), and Ribonuclease A (13,700 Da). Y-axis normalized to the largest peak.





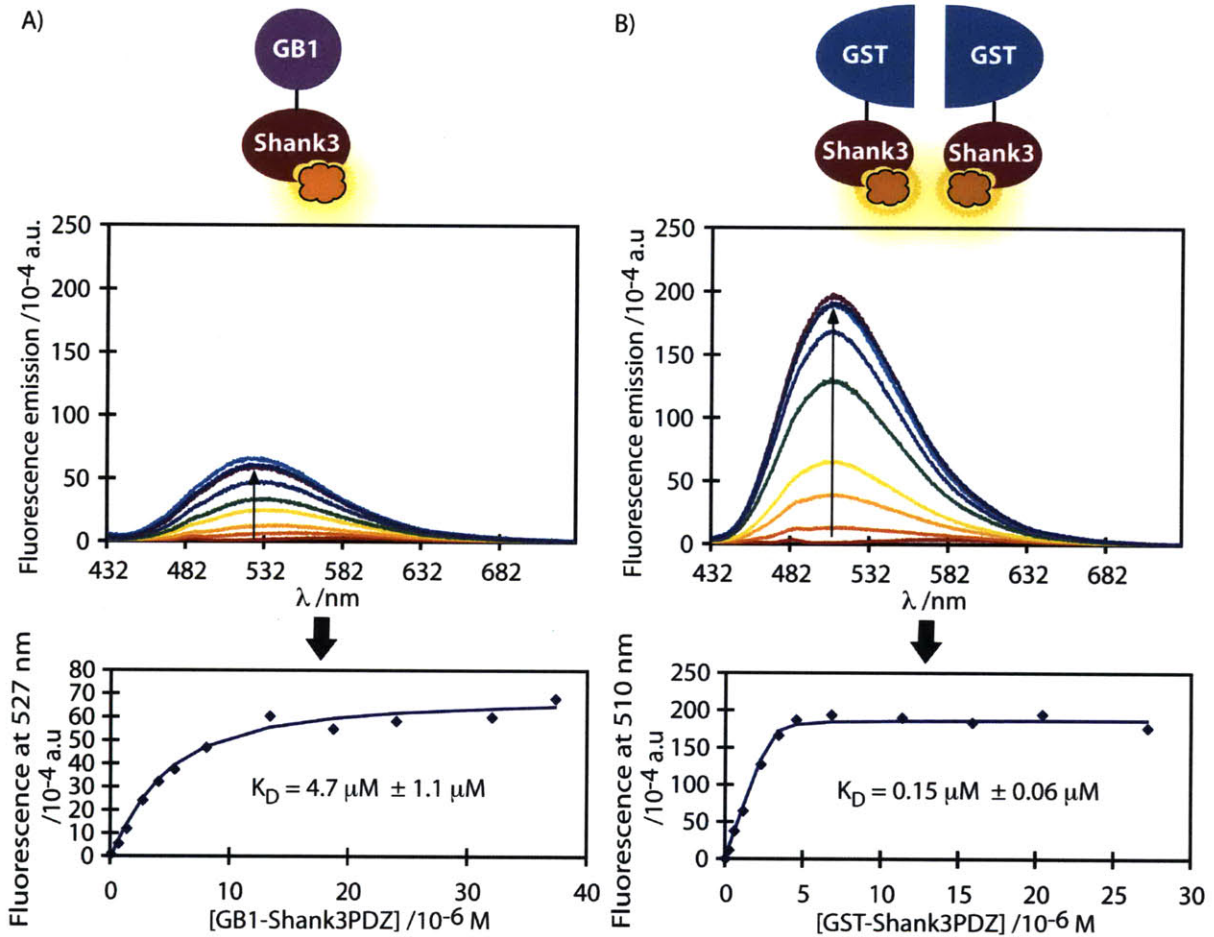
Monomer or Dimer?	Protein Construct	Expected MW (Da)	MW from SEC analysis (Da)
M	His <sub>6</sub> -TEV-Shank3PDZ	14,634	15,336
M	Shank3PDZ	12,685	15,137

**Figure 4.4.** Size exclusion chromatography results for Shank3 PDZ domain alone, with and without His<sub>6</sub> tag. These results confirm that the Shank3 PDZ domain with only a His<sub>6</sub> tag and without any tags is monomeric (i.e. the tags in the constructs used in this study have not perturbed any natural dimerization of the domain in solution). Molecular weights were calculated based on elution volume and comparison with Gel Filtration Standard (BioRad).

#### ***4-2. Probing the binding properties of the obligate Shank3 PDZ domain monomer and dimer<sup>1</sup>***

In the first dimerization approach, the ligand affinities of two fusion constructs of the Shank3 PDZ domain, GST-Shank3PDZ-FLAG and GB1-Shank3PDZ-His<sub>6</sub>, were compared. The GB1 domain is a small fusion partner that is known to enhance the expression and solubility of the corresponding protein constructs.<sup>15</sup> While the favorable properties of the GB1 fusion tag indicate monodispersity of GB1 fusion proteins, the GST fusion tag itself has been shown to homodimerize with a crystallographic 2-fold symmetry.<sup>16</sup> Interestingly, these two constructs showed very different ligand-binding affinities. Titrations with the fluorescent probe demonstrated a 30-fold difference in affinity between the two constructs, a significant difference in ligand-binding behavior. Size exclusion chromatography (SEC) confirmed that GB1-Shank3PDZ behaves exclusively as a monomer, while GST-Shank3PDZ is exclusively a

dimer (Figure 4.3). The probe bound to (monomeric) GB1-Shank3PDZ with a dissociation constant ( $K_D$ ) of  $4.7 \mu\text{M} \pm 1.1 \mu\text{M}$  and to (dimeric) GST-Shank3PDZ with a  $K_D$  of  $0.15 \mu\text{M} \pm 0.06 \mu\text{M}$  (Figure 4.5). This 30-fold difference in  $K_D$  would be significant in the context of protein / ligand-binding *in vivo*.



**Figure 4.5.** Ligand binding to obligate monomer and dimer of Shank3 PDZ domain. A) Titration of fluorescence-based probe with GB1-Shank3PDZ (monomer) and corresponding data fit. B) Titration of fluorescence-based probe with GST-Shank3PDZ (dimer) and corresponding data fit. Representative titrations,  $K_D$  and standard deviation from at least three trials. Peptide concentration  $3 \mu\text{M}$ , protein concentrations from 0 to  $> 25 \mu\text{M}$ .

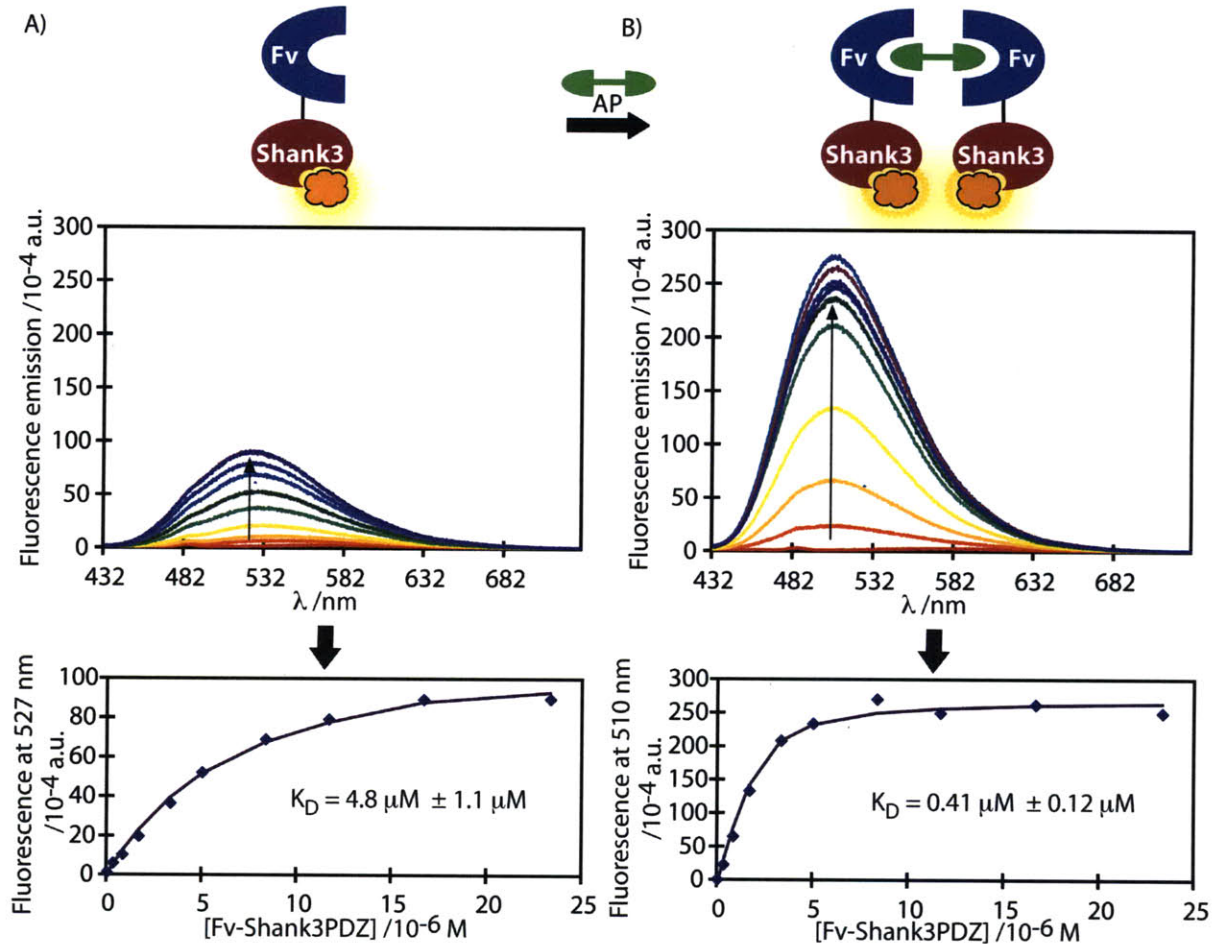
Several control experiments confirm that the higher affinity for the GST construct does not result from nonspecific interactions of the fluorophore. The fluorescent probe displays no appreciable fluorescence increase (less than 1.5-fold) or binding affinity when

titrated with GST domain alone. Non-optimal fluorescent probes, with the fluorophore at the (-)6 and (-)4 positions in the peptide, as opposed to the optimal (-)5 position, yield less than 4-fold increase in fluorescence with saturating GST-Shank3PDZ protein construct (see Figure 4.1 caption for numbering convention). This is less than 1.5% of the fluorescence increase observed with the optimized probe. These results establish that the greater fluorescence signal of the probe with the GST protein construct does not result from nonspecific interactions of the 4-DMAP fluorophore with the protein construct.

Additionally, ITC studies of the probe with GB1-Shank3PDZ and GST-Shank3PDZ constructs yield dissociation constants of  $2.9 \mu\text{M} \pm 1.3 \mu\text{M}$  and  $0.2 \mu\text{M} \pm 0.1 \mu\text{M}$ , respectively. These values agree well with those obtained by fluorescence titration, and demonstrate that the fluorescence titration method is still robust in this application. Overall, the probe does in fact bind with a significantly tighter affinity to dimeric GST-Shank3PDZ than to monomeric GB1-Shank3PDZ.

#### ***4-3. Probing Shank3 PDZ domain monomer and dimer via external regulator<sup>1</sup>***

A complementary chemical inducer of dimerization approach was taken, in order to isolate and study the effects of domain dimerization via an external small molecule regulator. The Fv (FKBP Phe36Val) domain affords this level of control over protein dimerization, through the chemical inducer of dimerization (CID) AP20187.<sup>14</sup> A Shank3 PDZ domain construct incorporating the Fv domain was engineered for this study and is shown in Figure 4.6. It should be noted that the design of this construct required semi-empirical optimization. Specifically, the Fv construct was augmented with a linker region, after preliminary titrations and a comparative analysis of the Fv dimer and GST dimer crystal structures indicated that additional flexibility might be necessary to fully allow for PDZ domain dimerization. This construct was expressed, purified, and tested in both the monomeric and dimeric forms.



**Figure 4.6.** Effects of Shank3 PDZ domain dimerization on ligand binding. A) Titration of fluorescence-based probe with Fv-Shank3PDZ (monomer) and corresponding data fit. B) Titration of probe with Fv-Shank3PDZ/AP20187 (dimer) and corresponding data fit. Representative titrations,  $K_D$  and standard deviation from at least three trials. Peptide concentration 2  $\mu\text{M}$ , protein concentrations from 0 to > 25  $\mu\text{M}$ .

A comparison of the resulting binding titrations highlights the fact that the maximum fluorescence increase of the fluorogenic probe is much larger for the dimeric Fv-Shank3PDZ than for the corresponding monomer (see Figure 4.6), a trend that was evident with the monomeric GB1 and dimeric GST fusion constructs. Since the fluorescence increase of the probe is correlated with the polarity of the surrounding environment, these results suggest that the Shank3 PDZ domain undergoes a structural reorganization upon dimerization, creating a more hydrophobic, higher-affinity binding pocket for the ligand. The fluorescent probe bound with an order-of-magnitude tighter affinity to the Fv-Shank3PDZ dimer ( $K_D$  of 0.41



$\mu\text{M} \pm 0.12 \mu\text{M}$ ) than to the corresponding monomer ( $K_D$  of  $4.8 \mu\text{M} \pm 1.1 \mu\text{M}$ ), mirroring the results from the GST-Shank3PDZ dimer and GB1-Shank3PDZ monomer, respectively (see entries 1 and 2 in Table 4.1).

Entry	State <sup>a</sup>	Protein Construct	Ligand <sup>b</sup>	$K_D$ <sup>c</sup>
1	M	GB1-Shank3PDZ	F1	$4.7 \pm 1.1$
	D	GST-Shank3PDZ	F1	$0.15 \pm 0.06$
2	M	Fv-Shank3PDZ	F1	$4.8 \pm 1.1$
	D	Fv-Shank3PDZ+AP	F1	$0.41 \pm 0.12$
3	M	Fv-Shank3PDZ	GKAP	$1.0 \pm 0.1$
	D	Fv-Shank3PDZ+AP	GKAP	$0.12 \pm 0.01$

**Table 4.1.** Dissociation constants from fluorescence titrations and competition experiments. <sup>a</sup>M = monomer, D = dimer. <sup>b</sup>F1 = fluorescent probe; GKAP = non-fluorescent GKAP ligand. <sup>c</sup> $K_D$  represented as mean  $\pm$  standard deviation, from at least three trials. Note that the numbers given in entry 3 are  $K_i$  values.

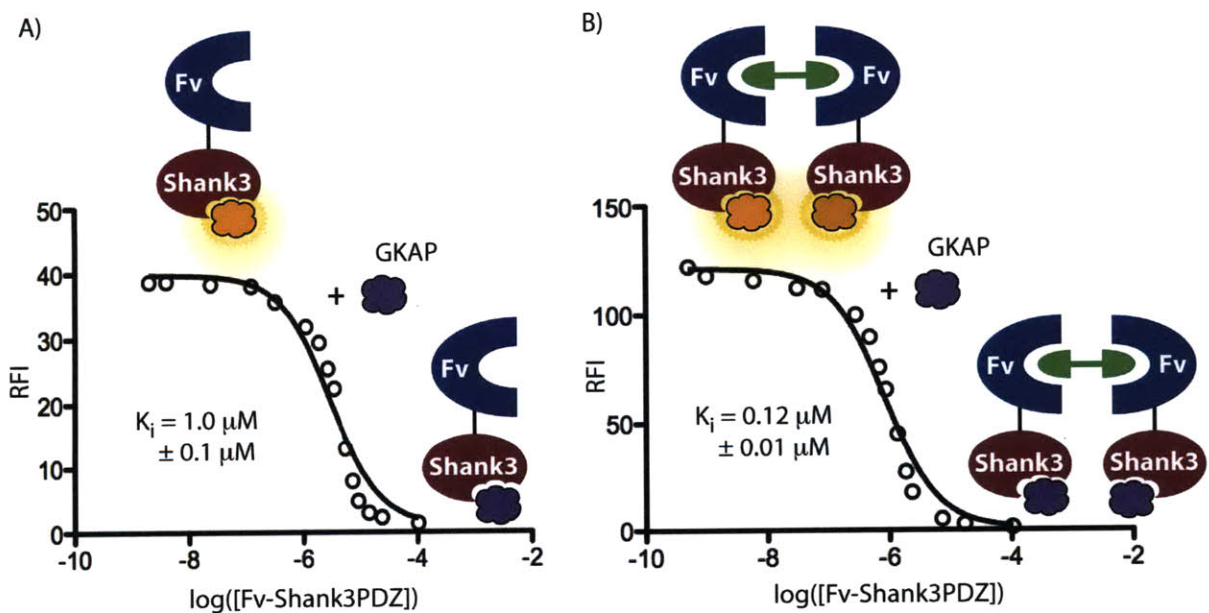
Control titrations of Fv domain alone and AP20187 alone with fluorescent probe show negligible fluorescence increase (less than 2-fold in each case). These results again confirm that the fluorescence increases observed in the Fv-Shank3PDZ monomer and dimer titrations arise from specific interactions between probe and the cognate PDZ domain. Size exclusion chromatography studies reveal that the Shank3 PDZ domain alone (with and without His<sub>6</sub> tag) is a monomer in solution, demonstrating that the N- and C-terminal tags do not interfere with any natural ability of the Shank3 PDZ domain to dimerize individually. Overall, these studies establish that *in vitro*, the designed fluorescent probe binds with an order of magnitude higher affinity to the dimerized Shank3 PDZ domain than to the monomeric form.

#### ***4-4. Effects of dimerization on a biologically-relevant ligand<sup>1</sup>***

The observed difference in affinity could be of potential significance in the PSD, if the magnitude of this effect also translates to natural (non-fluorescent) ligands. In this context, experiments were performed to investigate the effect of Shank3 PDZ domain dimerization on binding affinity of the GKAP ligand (C-terminal peptide of the GKAP protein). In order to determine these affinities via titration, competition experiments were performed in which non-fluorescent GKAP gradually displaced fluorescent probe bound to protein. As shown in

Figure 4.7, these data were fit to yield dissociation constants for the unlabeled GKAP peptide with both the monomeric and dimeric forms of Fv-Shank3PDZ.

The binding affinity of the unlabeled GKAP peptide is significantly affected by the oligomerization state of Shank3. GKAP binds with an 8-fold higher affinity to dimeric Shank3 PDZ domain ( $K_i = 0.12 \mu\text{M} \pm 0.01 \mu\text{M}$ ) than to the monomeric domain ( $K_i = 1.0 \mu\text{M} \pm 0.1 \mu\text{M}$ ). This difference in affinity follows the trend found in the previous studies utilizing the fluorescent probe (see Table 4.1), and demonstrates that the natural interaction between GKAP and the Shank3 PDZ domain can also be tuned by PDZ domain dimerization. In effect, dimerization of the Shank3PDZ domain pushes the equilibrium of its ligand-binding process significantly toward ligand-bound protein, a change that would have a large effect in the pool of PDZ domain-mediated interactions at the synapse. Taken together, the results presented here implicate PDZ domain dimerization as a potential modular control mechanism for ligand binding in the biological context.



**Figure 4.7.** Data fits of competitive titrations of unlabeled GKAP versus fluorescent probe / Shank3PDZ complexes. A) Fit of competition experiment in which 0 to 10 equivalents of the non-fluorescent GKAP was sequentially added to a  $10 \mu\text{M}$  solution of fluorescent probe and Fv-Shank3PDZ. B) Fit of competition experiment in which 0 to 25 equivalents of the non-fluorescent GKAP was sequentially added to a  $2.5 \mu\text{M}$  solution of fluorescent probe and Fv-Shank3PDZ/AP20187. Representative titrations,  $K_D$  and standard deviation from at least three trials. The imperfect fit is most likely caused by slightly different binding modes of the two different peptide ligands used in the competition experiments.

## Conclusions

Investigations into the macromolecular complexes that govern the structure and function of the neuronal PSD continue to point to the significance of multimeric interaction motifs, especially with respect to PDZ domains. The tandem PDZ domains of proteins such as PSD-95 have been shown to work together, relying upon this synergy to function properly in their interactions with protein ligands.<sup>17, 18</sup> Homodimers of glutamate receptor-interacting protein (GRIP1) PDZ6, zonula occludens (ZO) PDZ2, and Shank1 PDZ have been reported, all with different relative orientations of the monomer components.<sup>19, 20</sup> The dimerization of GRIP1-PDZ6 is proposed to enable the efficient clustering of synaptic GRIP-binding proteins,<sup>21</sup> and PDZ domain dimers of ZO proteins may provide a structural basis for the polymerization of claudin protein targets.<sup>20</sup>

In this study, the optimized fluorescent probe technology discussed in chapter 2 of this thesis has been used in conjunction with controlled protein dimerization to probe the effects of Shank3 PDZ domain dimerization on ligand binding, a relevant investigation in light of the functional effects of other PDZ domain homodimers that have been reported. Specifically, fluorescence titrations, isothermal titration calorimetry, and competition experiments all demonstrate that the Shank3 PDZ domain may function as a homodimer to achieve significant (8- to 30-fold) enhancement of ligand-binding affinities. Shank3 PDZ domain oligomerization state regulates the binding affinity of partner ligands *in vitro*, and the magnitude of the effect of dimerization suggests its potential as a regulatory mechanism for the formation of large synaptic multiprotein complexes in neurons. Since PDZ domains regulate signaling cascades that often rely upon the clustering of target proteins, it is perhaps not surprising that the dimerization of the Shank3 PDZ domain would exert significant control over the PDZ/ligand binding interaction.<sup>22</sup> Overall, these *in vitro* studies have led to a more comprehensive insight into the regulation of the Shank3 PDZ domain and have provided a foundation from which to begin designing experiments in cells. Future studies to probe the functional effects of Shank3 PDZ domain dimerization in neurons will clarify the degree to which this mechanism controls Shank3 PDZ domain ligand binding in the biological context.

## Experimental Methods

**Solid phase peptide synthesis (SPPS) and fluorophore insertion.** Peptides were obtained either by manual synthesis or with an automated synthesizer (Advanced ChemTech automated synthesizer model 396  $\Omega$ ) using standard Fmoc SPPS procedures. Manual syntheses were performed on a 0.02 mmol scale with TGT resin ( $\sim$ 0.2 mmol/g), using Fmoc-protected amino acid (6 equivalents), HBTU / HOBt coupling reagents (6 equivalents), and DIPEA (12 equivalents) in NMP. Couplings were done with a 50 mM amino acid solution for approximately 1 hour. Fmoc deprotection was achieved after each coupling step by treatment of the resin with a solution of 4-methyl-piperidine in DMF (20% vol/vol) for 3 x 5 minutes. The free amine of the N-terminal residue of the peptide was acetyl capped by treatment with a solution of acetic anhydride and pyridine (0.15 M each in DMF).

Peptides incorporating the 4-DMAP environment-sensitive fluorophore were synthesized using the on-resin derivatization approach previously reported.<sup>23</sup> An Alloc-protected diaminoacid (Fmoc-L-Dab(Alloc)-OH from AnaSpec) was incorporated into the peptide sequence during the initial coupling SPPS procedure described above. After N-terminal acetyl capping, the resin was resuspended in dry dichloromethane (20 mM in peptide) with a stream of N<sub>2</sub> bubbling through the solution for 5 minutes. Pd(PPh<sub>3</sub>)<sub>4</sub> (0.8 equivalents) and phenylsilane (25 equivalents) were then added to remove the Alloc group, and the bubbling of the solution was maintained for another 15 minutes. The resin was washed with dichloromethane, and the degassing/deprotection cycle was repeated two additional times. The free amine of the resulting peptide on resin was then reacted with a solution of the 4-DMAP anhydride (4-*N,N*-dimethylaminophthalic anhydride, 2 equivalents, 50 mM) and DIPEA (4 equivalents) in NMP overnight. The resin was washed with DMF and dichloromethane, and ring closure was performed by treatment with a solution of HBTU/HOBt (6 equivalents, 50 mM) in NMP with DIPEA (12 equivalents) for approximately 1 hour. The resin was then washed with DMF and dichloromethane, and the washing/coupling cycle was repeated two additional times in order to achieve complete ring closure.

Peptides were deprotected and cleaved from the resin with a TFA / H<sub>2</sub>O / TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours. TFA was evaporated until the peptides precipitated, and the residue was then triturated and precipitated in cold ether before purification by reverse

phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250X20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 350 nm. Peptides were characterized by analytical HPLC and MALDI-TOF mass spectroscopy. Peptides were lyophilized and stored at – 80 °C.

Fluorescent GluR1 peptide sequence: Ac-NRRSGλPβGATGL-COOH where λ = norleucine (replaced native methionine) and β = Dab(4-DMAP); expected mass = 1512.8, mass obtained = 1513.7.

Control Fluorescent GluR1 peptide sequence GluR1-4b: Ac-NRRSGλPLβATGL-COOH; expected mass = 1568.9, mass obtained = 1571.0.

Control Fluorescent GluR1 peptide sequence GluR1-6b: Ac-NRRSGλβLGATGL-COOH; expected mass = 1528.9, mass obtained = 1530.0.

GKAP peptide Ac-φNRRYIPEAQTRL-COOH where φ = *p*-nitrophenylalanine (which has been included to provide a handle for quantification purposes); expected mass = 1862.9, mass obtained = 1862.5.

***Cloning of Shank3 PDZ domain constructs.*** The PDZ domain of Shank3 (residues 635 to 749, from UniProtKB/Swiss-Prot entry Q9JLU4), was PCR-amplified using the primers 5' –CGG GAT CCG AGA ACC TGT ACT TCC AGG GCT CAC ACA GTG ATT ATG TCA TTG ATG and 5' –CCG GAA TTC TTA CTT ATC GTC ATC GTC CTT GTA GTC TGC GCC GAG CAC TAT CCT CCT CTG G, as described in Chapter 2 of this document. The PCR-amplified inserts were digested with BamHI/EcoRI and ligated into a BamHI/EcoRI-digested pGEX-4T-2 vector (GE Healthcare) for GST-TEV-Shank3PDZ-FLAG.

In order to replace the C-terminal FLAG tag with a His<sub>6</sub> tag, the following primers and their reverse complements were ordered and used in two 2-stage PCR protocol<sup>24</sup> for QuikChange® Site-directed Mutagenesis (the first to remove the FLAG tag, the second to insert the His<sub>6</sub> tag): 5' –GC CAG AGG AGG ATA GTG CTC GGC GCA TAA GAA TTC CCG GGT CGA CTC GAG CG- and 5' –GC CAG AGG AGG ATA GTG CTC GGC GCA CAC CAC CAC CAC CAC CAC TAA GAA TTC CCG GGT CGA CTC GAG CG-.

The PDZ domain of Shank3 (residues 635 to 749) was PCR-amplified using the primers 5' –CGG GAT CCG AGA ACC TGT ACT TCC AGG GCT CAC ACA GTG ATT ATG TCA TTG ATG- (same as above) and 5' -CCG CTC GAG GCG CCG AGC ACT ATC CTC CTC TGG-. The PCR-amplified inserts were digested with BamHI/XhoI and ligated into a BamHI/XhoI-digested pGBH vector (provided by Dr. H.-Y. Hu) for GB1-TEV-Shank3PDZ-His<sub>6</sub>.

The codon-optimized Fv-Shank3PDZ construct (residues 635 to 749 of Shank3) was ordered from Bio Basic Inc. A maxiprep of the DNA was digested with NdeI/XhoI, and the desired gene was ligated into a NdeI/XhoI-digested pET-24a(+) vector.

In order to insert a linker (10 amino acids, GGGENLYFQG) between the Fv and Shank3PDZ domains, the following primer and its reverse complement were ordered and used in a 2-stage PCR protocol<sup>24</sup> for QuikChange® Site-directed Mutagenesis: 5' –CTG GAG GGA TCC GAA TTC GGC GGT GGC GAG AAC CTG TAC TTC CAG GGC TCT CAT TCT GAT TAC GTT ATC-.

The full nucleotide sequence for the final Fv-Shank3PDZ construct used in these experiments is as follows: 5' –GAT ATA CAT ATG GGC GTT CAG GTT GAA ACT ATC TCT CCG GGT GAT GGC CGT ACC TTC CCG AAA CGC GGC CAG ACC TGC GTT GTT CAT TAT ACC GGT ATG CTG GAA GAT GGC AAA AAG GTG GAT TCT AGC CGT GAC CGC AAC AAA CCA TTC AAA TTT ATG CTG GGT AAA CAA GAA GTG ATC CGC GGT TGG GAA GAA GGT GTG GCG CAG ATG TCC GTT GGC CAG CGT GCA AAG CTG ACC ATT TCC CCG GAC TAC GCC TAC GGC GCG ACT GGT CAC CCG GGT ATC ATT CCG CCG CAC GCC ACC CTG GTT TTT GAC GTA GAA CTG CTG AAA CTG GAG GGA TCC GAA TTC GGC GGT GGC GAG AAC CTG TAC TTC CAG GGC TCT CAT TCT GAT TAC GTT ATC GAC GAT AAA GTT GCT ATC CTG CAG AAG CGT GAT CAC GAG GGT TTC GGT TTC GTG CTG CGT GGC GCT AAA GCG GAA ACT CCG ATT GAG GAA TTT ACC CCG ACT CCG GCT TTC CCA GCA CTG CAG TAT CTG GAA AGC GTG GAC GTA GAG GGT GTA GCT TGG AAA GCG GGT CTG CGT ACC GGT GAC TTC CTG ATC GAA GTT AAT GGC GTG AAC GTT GTT AAA GTT GGT CAC AAA CAG GTG GTA GGC CTG ATC CGC CAA GGT GGC AAC CGC CTG GTG ATG AAA GTT GTT TCC GTA ACT CGT AAA CCG GAA GAA GAC AGC GCA CGT CGT CTC GAG CAC CAC CAC CAC CAC TGA- 3'.

The PDZ domain of Shank3 (residues 635 to 749) was PCR-amplified using primers to introduce a 5' NdeI restriction site and a 3' XhoI restriction site. The PCR-amplified inserts were digested with NdeI/XhoI and inserted into a modified pET32a vector (Novagen), yielding the His<sub>6</sub>-TEV-Shank3PDZ construct.

***Expression and purification of Shank3 PDZ domain constructs.*** *E. coli* BL21 codon plus™ (DE3)-RIL or RP cells were transformed with each of the expression plasmids. Cells were amplified in 0.5 L of an auto-inducing medium (ZYM-5052, see reference for composition details) at 25 °C for 24 hours,<sup>25</sup> or in 1 L of Luria-Bertani media (at 37 °C until an optical density of 0.6-0.8 was achieved, then at 16 °C for 16 hours after induction with IPTG).

For GST-Shank3PDZ, cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (100 mM EDTA, 10% glycerol, 1% TritonX-100, 1 mg/mL Lysozyme in PBS, pH 7.4) containing protease inhibitor cocktail III (Calbiochem). Cells were lysed by ultrasonication (Branson Sonifier 450 at 40% power with a 40% duty cycle for 4 min at 4 °C). After addition of 1 mL of a 1 mM solution of DTT in water, the lysate was cleared by centrifugation (15,334 g, 40 minutes, 4 °C). The GST-Shank3 was purified with 5 mL of Glutathione Sepharose™ 4 Fast Flow (GE Healthcare) by batch binding. The bound proteins were washed with 20 volumes of PBS (pH 7.4) and eluted with 2 volumes of a 10 mM glutathione in 50 mM Tris (pH 8) solution.

For GB1-Shank3PDZ, Fv-Shank3PDZ, and His<sub>6</sub>-TEV-Shank3PDZ, cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 10 mM imidazole, 2 mM β-mercaptoethanol, 1% NP-40, 1 mg/mL Lysozyme, pH 8.0) containing protease inhibitor cocktail III (Calbiochem). Cells were lysed by ultrasonication (Branson Sonifier 450 at 40% power with a 40% duty cycle for 4 min at 4 °C). The lysate was cleared by centrifugation (15,334 g, 40 minutes, 4 °C). The proteins were purified with 5 mL of Ni-NTA Agarose (Qiagen) by batch-binding. The bound proteins were washed with 7 volumes of wash buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 20 mM imidazole, 2 mM β-mercaptoethanol, pH 8.0) and eluted with 2 volumes of elution buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 250 mM imidazole, 2 mM β-mercaptoethanol, pH = 8.0).

For all constructs, fractions were analyzed by 12% SDS-PAGE followed by Coomassie staining and Western blotting (anti-FLAG or anti-His<sub>6</sub>). Fractions containing the

protein were pooled, diluted and dialyzed against PBS (pH 7.4). Protein concentrations were measured by determining the absorption at 280 nm in 6 M guanidinium chloride. Purified protein solutions were aliquoted, flash-frozen and stored at  $-80^{\circ}\text{C}$  in PBS (pH 7.4) until use.

For TEV cleavage of the His<sub>6</sub>-TEV-Shank3 construct, the protein (in 10 mL volume) was dialyzed into 1x TEV buffer (50 mM Tris HCl pH 8, 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 2.5% glycerol), 1.5 mg of TEV protease was added to the dialysis cartridge, and the cleavage was allowed to proceed under dialysis conditions for 3 days at 4 degrees Celsius. Completion of the TEV cleavage reaction was monitored by Coomassie and Western blot analysis.

***Controlled dimerization of Fv-Shank3-His<sub>6</sub>.*** Fv-Shank3PDZ was dimerized with the AP20187 small molecule dimerizer supplied by Ariad. The material was reconstituted in ethanol and stored as specified in the Argent™ Regulated Homodimerization Kit v. 2.0. Prior to titrations, Fv-Shank3PDZ was incubated with AP20187 (exactly half the concentration of the Fv-Shank3 concentration, no greater than 60  $\mu\text{M}$  protein and 30  $\mu\text{M}$  AP20187) on ice for 5 minutes. This stock solution of dimerized Fv-Shank3PDZ was then used in titrations. Figure S2 shows that Fv-Shank3PDZ + AP20187 runs as a homogeneous dimer on an analytical SEC column, confirming that a homogeneous solution of dimerized protein is obtained by this method.

***Size exclusion chromatography (SEC).*** Each protein sample (100  $\mu\text{L}$ , 10-50  $\mu\text{M}$ ) was injected onto a 200  $\mu\text{L}$  injection loop of the FPLC. These samples were run with a solution of PBS buffer (pH 7.4) through a Superdex® 75 analytical SEC column. Blue Dextran was run on the column to determine the void volume. Either BioRad Gel Filtration Standard or Ribonuclease A, Ovalbumin, and Albumin standards were run, and a calibration curve of  $K_{av}$  vs.  $\log(\text{MW})$  allowed for subsequent determination of the molecular weight of the Shank construct protein samples through equation (1) below (where  $V_e$  = elution volume;  $V_o$  = void volume;  $V_c$  = column volume = 24 mL).

$$K_{av} = \frac{V_e - V_o}{V_c - V_o} \quad (1)$$



**Fluorescence titrations.** Fluorescence titrations were performed to determine the  $K_D$  of fluorescent peptide / protein pairs. For each titration, the peptide concentration was kept constant (between 1 and 20  $\mu\text{M}$ ) and the protein concentration was varied from values close to the expected  $K_D$  to saturation of fluorescence signal. A separate solution (150  $\mu\text{L}$ ) in PBS buffer (pH 7.4) was prepared for each protein concentration. Dissociation constants were evaluated using SPECFIT/32™ Global Analysis System for Windows (version 3.0.39), taking into account wavelengths from 432 nm to 730 nm. The results from at least three independent titrations were averaged together, and the average with the standard deviation is reported here.

**Competitive titrations.** Competition experiments were performed on a PBS (pH 7.4) solution (450  $\mu\text{L}$ , containing either 10  $\mu\text{M}$  fluorescent peptide and 10  $\mu\text{M}$  Fv-Shank3 PDZ domain or 2.5  $\mu\text{M}$  fluorescent peptide and 2.5  $\mu\text{M}$  Fv-Shank3 PDZ domain in the presence of 1.25  $\mu\text{M}$  AP20187). The concentrated ( $> 5 \text{ mM}$ ) non-fluorescent peptide solution was added sequentially in small volumes to achieve a range of concentrations (from 1 nM to 30  $\mu\text{M}$ ). The  $K_D$  of the non-fluorescent peptide was obtained by fitting the relative fluorescence increase (RFI, the quotient of the fluorescence intensity at the wavelength of maximum emission divided by the fluorescence intensity of peptide without the presence of protein at that wavelength) to the logarithm of the non-fluorescent peptide concentration ( $\text{Log}[\text{nFP}]$ , Molar). This was performed using the Competitive Binding: One Site model from GraphPad Prism 5 software, which uses the following equations:

$$\log(EC50) = \log\left(10^{\log(K_i)}\left(1 + \frac{[FP]}{K_D^{Fluo}}\right)\right) \quad (2)$$

$$RFI = \min_{RFI} + \frac{\max_{RFI} - \min_{RFI}}{1 + 10^{(\log[\text{nFP}] - \log(EC50))}} \quad (3)$$

where [FP] is the concentration of the fluorescent peptide,  $K_D^{Fluo}$  is the dissociation constant of the fluorescent peptide with the corresponding Fv-Shank3PDZ form (constant, same units as [FP]),  $\min_{RFI}$  is 1 and  $\max_{RFI}$  is the RFI in the absence of competitive ligand. The results of at least three independent competition titrations were averaged together (and the standard deviation of these values was calculated) to yield the values reported.

## Acknowledgements

I am grateful for the gift of the Shank3 cDNA from Dr. Daniel Choquet (CNRS), the gift of the pGBH vector from Dr. Hong-Yu Hu (Shanghai Institutes for Biological Sciences), and the gift of the Argent™ regulated homodimerization kit from Ariad ([www.ariad.com/regulationkits](http://www.ariad.com/regulationkits)). I would like to thank Dr. Matthieu Sainlos, Dr. Daniel Choquet, and Dr. Tim Clackson of Ariad for helpful discussion.

## References

1. Iskenderian-Epps, W. S.; Imperiali, B. "Modulation of Shank3 PDZ domain ligand-binding affinity by dimerization." *ChemBioChem* **2010**, *11*, 1979-1984.
2. Hayashi Mariko, K.; Tang, C.; Verpelli, C.; Narayanan, R.; Stearns Marissa, H.; Xu, R.-M.; Li, H.; Sala, C.; Hayashi, Y. "The postsynaptic density proteins Homer and Shank form a polymeric network structure." *Cell* **2009**, *137*, 159-171.
3. Hung, A. Y.; Sheng, M. "PDZ domains: structural modules for protein complex assembly." *J. Biol. Chem.* **2002**, *277*, 5699-5702.
4. Kim, E.; Sheng, M. "PDZ domain proteins of synapses." *Nat. Rev. Neurosci.* **2004**, *5*, 771-781.
5. Naisbitt, S.; Kim, E.; Tu, J. C.; Xiao, B.; Sala, C.; Valtschanoff, J.; Weinberg, R. J.; Worley, P. F.; Sheng, M. "Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin." *Neuron* **1999**, *23*, 569-582.
6. Tu, J. C.; Xiao, B.; Naisbitt, S.; Yuan, J. P.; Petralia, R. S.; Brakeman, P.; Doan, A.; Aakalu, V. K.; Lanahan, A. A.; Sheng, M.; Worley, P. F. "Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins." *Neuron* **1999**, *23*, 583-592.
7. Stiffler, M. A.; Grantcharova, V. P.; Sevecka, M.; MacBeath, G. "Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays." *J. Am. Chem. Soc.* **2006**, *128*, 5913-5922.
8. Sainlos, M.; Iskenderian, W. S.; Imperiali, B. "A general screening strategy for peptide-based fluorogenic ligands: Probes for dynamic studies of PDZ domain-mediated interactions." *J. Am. Chem. Soc.* **2009**, *131*, 6680-6682.
9. Romorini, S.; Piccoli, G.; Jiang, M.; Grossano, P.; Tonna, N.; Passafaro, M.; Zhang, M.; Sala, C. "A functional role of postsynaptic density-95-guanylate kinase-associated protein complex in regulating Shank assembly and stability to synapses." *J. Neurosci.* **2004**, *24*, 9391-9404.
10. Sheng, M.; Kim, E. "The Shank family of scaffold proteins." *J. Cell Sci.* **2000**, *113*, 1851-1856.
11. Im, Y. J.; Lee, J. H.; Park, S. H.; Park, S. J.; Rho, S.-H.; Kang, G. B.; Kim, E.; Eom, S. H. "Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization." *J. Biol. Chem.* **2003**, *278*, 48099-48104.

12. Im, Y. J.; Kang, G. B.; Lee, J. H.; Park, K. R.; Song, H. E.; Kim, E.; Song, W. K.; Park, D.; Eom, S. H. "Structural basis for asymmetric association of the betaPIX coiled coil and shank PDZ." *J. Mol. Biol.* **2010**, *397*, 457-466.
13. Spencer, D. M.; Wandless, T. J.; Schreiber, S. L.; Crabtree, G. R. "Controlling signal transduction with synthetic ligands." *Science* **1993**, *262*, 1019-1024.
14. Clackson, T.; Yang, W.; Rozamus, L. W.; Hatada, M.; Amara, J. F.; Rollins, C. T.; Stevenson, L. F.; Magari, S. R.; Wood, S. A.; Courage, N. L.; Lu, X.; Cerasoli, F., Jr.; Gilman, M.; Holt, D. A. "Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity." *Proc. Natl. Acad. Sci.* **1998**, *95*, 10437-10442.
15. Bao, W.-J.; Gao, Y.-G.; Chang, Y.-G.; Zhang, T.-Y.; Lin, X.-J.; Yan, X.-Z.; Hu, H.-Y. "Highly efficient expression and purification system of small-size protein domains in Escherichia coli for biochemical characterization." *Prot. Exp. Purif.* **2006**, *47*, 599-606.
16. Dirr, H.; Reinemer, P.; Huber, R. "X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function." *Eur. J. Biochem.* **1994**, *220*, 645-661.
17. Long, J.-F.; Tochio, H.; Wang, P.; Fan, J.-S.; Sala, C.; Niethammer, M.; Sheng, M.; Zhang, M. "Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95." *J. Mol. Biol.* **2003**, *327*, 203-214.
18. Wang, W. W., Jingwei; Zhang, Xu; Liu, Maili; Zhang, Mingjie " Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95 " *J. Am. Chem. Soc.* **2009**, *131*, 787-796.
19. Fanning, A. S.; Anderson, J. M. "Zonula occludens-1 and -2 are cytosolic scaffolds that regulate the assembly of cellular junctions." *Ann. N. Y. Acad. Sci.* **2009**, *1165*, 113-120.
20. Wu, J.; Yang, Y.; Zhang, J.; Ji, P.; Du, W.; Jiang, P.; Xie, D.; Huang, H.; Wu, M.; Zhang, G.; Shi, Y. "Domain-swapped dimerization of the second PDZ domain of ZO2 may provide a structural basis for the polymerization of claudins." *J. Biol. Chem.* **2007**, *282*, 35988-35999.
21. Im, Y. J.; Park, S. H.; Rho, S.-H.; Lee, J. H.; Kang, G. B.; Sheng, M.; Kim, E.; Eom, S. H. "Crystal structure of GRIPI PDZ6-peptide complex reveals the structural basis for class II PDZ target recognition and PDZ domain-mediated multimerization." *J. Biol. Chem.* **2003**, *278*, 8501-8507.
22. van Ham, M.; Hendriks, W. "PDZ domains-glue and guide." *Mol. Biol. Rep.* **2003**, *30*, 69-82.
23. Sainlos, M.; Imperiali, B. "Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization." *Nat. Protoc.* **2007**, *2*, 3201-3209.
24. Wang, W.; Malcolm, B. A. "Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis." *BioTechniques* **1999**, *26*, 680-682.
25. Studier, F. W. "Protein production by auto-induction in high-density shaking cultures." *Prot. Expr. Purif.* **2005**, *41*, 207-234.

## Chapter 5

# Development of Chemical Probes for Investigation of the *C. elegans* Lin-10 PDZ Domains

### Introduction

The study of neuronal receptors and proteins is aided immensely by *C. elegans* investigations involving genetic manipulation, electrophysiology, behavioral models, and fluorescence studies. *C. elegans* is a multicellular eukaryote, and is a well-defined model organism. Each worm contains exactly 302 neurons, and their connectivity pattern has been completely mapped. Novel proteins for neuronal receptor function have been discovered in the worm, and have led to a better understanding of neuronal signaling in vertebrates.

There are at least ten glutamate receptor subunits that have been found in *C. elegans*, two of which (GLR-1 and GLR-2) are clear homologues of AMPA-type mammalian receptor subunits based on sequence identity.<sup>1</sup> Studies toward understanding the localization of these receptors in *C. elegans* provide direct evidence that PDZ proteins are required for localization of the GLR-1 neurotransmitter receptor to central synapses *in vivo*. Specifically, these studies point to the PDZ-domain containing protein Lin-10, which is expressed in *C. elegans* vulval epithelial cells and in neurons.<sup>2</sup>

Utilizing GFP fusion constructs of GLR-1, researchers have shown that proper localization and function of GLR-1 in both neurons and epithelia requires Lin-10. However, the mechanism by which Lin-10 assists GLR-1 is not understood. The simplest model for the role of Lin-10 in the proper assembly of glutamatergic synapses invokes the direct binding of the Lin-10 PDZ domain(s) to the C-terminus of GLR-1. However, researchers have been unable to detect binding of Lin-10 to GLR-1 in a variety of *in vitro* assays, suggesting that Lin-10 plays a more indirect role in GLR-1 receptor targeting.<sup>3</sup>

GLR-1 is a class I PDZ domain ligand, which has the C-terminal consensus sequence [T/S]-X- $\phi$  (where  $\phi$  is a hydrophobic amino acid). This is in contrast to class II PDZ domain ligands, which display a  $\phi$ -Xaa- $\phi$  C-terminal consensus sequence. Interestingly, while GLR-1 synaptic localization is disrupted in Lin-10 null mutants, it can be restored via changes to the

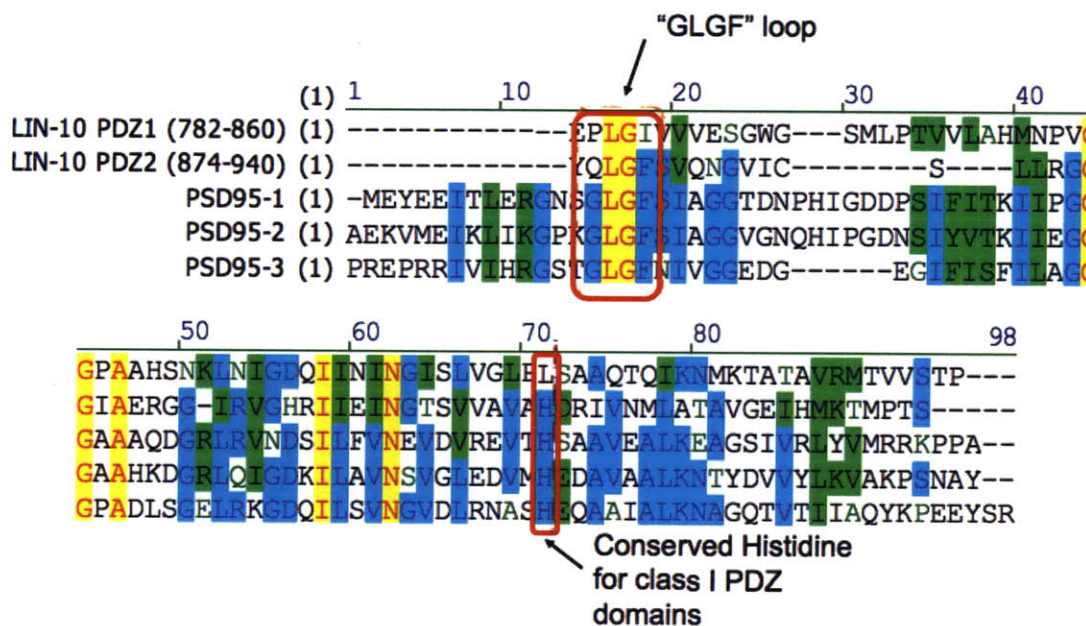
C-terminus of GLR-1. Specifically, replacement of the GLR-1 C-terminal class I consensus sequence (-TAV) with a class II sequence (-FYV) restores GLR-1 synaptic localization in these mutants. This demonstrates that these *C. elegans* neurons utilize at least two separate post-synaptic localization pathways, one of which involves the Lin-10 protein and both of which have yet to be completely defined.<sup>3</sup>

Due to the observed role of Lin-10 in receptor stabilization and function, and the likelihood of PDZ domain involvement in these interactions, probes for the PDZ domains of Lin-10 were targeted in this study. Specifically, the goal was to screen for optimal fluorescent probes, to adapt the probe design for *in vivo* studies, and to use them to both visualize Lin-10 localization in *C. elegans* and to study the effects of competing out natural Lin-10 PDZ-ligand interactions.

## **Results and Discussion**

### ***5-1. Initial ligand screening for Lin-10 probes***

In order to learn more about the function of the Lin-10 protein in *C. elegans*, probes were designed for its PDZ domains, Lin-10 PDZ-12. These PDZ domains are organized in tandem, similar to the first two PDZ domains of PSD-95. However, unlike the class I PSD-95 PDZ-12, this Lin-10 tandem appears to contain one class I domain and one class II domain (see sequence alignment, Figure 5.1).



**Figure 5.1.** Sequence alignment of PDZ domains from *C. elegans* Lin-10 and *R. norvegicus* PSD-95. The PDZ GLGF loop is present in Lin-10, albeit with modifications to the first and last residue of the consensus sequence. The class I conserved His residue is present in Lin-10 PDZ-2, but not PDZ-1, indicating that PDZ-1 is a class II-like PDZ domain. Lin-10 PDZ-2 displays 89% sequence homology with the class I PDZ-2 of PSD-95.

The Lin-10 PDZ domains were cloned as GB1 fusion proteins, since the GB1 domain has shown favorable properties as a fusion tag for the PDZ domain constructs. Due to the proximity of the two domains in Lin-10, they could function as a cooperative tandem; thus, the domains were tested both as a tandem and in isolation. The following constructs were cloned, expressed, and purified: GB1-Lin-10PDZ-12-His<sub>6</sub>, GB1-Lin-10PDZ-1-His<sub>6</sub>, and GB1-Lin-10PDZ-2-His<sub>6</sub>.

Due to the lack of knowledge regarding Lin-10 PDZ domain binding partners, a variety of class I and II probes were tested. The probes listed in Table 2.1 of chapter 2 showed negligible fluorescence increase with the Lin-10 PDZ domains. The remaining probes tested are listed in Table 5.1, along with their characterization. Interestingly, most ligands showed negligible fluorescence increase in the presence of the Lin-10 PDZ constructs, except for those based on the class I C-terminal peptide of the AN2 protein, a transmembrane protein implicated in the control of cell migration.<sup>4,5</sup> Even the GKAP probes, which like AN2 probes bind to the Shank3 PDZ domain, showed less than 2-fold increase in the presence of



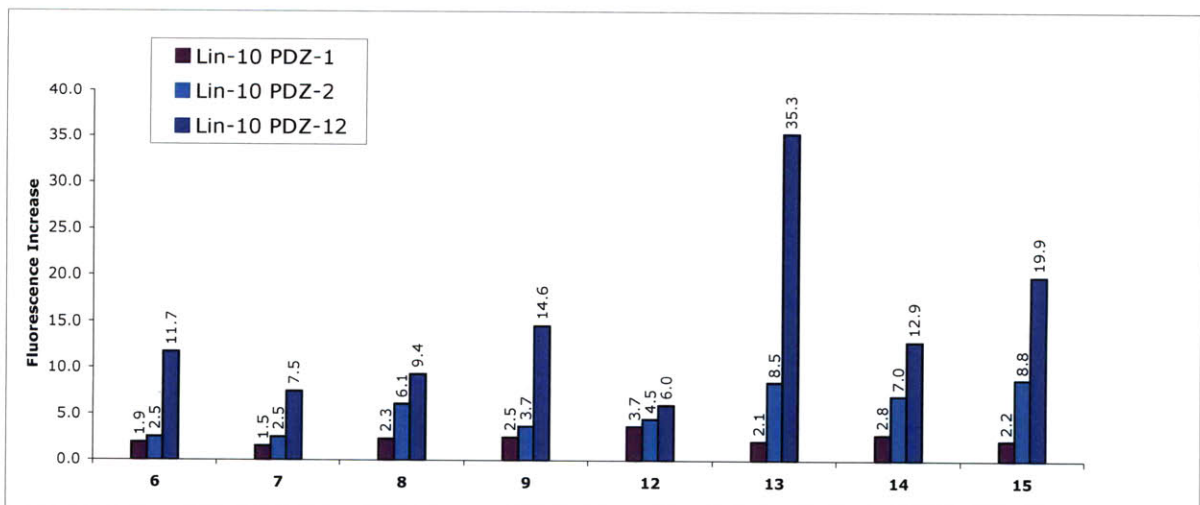
Lin-10 protein. The class I Kv1.4,<sup>4</sup> CRIPT, and Stargazin probes also exhibited negligible fluorescence increase, as did the class II GluR2<sup>4</sup> probe.

#	Origin	Peptide Sequence <sup>a</sup>	Formula	Mass expected	Mass <sup>b</sup> obtained [MH] <sup>+</sup>
1	Kv1.4	Ac-SNAK $\gamma$ ETDV-COOH	C <sub>54</sub> H <sub>82</sub> N <sub>14</sub> O <sub>21</sub>	1262.6	1262.9
2	Kv1.4	Ac-SNAK $\beta$ VETDV-COOH	C <sub>55</sub> H <sub>84</sub> N <sub>14</sub> O <sub>21</sub>	1276.7	1278.7
3	Kv1.4	Ac-SNAK $\gamma$ VETDV-COOH	C <sub>56</sub> H <sub>86</sub> N <sub>14</sub> O <sub>21</sub>	1290.7	1292.7
4	GKAP	Ac-NRRYIP $\beta$ AQTRL-COOH	C <sub>82</sub> H <sub>130</sub> N <sub>26</sub> O <sub>21</sub>	1815.4	1817.1
5	AN2	Ac-NRRPARLNGQY $\alpha$ V-COOH	C <sub>76</sub> H <sub>117</sub> N <sub>27</sub> O <sub>21</sub>	1744.3	1741.0
6	AN2	Ac-NRRPARL $\alpha$ GQYWV-COOH	C <sub>83</sub> H <sub>121</sub> N <sub>27</sub> O <sub>20</sub>	1816.4	1814.4
7	AN2	Ac-NRRPARL $\beta$ GQYWV-COOH	C <sub>84</sub> H <sub>123</sub> N <sub>27</sub> O <sub>20</sub>	1830.4	1828.4
8	AN2	Ac-NRRPARL $\beta$ GQYWL-COOH	C <sub>85</sub> H <sub>125</sub> N <sub>27</sub> O <sub>20</sub>	1844.4	1842.7
9	AN2	Ac-NRRPARL $\gamma$ GQYWV-COOH	C <sub>85</sub> H <sub>125</sub> N <sub>27</sub> O <sub>20</sub>	1844.4	1842.6
10	AN2	Ac-NRRPARL $\delta$ GQYWV-COOH	C <sub>86</sub> H <sub>127</sub> N <sub>27</sub> O <sub>20</sub>	1858.4	1856.6
11	GluR2	Ac-NRRNVY $\beta$ ESVKI-COOH	C <sub>77</sub> H <sub>120</sub> N <sub>24</sub> O <sub>23</sub>	1749.2	1754.4

**Table 5.1.** Probes used in the initial screening of Lin-10 PDZ domain constructs. <sup>a</sup> $\alpha$  = Dap(4-DMAP),  $\beta$  = Dab(4-DMAP),  $\gamma$  = Orn(4-DMAP),  $\delta$  = Lys(4-DMAP), the NRR sequence was added on the N-terminus of peptides to improve solubility at high concentrations, as low solubility of the probe leads to high background fluorescence signal due to aggregation.

<sup>b</sup> Masses were determined by MALDI-TOF spectroscopy.

AN2 probes 6-9 from Table 5.1 yielded reasonable fluorescence increases with the Lin-10 PDZ domain constructs, and these results are displayed in the first four entries of the Figure 5.2 bar graph. It is evident from these results that the fluorescence increases for Lin-10 PDZ-1 and Lin-10 PDZ-2 were not additive, since the fluorescence increase of the Lin-10 PDZ-12 tandem was consistently greater than the sum of its single domain counterparts. This points to the fact that the domains may act in tandem, either (1) by structurally stabilizing one another and allowing for proper domain folding, or (2) by enhancing one another's ligand binding ability.



**Figure 5.2.** Fluorescence increase results for first generation probes **6-9** (Table 5.1) and second generation probes **12-15** (Table 5.3) with Lin-10 constructs.

Fluorescence titrations of these four first generation probes yielded dissociation constants and overall fluorescence increases reported in Table 5.2. Based upon both dissociation constant and relative fluorescence intensity increases, probes **6** and **9** were carried further.

#	Origin	Peptide Sequence <sup>a</sup>	K <sub>D</sub> (μM)	RFI <sup>b</sup>
<b>6</b>	AN2	Ac-NRRPARL $\alpha$ GQYWV-COOH	17.3 ± 1.6	19
<b>7</b>	AN2	Ac-NRRPARL $\beta$ GQYWV-COOH	36.2 ± 3.0	15
<b>8</b>	AN2	Ac-NRRPARL $\beta$ GQYWL-COOH	31.4 ± 2.3	18
<b>9</b>	AN2	Ac-NRRPARL $\gamma$ GQYWV-COOH	34.2 ± 3.1	40

**Table 5.2.** Dissociation constants and overall fluorescence increases from fluorescence titrations of probes **6-9** with Lin-10 PDZ-12. <sup>a</sup> $\alpha$  = Dap(4-DMAP),  $\beta$  = Dap(4-DMAP),  $\gamma$  = Orn(4-DMAP), the NRR sequence was added on the N-terminus of peptides to improve solubility. <sup>b</sup>RFI = relative fluorescence increase.

Interestingly, the AN2 peptide sequence could be classified as both a class I and class II ligand, since the Tyr residue at the -2 position of the peptide contains a hydrophobic aromatic ring and a hydroxyl group. Therefore, a second generation of probes was pursued, in which the -2 residue of probes **6** and **9** was replaced with both Ser and Phe to mimic class I



and class II ligands, respectively. Characterization of these second generation probes **12-15** is provided in Table 5.3, and initial fluorescence results with the Lin-10 protein constructs are shown in the last four entries of Figure 5.2.

#	Origin	Peptide Sequence <sup>a</sup>	Formula	Mass expected	Mass <sup>b</sup> obtained [MH] <sup>+</sup>
12	AN2	Ac-NRRPARL $\alpha$ GQYFV-COOH	C <sub>81</sub> H <sub>120</sub> N <sub>26</sub> O <sub>20</sub>	1777.4	1779.7
13	AN2	Ac-NRRPARL $\alpha$ GQYSV-COOH	C <sub>75</sub> H <sub>116</sub> N <sub>26</sub> O <sub>21</sub>	1717.3	1719.5
14	AN2	Ac-NRRPARL $\gamma$ GQYFV-COOH	C <sub>83</sub> H <sub>124</sub> N <sub>26</sub> O <sub>20</sub>	1805.4	1809.8
15	AN2	Ac-NRRPARL $\gamma$ GQYSV-COOH	C <sub>77</sub> H <sub>120</sub> N <sub>26</sub> O <sub>21</sub>	1745.3	1744.4

**Table 5.3.** Second generation probes **12-15** and their characterization. <sup>a</sup> $\alpha$  = Dap(4-DMAP),  $\gamma$  = Orn(4-DMAP), the *NRR* sequence was added on the N-terminus of peptides to improve solubility. <sup>b</sup>Masses were determined by MALDI-TOF spectroscopy.

The results from the second generation probes suggested that the Lin-10 tandem binds class I ligands, as probes **13** and **15** exhibit greater fluorescence increases than probes **12** and **14**. This was further corroborated by the dissociation constants of probes **13** and **14** for the Lin-10 PDZ-12 tandem, shown in Table 5.4. Overall, first generation probe **6** and the corresponding class I derivative **13** yielded the lowest dissociation constants for the Lin-10 tandem, and probe **13** yielded the highest fluorescence increase. Hence, these two probes were taken forward for further derivatization and application *in vivo*.

#	Origin	Peptide Sequence <sup>a</sup>	K <sub>d</sub> ( $\mu$ M)	RFI <sup>b</sup>
13	AN2	Ac-NRRPARL $\alpha$ GQYSV-COOH	18.7 $\pm$ 2.3	41
14	AN2	Ac-NRRPARL $\gamma$ GQYFV-COOH	29.7 $\pm$ 4.5	26

**Table 5.4.** Dissociation constants and overall fluorescence increases from fluorescence titrations of probes **13** and **14** with Lin-10 PDZ-12. <sup>a</sup> $\alpha$  = Dap(4-DMAP),  $\gamma$  = Orn(4-DMAP), the *NRR* sequence was added on the N-terminus of peptides to improve solubility. <sup>b</sup> RFI = relative fluorescence increase.

## 5-2. Cell-permeant second generation probes

In order to apply these Lin-10 probes to *in vivo* tests in *C. elegans*, probes **6** and **13** were functionalized with the 11-residue TAT sequence,<sup>6</sup> a truncated portion of the HIV-1 Tat protein basic domain that can confer cell permeability to appended ligands. The TAT sequence was appended to the N-terminus of probes **6** and **13**, and the resulting probes **16** and **17** in Table 5.5 were tested in titrations with the Lin-10 PDZ-12 tandem. In addition to these two probes, a third TAT-control probe, **18**, was synthesized. The C-terminal amide and Arg residues at positions (-)1 and (-)3 of the control probe prohibit it from binding to PDZ domains.

#	Origin	Peptide Sequence <sup>a</sup>	K <sub>D</sub> (μM)	RFI <sup>b</sup>
<b>16</b>	<b>AN2</b>	Ac- <b>YGRKKRRQRRRPARL</b> α <b>GQYWV</b> -COOH	33 ± 1.8	23
<b>17</b>	<b>AN2</b>	Ac- <b>YGRKKRRQRRRPARL</b> α <b>GQYSV</b> -COOH	29.0 ± 2.2	33
<b>18</b>	<b>control</b>	Ac- <b>YGRKKRRQRRRNTAN</b> α <b>RTTRPR</b> -NH <sub>2</sub>	560 ± 300	1.5

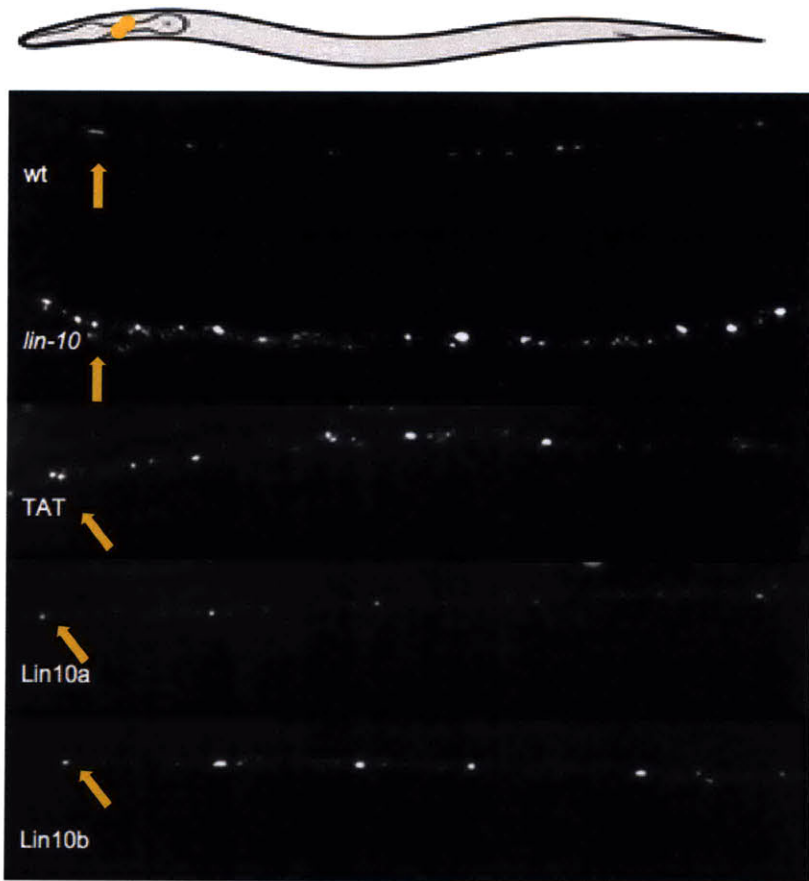
**Table 5.5.** Dissociation constants and overall fluorescence increases from fluorescence titrations of TAT probes **16** and **17** and control **18** with Lin-10 PDZ-12. <sup>a</sup>α = Dap(4-DMAP), *NRR* residues have been replaced with TAT sequence. <sup>b</sup> RFI = relative fluorescence increase.

The TAT-control peptide did not bind to the Lin-10 PDZ-12 construct or show any appreciable fluorescence increase in the presence of the protein. Both probes **16** and **17** showed slightly increased dissociation constants in comparison to their parent probes **6** and **13**, which was attributed to some small interference of the TAT sequence with binding to the tandem domains. However, the dissociation constants and relative fluorescence increases for both probes were still reasonable for initial trials in *C. elegans*, and therefore these probes were prepared for testing *in vivo* while additional avenues were pursued to increase probe affinity in a third generation of Lin-10 PDZ-12 probes.

### **5-3. Initial probe testing in *C. elegans***

Since Lin-10 has been shown to be involved in the localization of GLR-1 to synapses,<sup>3</sup> peptides **16** and **17** were tested for inhibitory activity against GLR-1 synaptic localization in *C. elegans*. Researchers in the Maricq laboratory injected these peptides into *C. elegans* expressing GLR1-GFP in a single pair of AVA neurons. Wild type expression in these neurons is shown in Figure 5.3, as is expression in a Lin-10 null mutant (*lin-10*). Note that in the AVA neurons, the expression of GLR1-GFP was more diffuse in the Lin-10 null background than in the wild type *C. elegans* (arrows point to AVA neurons). Also note that there was background expression of GLR1-GFP in other neurons of the worms, which accounts for the fluorescent signal elsewhere in the images. Based on these results, if the Lin-10 probes did indeed bind to the Lin-10 PDZ domain(s) and out-compete natural interactions with the GLR1 receptor, lack of localization of these receptors, i.e. more diffuse signal from GLR1-GFP, would be expected.

Unfortunately, treatments with the inhibitory peptides **16** and **17** did not have much of an effect on the localization of GLR1-GFP at the AVA neurons (and neither did treatment with control peptide **18**). TAT control, Lin10a, and Lin10b panels still showed punctae of GLR1-GFP localized at the AVA neurons, similar to the localization in the wild type worms. It should be noted that in these imaging experiments, a significantly strong effect was needed to ensure accurate interpretation and observation. Most likely the peptides used in these studies did not bind strongly enough to the Lin-10 PDZ domains to elicit an effect on GLR1-GFP localization. Hence, in order to repeat these types of experiments and obtain interpretable results, inhibitory peptides with dissociation constants in the low micromolar range, about an order of magnitude tighter than those of peptides **16** and **17**, would be necessary.



**Figure 5.3.** Effect of peptides on inhibition of GLR1-GFP synaptic punctae in a single pair of AVA neurons. AVA shown in yellow in the worm schematic at the top. Wild type expression of GLR1-GFP (wt), expression in Lin-10 null background (*lin-10*), wild type injected with control peptide **18** (TAT), wild type injected with inhibitory peptide **16** (Lin10a), wild type injected with inhibitory peptide **17** (Lin10b).



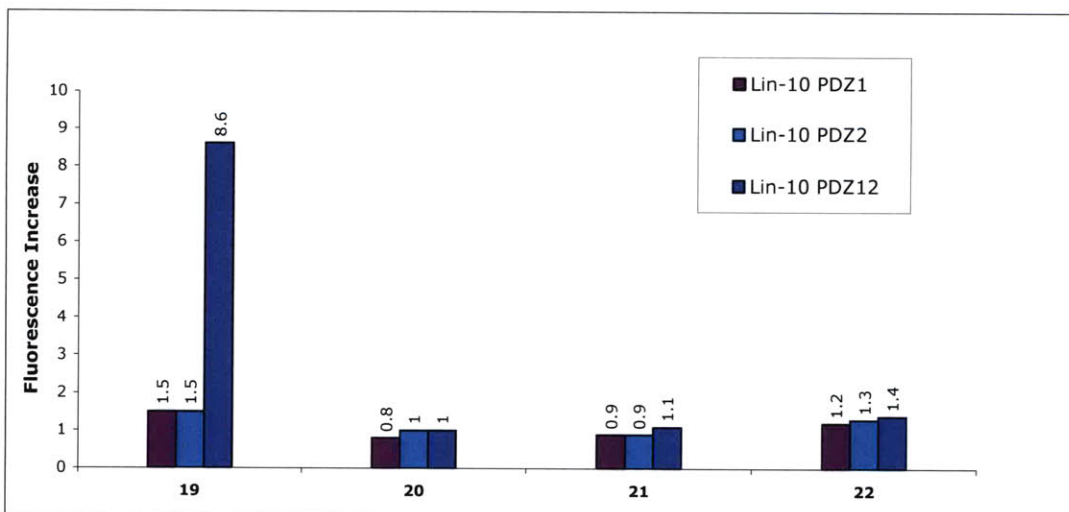
#### 5-4. Efforts toward third generation probes

Further investigations in *C. elegans* necessitate Lin-10 probes with greater affinity. In order to identify other potential peptide sequences for probes, ligands for the mammalian Lin-10 homologues (Mint proteins) were investigated. Two such ligands are presenilin<sup>7</sup> of the gamma-secretase protease complex and the motor protein kinesin Kif17.<sup>8</sup> Additionally, it has been reported that an autoinhibited conformation of the PDZ domains of Mint1 is mediated by molecular interaction between the conserved C-terminal tail of the protein and the first domain of the PDZ tandem.<sup>9</sup> Interestingly, the -3 residue of the Mint1 and Lin-10 C-terminal peptides is a proline. This residue is not commonly found at the -3 position among PDZ domain ligands. Given the conformational constraint imposed by a Pro residue, it is possible that the Pro allows for autoinhibition in both MINT1 and Lin-10 proteins. Hence, Mint1 and Lin-10 C-terminal peptides (C-peptides) were also tested as potential ligands. Since the Dap residue for the 4-DMAP fluorophore stood out in the previous screen of ligands for the Lin-10 domains, that linker length was kept constant in the new screen.

#	Origin	Peptide Sequence <sup>a</sup>	Formula	Mass expected	Mass <sup>b</sup> obtained [MH] <sup>+</sup>
19	Presenilin	Ac-DQLA $\alpha$ HQFYI-COOH	C <sub>68</sub> H <sub>90</sub> N <sub>16</sub> O <sub>19</sub>	1434.9	1430.7
20	Kinesin	Ac-SKSN $\alpha$ GSEPL-COOH	C <sub>52</sub> H <sub>78</sub> N <sub>14</sub> O <sub>20</sub>	1218.6	1220.6
21	Lin-10	Ac-LTGQ $\alpha$ QPQYI-COOH	C <sub>62</sub> H <sub>89</sub> N <sub>15</sub> O <sub>19</sub>	1337.8	1337.3
22	Mint-1	Ac-LTAQ $\alpha$ QPVYI-COOH	C <sub>63</sub> H <sub>92</sub> N <sub>14</sub> O <sub>18</sub>	1333.8	1338.7

**Table 5.6.** Probes based on Mint PDZ domain ligands, to screen with Lin-10 PDZ domain constructs. <sup>a</sup> $\alpha$  = Dap(4-DMAP). <sup>b</sup>Masses were determined by MALDI-TOF spectroscopy.

Probes **19-22** were screened against all three Lin-10 PDZ domain constructs, and the relative fluorescence intensity results are summarized in Figure 5.4. The presenilin probe **19** was the only one to yield promising results. Although the fluorescence increase for this probe with the Lin-10 tandem was lower than that of the other probes, titration analysis demonstrated low micromolar affinity for the Lin-10 PDZ-12 tandem (see Table 5.7), the lowest affinity of any of the Lin-10 probes yet.



**Figure 5.4.** Fluorescence increase results for probes **19-22** with Lin-10 constructs.

These promising results led to the derivatization of probe **19** with the TAT sequence (see probe **23**), to render ligand cell-permeant. Unfortunately, addition of the TAT sequence reduced the affinity of the probe for the Lin-10 tandem by more than 25-fold. Further efforts to ameliorate this undesired effect of the TAT sequence by lengthening the probe (to 17 residues of presenilin followed by the TAT sequence) helped somewhat, but the resulting probe **24** still bound more than 15-fold weaker to the Lin-10 tandem than probe **19**, and with a smaller fluorescence increase as well. Further efforts are required to obtain a probe for which cell permeability and tight binding (a  $K_D$  in the low micromolar range or less) are not exclusive properties. This may necessitate the use of a different method to confer cell permeability to the probes, perhaps one that does not require covalent attachment of the carrier to the cargo.

#	Origin	Peptide Sequence <sup>a</sup>	$K_D$ ( $\mu\text{M}$ )	RFI <sup>b</sup>
<b>19</b>	Pre	Ac-NRRDQLA $\alpha$ HQFYI-COOH	$1.9 \pm 0.6$	<b>33</b>
<b>23</b>	Pre	Ac-YGRKKRRQRRRDQLA $\alpha$ HQFYI-COOH	$53 \pm 4$	<b>27</b>
<b>24</b>	Pre	Ac-YGRKKRRQRRRYLVQPFADQLA $\alpha$ HQFYI-COOH	$31 \pm 4$	<b>10</b>

**Table 5.7.** Dissociation constants and RFI of probes **19**, **23**, and **24**. <sup>a</sup> $\alpha$  = Dap(4-DMAP),  $\lambda$  = norleucine (replaces native methionine), probes contain either *NRR* residues or TAT sequence. <sup>b</sup> RFI = relative fluorescence increase.

## Conclusions

The reach of the fluorescent probes for PDZ domains has been extended to those of *C. elegans*. Studies of these critical elements of the neural network in the worm model can afford information on learning and memory that is relevant to other organisms. Second generation probes based on the AN2 ligand sequence were tested in *C. elegans*, but the results were inconclusive. Eventually the effects of out-competing the Lin-10 PDZ tandem *in vivo*, which necessitates probes with tighter affinity for the Lin-10 tandem, will be studied. Toward this end, a probe based on the presenilin ligand sequence has been developed. This probe has the desired low micromolar affinity for Lin-10 PDZ-12, but unfortunately this affinity is greatly affected by the appendage of the cell-permeabilizing TAT sequence. Future directions should focus on the development of TAT-free methods of probe delivery, and / or the design of other TAT derivatives that do not interfere with protein binding. In addition, the monomeric probe may benefit from extension to a dimeric probe, since it binds to a tandem of two PDZ domains; this option has been explored for other PDZ domains and will be discussed in Chapter 6.

## Experimental Methods

### *Peptide Synthesis (SPPS), Fluorophore Synthesis, and Fluorophore Insertion.*

Peptides were obtained either manually or with an automated synthesizer (Advanced ChemTech automated synthesizer model 396  $\Omega$ ) by using standard Fmoc-based solid phase peptide synthesis procedures. Typically manual synthesis was performed on a 0.04 to 0.02 mmol scale with Fmoc-Val/Leu/Ile-NovaSyn® TGT resin (0.2 mmol/g) or Fmoc-PAL-PEG-PS resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents (6 equivalents each) and DIPEA (12 equivalents) in DMF or NMP. Coupling steps were conducted with a 50 mM solution of amino acid (~1 mL per 100 mg of resin) for 1 hour. Fmoc group removal was performed with a 20% solution of 4-methyl-piperidine in DMF (vol/vol) for 3 x 5 minutes. After removal of the N-terminal Fmoc group, the resulting free amine was capped with an acetyl group by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

Peptides incorporating the 4-DMAP environment-sensitive fluorophore were synthesized using the on-resin derivatization approach previously reported.<sup>10</sup> An Alloc-protected diaminoacid (either Fmoc-Dap(Alloc)-OH, Fmoc-Dab(Alloc)-OH or Fmoc-Orn(Alloc)-OH from AnaSpec) was initially inserted during the SPPS steps. After capping of the N-terminal amino group of the peptide sequence, the Alloc-protecting group was removed. The resin was resuspended in dry dichloromethane (20 mM in peptide) with a stream of N<sub>2</sub> bubbling through the solution for 5 minutes. Pd(PPh<sub>3</sub>)<sub>4</sub> (0.8 eq.) and phenylsilane (25 eq.) were then added to the solution, which was maintained under N<sub>2</sub> bubbling for another 15 minutes. The resin was then washed with dichloromethane and DMF, and the degassing/deprotection cycle was repeated two additional times. The resin containing the free amine was reacted with a solution of the anhydride form of 4-DMAP (4-dimethylaminophthalenedicarboxylic anhydride, 2 equiv., 50 mM) and DIPEA (4 equiv.) in NMP. The mixture was allowed to stir overnight. The resin was then washed with DMF and dichloromethane and ring closure was performed by using a solution of HBTU/HOBt (6 equiv., 50 mM) in NMP or DMF with DIPEA (12 eq.) for 2 hours. The resin was then washed with DMF and dichloromethane. The coupling/washing cycle was repeated two additional times in order to achieve full ring closure.



Peptides were deprotected and cleaved from the resin with a TFA/H<sub>2</sub>O/TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours. TFA was evaporated until the peptides precipitated. The peptides were then triturated and precipitated in cold ether before purification by reverse phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250x20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 350 nm. Peptides were stored lyophilized at -80 °C until use.

**Peptide Characterization and Quantification.** Peptide identity was confirmed by MALDI-TOF mass spectroscopy (MALDI-TOF, PerSeptive Biosystems Voyager) using DHB as a matrix. Purity was assessed by analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 min followed by 5-95% acetonitrile containing 0.1% TFA over 35 min in water containing 0.1% TFA at a flow rate of 1 mL/min). The final peptides were quantified using molar extinction coefficients of either 4-DMAP ( $\epsilon_{421\text{nm}} = 6480 \text{ M}^{-1}\text{cm}^{-1}$  in water) or *p*-nitrophenylalanine ( $\epsilon_{280\text{nm}} = 12500 \text{ M}^{-1}\text{cm}^{-1}$ ).

**PDZ Domain Cloning.** The cDNA encoding for the genes used for PDZ domain subcloning were provided by Villu Maricq (University of Utah; Lin-10). The PDZ domains of interest were cloned into the pGBH vector,<sup>11</sup> which was provided by Dr. Hong-Yu Hu (Shanghai Institutes for Biological Sciences), using EcoRI and XhoI restriction sites. A TEV cleavage site was incorporated between GB1 and the domain(s).

The first two PDZ domains of Lin-10, Lin-10 PDZ-12 (residues 797 to 976, from UniProtKB/Swiss-Prot entry O17583), were PCR-amplified and inserted into the pGBH vector.

The first PDZ domain of Lin-10, Lin-10 PDZ-1 (residues 797 to 893, from UniProtKB/Swiss-Prot entry O17583), was PCR-amplified and inserted into the pGBH vector.

The second PDZ domain of Lin-10, Lin-10 PDZ-2 (residues 887 to 976, from UniProtKB/Swiss-Prot entry O17583), was PCR-amplified and inserted into the pGBH vector.

**Expression and Purification of PDZ Domain Constructs.** *E. Coli* BL21 codon plus™ (DE3)- RP cells were transformed with each of the GB1 fusion protein expression plasmids. Cells were amplified in 0.5 L of an auto-inducing medium<sup>12</sup> (ZYM-5052, see reference for detailed composition) first at 37 °C for 4 hours followed by 12 hours at 16 °C. Cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 10 mM imidazole, 2 mM β-mercaptoethanol, 1% NP-40, 1 mg/mL Lysozyme, pH 8.0) containing protease inhibitor cocktail III (Calbiochem). Cells were lysed by ultrasonication (Branson Sonifier 450 at 40% power with a 40% duty cycle for 4 min at 4 °C). The lysate was cleared by centrifugation (15,334 g, 40 minutes, 4 °C). The proteins were purified with 5 mL of Ni-NTA Agarose (Qiagen) by batch-binding. The bound proteins were washed with 7 volumes of wash buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 20 mM imidazole, 2 mM β-mercaptoethanol, pH 8.0) and eluted with 2 volumes of elution buffer (300 mM NaCl, 50 mM PO<sub>4</sub><sup>3-</sup>, 250 mM imidazole, 2 mM β-mercaptoethanol, pH = 8.0). Fractions were analyzed by 12% SDS-PAGE followed by Coomassie staining and Western blotting (anti-FLAG). Fractions containing the protein were pooled and dialyzed against PBS (pH 7.4). Protein concentrations were measured using either the BCA assay (Pierce) with BSA as the reference standard or by determining the absorption at 280 nm in 6 M guanidinium chloride. Purified PDZ domains were aliquoted, flash-frozen and stored at -80 °C in PBS (pH 7.4) until use.

**Fluorescence Studies.** Fluorescence spectra were recorded on a Fluoromax 3 instrument (Horiba Jobin Yvon) in 1 cm path length quartz cells (100 mL nominal volume from Starna Cells). All measurements were performed at a constant temperature of 25 °C. Slit widths were 3 nm for excitation and 6 nm for emission. The 4-DMAP fluorophore was excited at 421 nm and the spectra were recorded between 432 nm and 730 nm (0.5 nm increments and 0.1 s integration time).

**Relative Fluorescence Increases.** The fluorescence increase measurements were performed at 25 °C in PBS buffer (pH 7.4) comparing solutions of (1) 10 μM of 4-DMAP-containing peptide alone in PBS and (2) 10 μM of 4-DMAP-containing peptide with 20 μM of PDZ domain construct in PBS. Blanks consisting of PBS for (1) and PBS with 20 μM of the

corresponding PDZ domain construct for (2) were subtracted from the respective spectra. The fluorescence increase for each series of peptide and PDZ domain construct was evaluated by comparing the fluorescence emission intensities of (2) and (1) at the wavelength of maximal emission of (2). Ratios were calculated over a 5 nm range centered on the wavelength of maximal emission of (2) and the values were averaged to yield the reported final ratio.

**Fluorescence Titrations.** For each titration, the peptide concentration was kept constant (5  $\mu$ M) and the protein concentration was varied from values lower than or close to the anticipated  $K_D$  to saturation of the fluorescence signal increase. A 150  $\mu$ L solution in PBS (pH 7.4) was prepared for each protein concentration. Dissociation constants were evaluated using SPECFIT/32™ Global Analysis System for Windows (version 3.0.39) after averaging at least three independent titrations and taking into account wavelengths from 432 to 730 nm.

**Injections in *C. elegans*.** Researchers in the Maricq laboratory performed pseudocoelomic injections on multiple worms expressing a GLR1-GFP construct. A 3 mM stock solution of either peptide **16**, **17**, or **18** was injected into the worm. After a short incubation time, the localization and clustering of the GFP signal in the worms was imaged and recorded.

## **Acknowledgements**

I would like to thank our collaborators Professor Villu Maricq, Dr. Fred Hörndli, and Dane Maxfield from the University of Utah, who performed the initial injections in *C. elegans*. Additionally, I would like to thank Dave Madsen from the University of Utah, who cloned the Lin-10 PDZ constructs discussed in this chapter.

## References

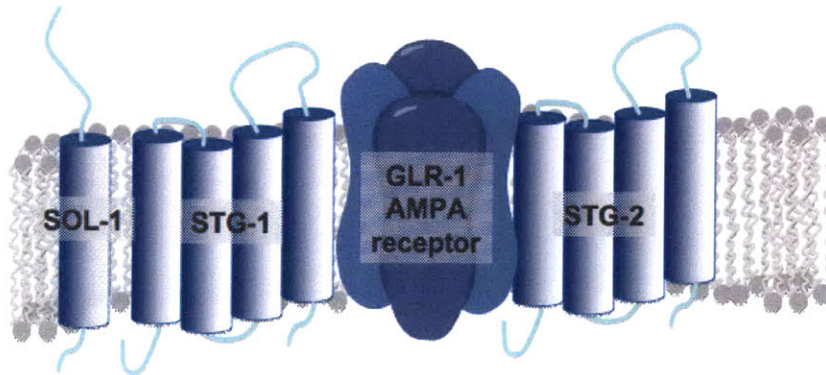
1. Brockie, P. J.; Madsen, D. M.; Zheng, Y.; Mellem, J.; Maricq, A. V. "Differential expression of glutamate receptor subunits in the nervous system of *Caenorhabditis elegans* and their regulation by the homeodomain protein UNC-42." *J. Neurosci.* **2001**, *21*, 1510-1522.
2. Whitfield, C. W.; Benard, C.; Barnes, T.; Hekimi, S.; Kim, S. K. "Basolateral localization of the *Caenorhabditis elegans* epidermal growth factor receptor in epithelial cells by the PDZ protein LIN-10." *Mol. Biol. Cell* **1999**, *10*, 2087-2100.
3. Rongo, C.; Whitfield, C. W.; Rodal, A.; Kim, S. K.; Kaplan, J. M. "LIN-10 is a shared component of the polarized protein localization pathways in neurons and epithelia." *Cell* **1998**, *94*, 751-759.
4. Stiffler, M. A.; Grantcharova, V. P.; Sevecka, M.; MacBeath, G. "Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays." *J. Am. Chem. Soc.* **2006**, *128*, 5913-5922.
5. Stegmuller, J.; Schneider, S.; Hellwig, A.; Garwood, J.; Trotter, J. "AN2, the mouse homologue of NG2, is a surface antigen on glial precursor cells implicated in control of cell migration." *J. Neurocytol.* **2002**, *31*, 497-505.
6. 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.* **1997**, *272*, 16010-16007.
7. Lau, K. F.; McLoughlin, D. M.; Standen, C.; Miller, C. C. "X11 alpha and x11 beta interact with presenilin-1 via their PDZ domains." *Mol. Cell Neurosci.* **2000**, *16*, 557-565.
8. Setou, M.; Nakagawa, T.; Seog, D. H.; Hirokawa, N. "Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport." *Science* **2000**, *288*, 1796-1802.
9. Long, J. F.; Feng, W.; Wang, R.; Chan, L. N.; Ip, F. C.; Xia, J.; Ip, N. Y.; Zhang, M. "Autoinhibition of X11/Mint scaffold proteins revealed by the closed conformation of the PDZ tandem." *Nat. Struct. Mol. Biol.* **2005**, *12*, 722-728.
10. Sainlos, M.; Imperiali, B. "Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization." *Nat. Protoc.* **2007**, *2*, 3201-3209.
11. Bao, W.-J.; Gao, Y.-G.; Chang, Y.-G.; Zhang, T.-Y.; Lin, X.-J.; Yan, X.-Z.; Hu, H.-Y. "Highly efficient expression and purification system of small-size protein domains in *Escherichia coli* for biochemical characterization." *Prot. Exp. Purif.* **2006**, *47*, 599-606.
12. Studier, F. W. "Protein production by auto-induction in high-density shaking cultures." *Prot. Expr. Purif.* **2005**, *41*, 207-234.

## Chapter 6

# Development and Evaluation of Dimeric Chemical Probes for the *C. elegans* Glutamate Receptor Complex

### Introduction

Neurotransmission is achieved through signaling between synapses, which is partially mediated by  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors (AMPA). In vertebrates and *Caenorhabditis elegans* (*C. elegans*), these receptors are associated with tetraspanning transmembrane proteins (TARPs), which are necessary for receptor function. *C. elegans* is a multicellular eukaryote that serves as an ideal model system for higher eukaryotes. The worm is one of the simplest organisms with a nervous system, and the connectivity and functions of its 302 neurons has been completely mapped, making it an ideal organism for *in vivo* neurobiological investigations. The critical components of the *C. elegans* AMPAR complex, including TARPs STG-1 and STG-2, are depicted in Figure 6.1. Synaptic transmission in *C. elegans*, mediated by glutamate-gated current (current evoked by stimulation with glutamate) depends on four proteins: *C. elegans* Stargazin-like proteins (STG-1 and STG-2), suppressor of lurcher movement defect protein (SOL-1), and glutamate receptor family protein (GLR-1).<sup>1-3</sup> STG-1 and STG-2, although distantly related by primary sequence to vertebrate TARPs, can functionally substitute for vertebrate TARPs in reconstitution experiments, displaying evolutionarily conserved roles for TARPs in the regulation of AMPARs. STG-1 and STG-2 can also substitute for one another in functional *C. elegans* receptors. Studies have clearly shown that STG-1 and STG-2 are necessary for the proper function of the GLR-1 receptor, since glutamate-gated current is not observed in the absence of these two proteins.<sup>4</sup> However, the exact functions of these TARP auxiliary proteins and their chemical modes of interaction with the GLR-1 receptor remain unclear. Further elucidation of the roles of these proteins in *C. elegans* may help to provide insight into the interactions and functions of vertebrate glutamate receptor complexes.



**Figure 6.1.** Proteins involved in the function of the ionotropic glutamate receptor GLR-1 complex in *C. elegans*.

Although the modes of interaction between GLR-1, SOL-1, STG-1, and STG-2 are currently unknown, it is possible that they interact via a PDZ domain-containing protein (to which the C-terminus of STG-2 binds). The C-terminus of STG-2 is a putative PDZ domain ligand, with a conserved hydrophobic C-terminal residue (see Figure 6.2). The Stargazin homologue of STG-2 in vertebrates is a known PDZ domain ligand, hence STG-2 may interact with a PDZ domain-containing protein as well. Note that, although STG-1 and STG-2 are functionally related in *C. elegans*, the STG-1 C-terminal sequence does not resemble a canonical PDZ domain ligand (Figure 6.2). For this reason, STG-2-mediated interactions were targeted in this study.

The interactions of the C-terminus of STG-2 may play a critical role in its function and also in the clustering of GLR-1, SOL-1, and STG-2 proteins of the AMPAR complex in *C. elegans*, similar to the role Stargazin plays in vertebrates. Dimeric Stargazin-based probes for the rat neuronal system have been utilized successfully to disrupt the Stargazin – PDZ domain interaction.<sup>5</sup> Specifically, dimeric probes (as opposed to the corresponding monomeric counterparts) are necessary to disrupt the native interactions, since Stargazin ligands bind cooperatively to the tandem PDZ domains (PDZ1 and PDZ2) of PSD-95. Tandem PDZ domains, which are linked closely together within a protein (typically about 9-12 residues apart), are known to bind synergistically to their target ligands; PDZ domains can also work together in concert by dimerization or multimerization (see Figure 6.3 for a schematic comparison). Synergistic binding of multiple PDZ domains thermodynamically enhances the binding affinity between PDZ domains and target monomeric ligands.<sup>6</sup> The presentation of two ligands in the form of a dimer mimics the stronger native multivalent interaction through

the same thermodynamic principles. Since Stargazin and STG-2 are evolutionarily conserved, it is likely that STG-2 may exhibit similar multivalent binding interactions with a PDZ domain-containing protein. This chapter describes the synthesis and evaluation of dimeric probes based on STG-2, designed to disrupt the binding of the STG-2 C-terminus to either monomeric or dimeric partner PDZ domains in *C. elegans*.

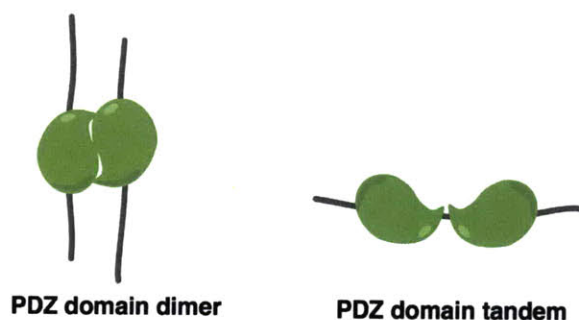
	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	
<b>Stargazin (309-323)</b>	D	S	L	H	A	N	T	A	N	R	R	T	T	P	V	COOH
<b>STG-1 (352-366)</b>	R	R	F	A	I	P	V	S	Q	G	Y	P	S	L	S	COOH
<b>STG-2 (365-379)</b>	Q	A	S	F	S	N	P	S	L	L	F	M	Q	D	V	COOH
<b>STG-2 Dimer</b>	<b>N</b>	<b>R</b>	<b>R</b>	<b>Y</b>	S	N	P	S	L	L	F	Nle	Q	D	V	COOH

**Figure 6.2.** Alignment of *R. norvegicus* Stargazin (Uniprot Q71RJ2), *C. elegans* STG-1 (Uniprot Q7YX75), and *C. elegans* STG-2 (Uniprot D9N123) C-terminal 15 residues. The sequence used for the dimer based on STG-2 is also shown (last entry); the presence of the tyrosine (Y) residue allows for quantification, and the *NRR* residues aid in solubility, since the *C. elegans* STG-2 C-terminal sequence is rather hydrophobic. Note that the methionine residue in the dimer entry is replaced with the isosteric norleucine in the dimer sequence, due to complications with methionine oxidation.

## Results and Discussion

### 6-1. Design of dimeric competitor for STG-2 protein

The binding partner(s) of *C. elegans* STG-2 are currently unknown, although the C-terminus of STG-2 (Figure 6.2) is a putative PDZ domain ligand. The chemical interaction between STG-2 and the GLR-1 complex in the worm model system is unknown, hence investigations were undertaken to determine whether or not the C-terminus of STG-2 is a critical component of this interaction.

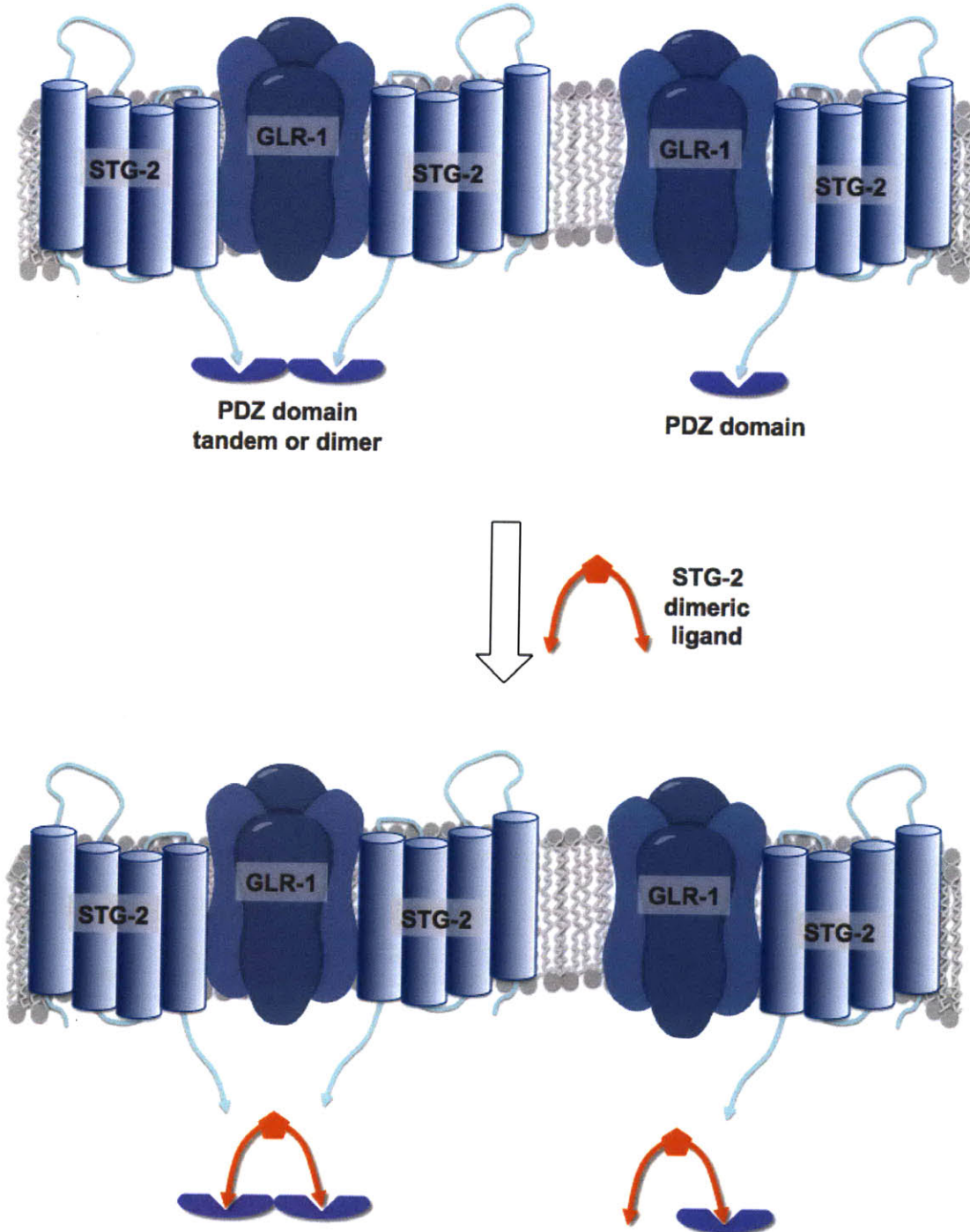


**Figure 6.3.** Schematic of dimer versus tandem PDZ domains.



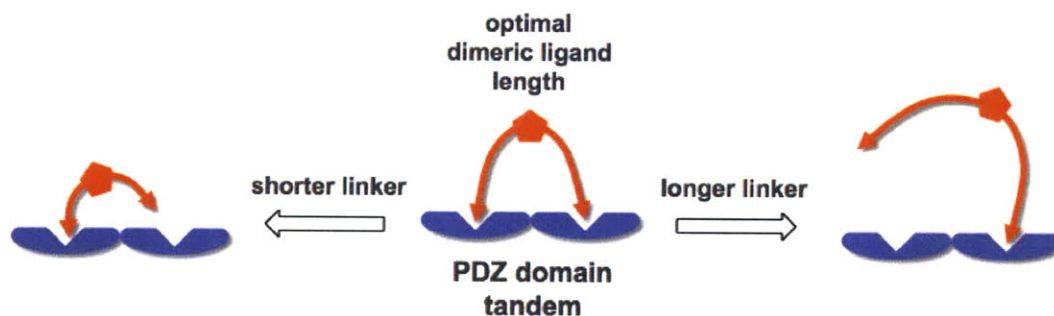
As discussed above, it has been shown that dimeric Stargazin-based PDZ domain ligands have been used successfully to out-compete out native multivalent interactions between Stargazin and PSD-95 in rat neurons. In this case, since the native interaction was multivalent (involving the binding of Stargazin ligands to a tandem of two PDZ domains), monovalent Stargazin-based ligands were not capable of inhibiting the native interaction.<sup>5</sup> Based on these results, a dimeric ligand of the Stargazin-related protein STG-2 was pursued. If STG-2 is not involved in multivalent interactions, but instead binds a partner PDZ domain in a monovalent fashion, as opposed to a multivalent interaction, the STG-2 dimer would still be capable of competing for this binding site, since it is composed of two monomeric units. In this scenario, the higher local concentration of monomer presented by the dimeric ligand could potentially enhance its ability to inhibit an interaction between STG-2 and the single PDZ domain. Both monovalent and multivalent modes of STG-2 interactions can therefore be probed with the STG-2 dimer, and these models are illustrated graphically in Figure 6.4.





**Figure 6.4.** Schematic of the proposed STG-2 dimeric ligand function. The dimeric ligand should out-compete interactions between the STG-2 C-terminus and PDZ domain(s) in either a multivalent or monovalent fashion. Note that multivalent interactions can arise between the ligand and either a tandem of PDZ domains (located within the same protein macromolecule) or a dimer or multimer of single PDZ domains from multiple proteins.

While the fluorescence-based monomeric probes developed for PDZ domains (described in detail in Chapter 2) demonstrated native sequence specificity with the incorporation 10 native PDZ domain ligand amino acid residues, the linker length considerations for dimeric probe design introduce an added layer of complexity. As depicted in Figure 6.5, dimeric ligands with too short of a linker length would not be able to bind both PDZ domains simultaneously, and dimeric ligands with longer linkers might have to pay an entropic penalty and consequently behave more like monomers in solution.

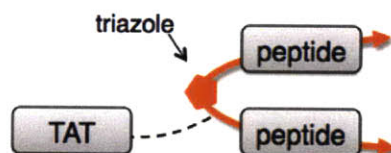


**Figure 6.5.** Schematic of the importance of linker length for the function of dimeric probes.

Previous studies have shown that dimeric ligands based on the Stargazin C-terminal sequence, composed of 15-residue monomeric units, were most effective at binding to the PSD-95 tandem PDZ domains in rat neurons.<sup>5</sup> With the hypothesis that the homologous TARPs, Stargazin and STG-2, share similarities in their cognate protein domains and binding modes, monomers of 15 residues were used to construct STG-2 dimers for the *C. elegans* target (see Figure 6.2). The STG-2 dimer was modified from the *C. elegans* STG-2 sequence, as shown in Figure 6.2; a tyrosine residue (replacing a phenylalanine residue) allows for quantification of the peptide by UV absorption analysis, and the *NRR* (Asn-Arg-Arg) sequence appended to the N-terminus aids in the solubility of the fairly hydrophobic peptide.

The application of these probes to studies *in vivo* necessitates a method of delivering the exogenous ligands into cells. Therefore, the probes were designed to incorporate a cell penetrating peptide sequence. Specifically, the HIV-1 TAT sequence (TAT<sub>47-57</sub>, YGRKKRRQRRR) was utilized for this purpose,<sup>7, 8</sup> since it has been shown to confer cell permeability to covalently-attached cargo. This sequence could be incorporated at the N-terminus of one of the dimer “arms,” far enough from the critical PDZ ligand – PDZ domain interactions to avoid interfering with binding of the dimer. A schematic of this general design

is shown in Figure 6.6, and details of the synthetic routes pursued and the optimized dimer synthesis will be presented in section 6-2. In general, the monomeric units are linked at their N-terminal residues to yield a dimer with two free carboxylic acids at the C-terminal amino acids; click chemistry of the alkyne of one monomer with the azide of another yields a triazole linker. The ability to further extend one of the monomeric units of the dimer (through reaction with a free N-terminus after click chemistry) is critical to the design of the probe, since this allows for covalent linkage of the TAT sequence to the dimer.



**Figure 6.6.** Modular TAT-modified dimeric ligand design. Note the presence of the triazole in the linker, which results from the click chemistry ligation of the monomeric units.

A control dimeric ligand of similar design was also pursued. The TAT-control dimer would be used for *in vivo* analyses to control for any non-specific or toxic effects of the TAT sequence. The peptide sequence designed for the control is shown in Figure 6.7, aligned with the corresponding STG-2 residues. Multiple design features of the control sequence should exclude binding of the control to PDZ domains. Firstly, the C-terminal amide (in place of the C-terminal carboxylate) should greatly reduce binding to any PDZ domain, since the C-terminal carboxylate of PDZ domain ligands forms critical contacts within the PDZ domain binding site. The arginine residue at the conserved 0 position replaces a residue that is canonically hydrophobic in all classes of PDZ domain ligands, and should not be tolerated in PDZ domain binding sites. Together with the multiple arginine residues at other positions near the C-terminus, these elements should render the probe unable to bind to any PDZ domain.

	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
<b>STG-2 Dimer</b>	N	R	R	Y	S	N	P	S	L	L	F	Nle	Q	D	V COOH
<b>Control Dimer</b>	Y	S	L	H	A	N	T	A	N	R	R	T	R	P	R NH <sub>2</sub>

**Figure 6.7.** Modified STG-2 sequence used for monomeric units of the TAT-STG-2 dimer, and control sequence used for monomeric units of TAT-control dimer.

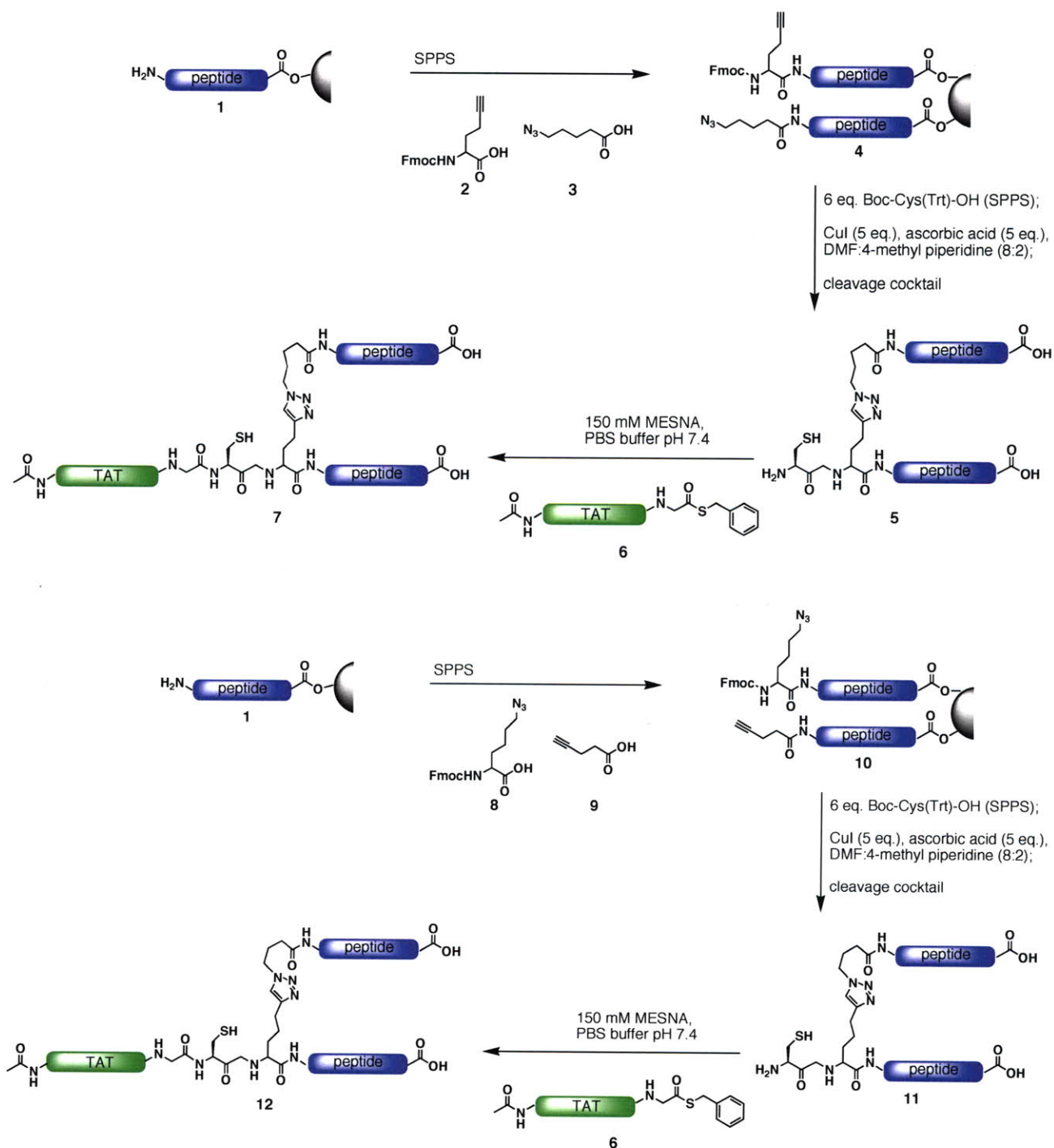
## ***6-2. Synthesis of TAT-STG-2 dimer and TAT-control dimer***

Previous unpublished results from the Imperiali laboratory have shown that an on-resin [3+2] dipolar Cu(I)-catalyzed Huisgen cyclization (“click” chemistry)<sup>9</sup> approach to forming triazole-linked dimeric PDZ domain ligands is favorable relative to click chemistry in solution. This is mainly due to solubility limitations of the monomers and the resulting dimer, which are often fairly hydrophobic (see STG-2 dimer entry, Figure 6.7), as well as the higher effective concentration of alkyne- and azide-modified peptides that are realized in an on-resin approach. Therefore, an on-resin synthetic approach was explored.

Two possible synthetic routes are presented in Schemes 6.1 and 6.2, which differ in the mode of TAT sequence introduction onto the dimeric peptide. In route A, the TAT sequence is incorporated via native chemical ligation, after the dimers have been synthesized and purified. This route has been utilized previously,<sup>5</sup> and can be advantageous since the dimer and TAT-thioester are purified before introduction of the TAT sequence, hopefully yielding a more pure TAT-dimer product before the final HPLC purification. In route B, the TAT sequence is incorporated via standard solid phase peptide synthesis before click chemistry. This route has fewer synthetic steps overall, and can be advantageous when the dimeric peptides alone are hydrophobic and not very soluble.

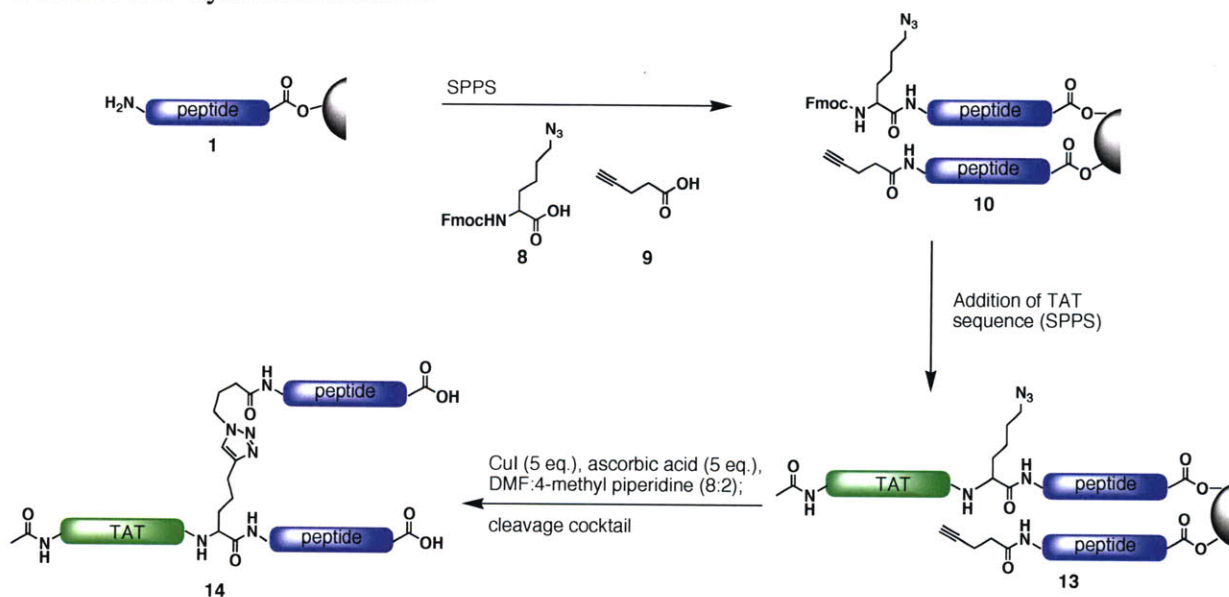


**Scheme 6.1.** Synthetic Route A, two variations.



**Scheme 6.1.** Synthetic Route A, two variations. Native chemical ligation after dimerization to introduce the TAT sequence, using either Fmoc-homopropargyl glycine **2** and 5-azidopentanoic acid **3** as the alkyne and azide (top) or Fmoc-Lys(N<sub>3</sub>) **8** and 4-pentynoic acid **9** (bottom). Note: SPPS = solid phase peptide synthesis, reagents and conditions are described in the *Experimental Methods* section.

**Scheme 6.2.** Synthetic Route B.



**Scheme 6.2.** Synthetic Route B. Standard SPPS before dimerization to introduce the TAT sequence. Note: SPPS = solid phase peptide synthesis, reagents and conditions used are described in the *Experimental Methods* section.

The options in synthetic route A were examined first, since the intermediates **5** and **11** are HPLC purified before reaction with purified TAT thioester, which should aid in the final TAT-dimer purification. Homopropargyl glycine **2** and Fmoc-Lys(N<sub>3</sub>) **8** were synthesized according to published literature procedures,<sup>10, 11</sup> until Fmoc-Lys(N<sub>3</sub>) became commercially available during the time period of our investigations. Initial testing revealed that the click chemistry reaction between **8** and **9** was higher yielding than the reaction between **2** and **3** in the context of test peptides. However, the STG-2 dimeric ligand **11** was too hydrophobic to dissolve in aqueous media, even with up to 10% dimethylformamide. The intermediate HPLC purification of the STG-2 dimeric ligand **11** was therefore not possible, and neither was the subsequent native chemical ligation reaction to yield compound **12**. This led to the exploration of synthetic route B.

Synthetic route B (Scheme 6.2) was successfully implemented for the formation of TAT-STG-2 and TAT-control dimers, and the full details of both dimeric ligand syntheses are described in the *Experimental Methods* section of this chapter. Note that for the synthesis of the control dimeric peptide, PAL-PEG-PS resin is used to afford C-terminal amides. Appendage of the basic TAT sequence via solid phase peptide synthesis (SPPS) before click chemistry avoided the STG-2 dimer solubility issues. For the synthesis of the TAT-control

dimer, optimizations to the SPPS conditions were necessary, as the major product initially observed was a truncation product (residue incorporation halted after coupling of the third TAT amino acid). At that point, 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) and 1-hydroxy-7-azabenzotriazole (HOAt) were explored as peptide coupling reagents in place of O-Benzotriazole-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and N-Hydroxybenzotriazole (HOBt). The HATU aza derivatives have been shown to be more efficient than HBTU, especially in syntheses involving hindered amino acids. Indeed, triple coupling of the TAT sequence residues with HATU/HOAt yielded the completed target sequence.

Since the azide, **8**, and the alkyne, **9**, are coupled onto the resin simultaneously in this synthesis, optimization of the ratio of each of these precursors to the resin was necessary. In initial trials, 4:3 equivalents of **8**:**9** were used, with the expectation that the more bulky reagent, **8**, might not couple quite as efficiently as precursor **9**. While this was indeed true, under these conditions very little of the coupling product with azide **8** was available for extension with the TAT sequence. In order for the dimerization to proceed with a reasonable yield, the TAT-azide-monomer and alkyne-monomer needed to be present in fairly equal quantities. Multiple different ratios of these compounds in the coupling step were therefore tested. Comparisons were made after the click chemistry step by preparative HPLC analysis of resulting TAT-dimer and excess unreacted alkyne-monomer. The results from these optimization studies are summarized in Table 6.1. Ultimately, the ratio of 3:1 (azide **8**: alkyne **9**) led to both a low amount of excess alkyne-monomer and the highest yield of TAT-dimer.

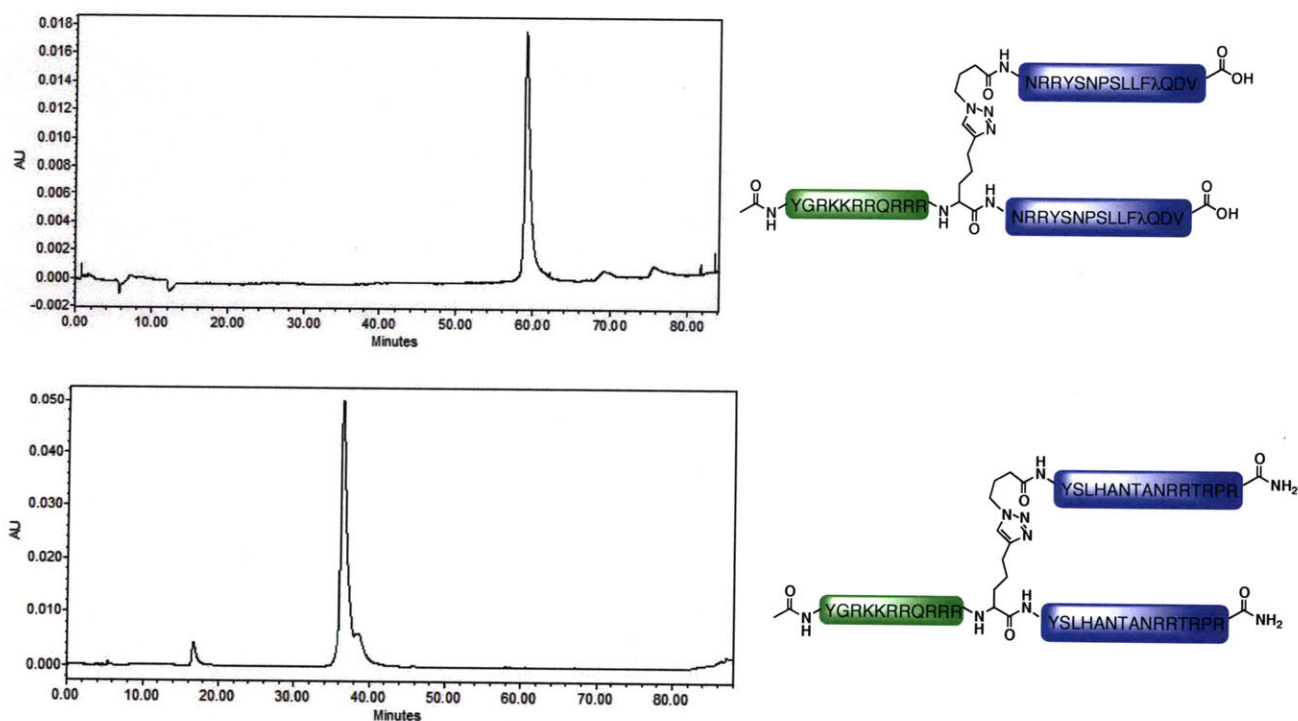
Entry	Ratio of reagents 8:9:resin <sup>a</sup>	Results <sup>b</sup>
1	4:3:1	monomer <sub>9</sub> >> TAT-dimer
2	6:2:1	monomer <sub>9</sub> > TAT-dimer
3	3:1:1	monomer <sub>9</sub> < TAT-dimer
4	2:1:1	monomer <sub>9</sub> < TAT-dimer
5	2:0.75:1	monomer <sub>9</sub> < TAT-dimer

**Table 6.1.** Optimization of the ratio of **8**:**9** used in the coupling step of the dimer synthesis. The optimal ratio was found to be 3:1, as highlighted in the table. <sup>a</sup> Ratio of equivalents of **8** and **9** with respect to the resin. <sup>b</sup> monomer<sub>9</sub> = alkyne monomer unit; comparative analysis of excess monomer<sub>9</sub> and TAT-dimer resulting from click chemistry.



During initial HPLC analysis, it was observed that the TAT-azide-monomer and TAT-dimer eluted under similar conditions. The following HPLC conditions were found to be optimal for the separation of TAT-dimer from excess monomer. For TAT-STG-2 dimer, the gradient was 5-25% acetonitrile containing 0.1% TFA for 6 min followed by 25-45% acetonitrile containing 0.1% TFA over 70 min in water containing 0.1% TFA at a flow rate of 10 mL/min. For TAT-control dimer, the gradient was 5-10% acetonitrile containing 0.1% TFA for 6 min followed by 10-25% acetonitrile containing 0.1% TFA over 70 min in water containing 0.1% TFA at a flow rate of 10 mL/min. Typically, two to three rounds of purification were necessary.

After multiple rounds of purification, ESI-MS and MALDI-TOF analyses show significantly reduced monomer signals as compared to the dimer signal. Analytical HPLC allowed for resolution of the monomer and dimer signals, hence final TAT-STG-2 and TAT-control dimeric ligand purity (>95% dimer) could be confirmed by analytical HPLC (see Figure 6.8 below).



**Figure 6.8.** Dimer HPLC traces. Top: TAT-STG-2 dimer, >95% pure by 280 nm analytical HPLC trace;  $\lambda$  = norleucine. Bottom: TAT-control dimer, >95% pure by 280 nm analytical HPLC trace.

Overall, the desired dimeric peptides were successfully synthesized and purified, and their characterization is presented in Table 6.2. While only small amounts of very pure material were produced, it was estimated that 0.15  $\mu\text{mol}$  of each dimeric ligand would allow for multiple tests in over 200 worms; 0.16  $\mu\text{mol}$  of the TAT-STG-2 dimer and 0.22  $\mu\text{mol}$  of the TAT-control dimer were obtained. These compounds were lyophilized into 0.02  $\mu\text{mol}$  aliquots and sent to the Maricq laboratory for testing in *C. elegans*, which will be described in section 6.3.

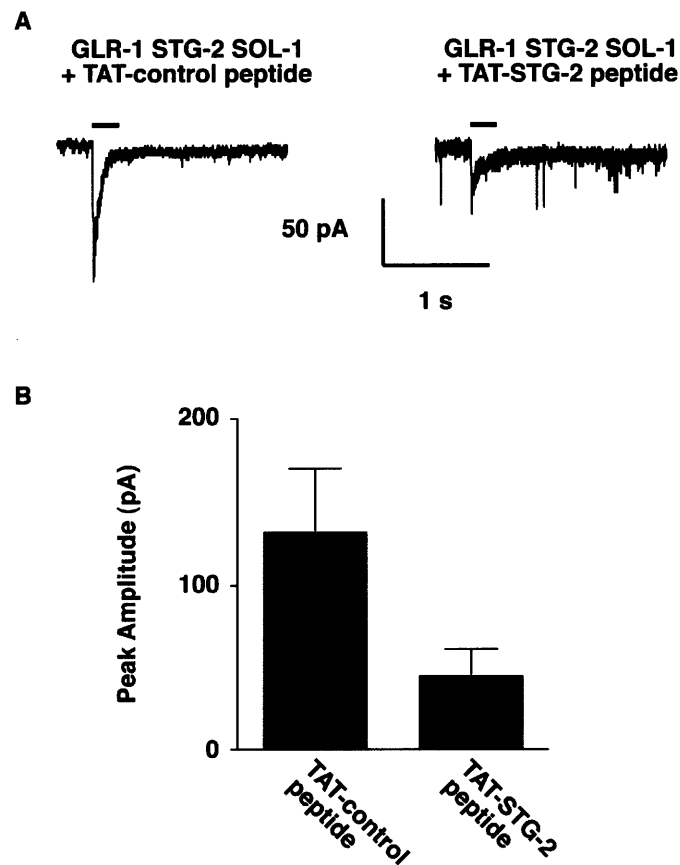
Entry	Peptide	Molecular Formula	Mass expected	ESI Mass obtained [MH] <sup>+</sup>	MALDI Mass obtained [MH] <sup>+</sup>	t <sub>R</sub> <sup>a</sup>
1	TAT-STG-2	C <sub>239</sub> H <sub>389</sub> N <sub>84</sub> O <sub>64</sub>	5462.4	1093.6 <sup>5+</sup> , 911.4 <sup>4+</sup>	5461.6	59.5
2	TAT-control	C <sub>227</sub> H <sub>383</sub> N <sub>98</sub> O <sub>60</sub>	5444.3	1089.7 <sup>5+</sup> , 908.7 <sup>4+</sup> , 778.7 <sup>3+</sup> , 681.5 <sup>2+</sup>	5437.9	36.8

**Table 6.2.** Characterization of dimers synthesized. <sup>a</sup> Retention time from analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using the following gradient: 5-10% acetonitrile containing 0.1% TFA for 6 min followed by 10-60% acetonitrile containing 0.1% TFA over 70 min in water containing 0.1% TFA at a flow rate of 0.9 mL/min.

### 6-3. Testing of probes *in vivo* in *C. elegans*

In collaboration with the Maricq laboratory, the STG-2 dimeric peptide was tested for its ability to disrupt GLR-1 mediated current in *C. elegans*. The complex of GLR-1, STG-2, and SOL-1 was reconstituted in *C. elegans* muscle, and the effect of the STG-2 dimeric peptide on current inhibition was tested. Proper function of the GLR-1 glutamate receptor is dependent upon its interaction with STG-2 and SOL-1, and disruption of these auxiliary subunits of the receptor complex has been shown to inhibit glutamate-gated currents.<sup>4</sup> Therefore, if the STG-2 C-terminal sequence is part of a critical interaction within the GLR-1 complex, the STG-2 dimeric inhibitor should disrupt this interaction and the GLR-1 synaptic currents.

Electrophysiological measurements were gathered after delivery of the TAT-STG-2 or TAT-control dimeric peptide to the *C. elegans* muscle via patch-clamp techniques, as previously described.<sup>12</sup> The peptide was diluted to a final concentration of 25  $\mu\text{M}$  in internal cellular fluid and patched into the muscle; glutamate was applied and currents were recorded beginning 1-2 minutes after introduction of the peptide. Initial results from multiple recordings show that the current response of the GLR-1 receptors to glutamate is indeed reduced after application of the TAT-STG-2 dimeric peptide (in comparison to the current response after application of TAT-control dimeric peptide). Figure 6.9 shows a sample of the currents from voltage-clamped body wall muscle cells after application of either TAT-STG-2 or TAT-control dimeric peptide, and also the mean peak current responses from multiple trials.



**Figure 6.9.** Current responses to glutamate, from voltage-clamped body wall muscle cells. A) Current responses of worms treated with (left) TAT-control dimeric peptide and (right) TAT-STG-2 dimeric peptide. Duration of glutamate application is signified by the horizontal bar above each current record. B) Mean peak current responses to glutamate application ( $\pm$  standard error of the mean);  $n = 4$ .

The results garnered from these experiments are promising, suggesting that the C-terminus of STG-2 is indeed critical to the function of the GLR-1 receptor complex. However, while this experimental approach allows for clean electrophysiological measurements due to the high copy numbers of the reconstituted receptor complex in the *C. elegans* muscle, the concentrations of protein in this system are much higher than they are in the native system. Additionally, there may be other scaffolding proteins that play critical roles in the GLR-1 receptor complex function, which are not present in the *C. elegans* muscle. Therefore, future investigations should focus on the testing of the dimeric peptides in the native *C. elegans* system (with endogenous levels of all proteins present), most likely with specific focus on the well-characterized AVA neurons. This pair of neurons is required for backward movement in the worm, and therefore behavioral phenotypic analysis could also be coupled with electrophysiological investigations.

## Conclusions

The on-resin syntheses of *C. elegans* TAT-STG-2 and TAT-control dimeric probes have been developed and optimized, as reported here. The dimeric probes have been successfully utilized in initial investigations in *C. elegans*, specifically with the aim of understanding the protein components of the GLR-1 receptor complex, their interactions with each other and with other proteins, and the functional significance of these interactions. Electrophysiological investigations with the STG-2 dimeric peptide have begun to shed light on the critical components governing GLR-1 receptor complex function. Specifically, the current response of the GLR-1 receptor to glutamate is significantly reduced in the presence of the TAT-STG-2 dimeric inhibitory peptide. While these results are a first step toward better understanding the role of STG-2 in the GLR-1 protein complex, the binding partner of the STG-2 C-terminus (and the TAT-STG-2 dimeric peptide) remains unidentified. The targeted interaction most likely involves a PDZ domain, and future work should focus on the identification of this protein as well as the role it plays in the function of the GLR-1 complex as a whole.

## Experimental Methods

**Optimal Synthesis of TAT-STG-2 Precursor Monomers.** Syntheses were performed on a 0.04 to 0.02 mmol scale with Fmoc-Val-NovaSyn® TGT resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents except where indicated differently (6 equivalents each), and DIPEA (12 equivalents) in DMF or NMP. Fmoc group removal was performed with a 20% solution of 4-methyl-piperidine in DMF (vol/vol) for 3 x 5 minutes. Capping with an acetyl group was achieved by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

Ac-TAT-Lys(N<sub>3</sub>)-STG-2 and alkyne-STG-2 monomers were synthesized on one batch of resin using an automated synthesizer (Applied Biosystems Inc. automated synthesizer model 431A) by using “difficult chemistry” Fmoc-based solid phase peptide synthesis procedures (double coupling, 100 minutes each, and 4 x 5 minutes Fmoc-deprotection before each coupling step). A mixture of Fmoc-azidolysine (Anaspec) and 4-pentynoic acid was coupled onto the resin after the STG-2 peptide residues were added, at an optimal ratio of 3:1 (azide:alkyne) equivalents as compared to the resin. This yielded a mixture of azide-containing and alkyne-containing monomers, with only the azide-containing monomers available for further Fmoc deprotection and extension. Further “difficult chemistry” synthesis allowed for the covalent addition of the TAT peptide residues at the N-terminus of the azide-containing monomer, which was acetyl capped at the end of the synthesis.

**Optimal Synthesis of TAT-Control Precursor Monomers.** Syntheses were performed on a 0.04 to 0.02 mmol scale with Fmoc-PAL-PEG-PS resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents except where indicated differently (6 equivalents each), and DIPEA (12 equivalents) in DMF or NMP. Fmoc group removal was performed with a 20% solution of 4-methyl-piperidine in DMF (vol/vol) for 3 x 5 minutes. Capping with an acetyl group was achieved by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

The 15-mer control sequence of the monomers was synthesized using an automated synthesizer (Applied Biosystems Inc. automated synthesizer model 431A) by using “difficult chemistry” Fmoc-based solid phase peptide synthesis procedures (double coupling, 100 minutes each, and 4 x 5 minutes Fmoc-deprotection before each coupling step). A mixture of

Fmoc-azidolysine (Anaspec) and 4-pentynoic acid was coupled onto the resin after the control sequence residues were added, at an optimal ratio of 3:1 (azide:alkyne) equivalents as compared to the resin. This yielded a mixture of azide-containing and alkyne-containing monomers, with only the azide-containing monomers available for further Fmoc deprotection and extension.

Addition of the TAT sequence to the azide-control monomer using the method described for the STG-2 peptide synthesis did not provide the desired compound, but instead a truncation product. Therefore, the TAT sequence was added via triple coupling steps with HATU/HOAt coupling reagents (6 equivalents each) on an automated synthesizer (Advanced ChemTech automated synthesizer model 396  $\Omega$ ), and this monomer was acetyl capped.

***Click Chemistry Ligation and Purification.*** The Ac-TAT-azide- and alkyne-containing monomers (for both the TAT-STG-2 and the TAT-control dimers) were then ligated via Cu<sup>I</sup>-catalyzed 1,3-dipolar cycloaddition (click chemistry)<sup>9</sup> on resin. Ligation was performed in the dark overnight by addition of CuI (55 mM, 5 equiv.) and ascorbic acid (5 equiv.) in DMF/4-methyl-piperidine (8:2, dried by bubbling with Argon) to the resin. After overnight shaking, the resin was dried. Drying of the resin and exposure to air resulted in a bluish tint to the resin, which was attributed to excess copper coordinated to the resin. Therefore, it was washed with 5 x 5 mL sodium diethyldithiocarbamate trihydrate (Et<sub>2</sub>NCSSNa•3H<sub>2</sub>O, 1% w/v), containing 1% DIPEA (v/v) in DMF to remove excess coordinated copper.<sup>13</sup> This was repeated 3-5 times until the resin was no longer blue or dark brown, and then the resin was washed extensively with DMF.

Peptides were deprotected and cleaved from the resin with a TFA/H<sub>2</sub>O/TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours. TFA was evaporated until the peptides precipitated. The peptides were then triturated and precipitated in cold ether before purification by reverse phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250x20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 280 nm. Purification was performed on 0.02 mmol or less of material at once, due to solubility and difficulty of HPLC separation of dimer from monomeric species. Typically, three rounds of HPLC purification were necessary to achieve pure (>95%) dimer. Peptides were stored lyophilized at -80 °C until use.

***Peptide Characterization and Quantification.*** Peptide identity was confirmed by MALDI-TOF mass spectroscopy (MALDI-TOF, PerSeptive Biosystems Voyager) using DHB as a matrix, and by ESI-MS analysis. Purity was assessed by analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5-10% acetonitrile containing 0.1% TFA over 6 minutes followed by 10-60% acetonitrile containing 0.1% TFA over 71 minutes in water containing 0.1% TFA at a flow rate of 0.9 mL/min). The final peptides were quantified using molar extinction coefficients of the peptides (each peptide contained at least one tyrosine residue, in order to allow for this quantification).

***Patching into Reconstituted C. elegans Muscle Cells, and Electrophysiology.*** All recordings were obtained from the body wall muscles of adult worms. For the worm preparation, histoacryl glue was applied to the dorsal aspect of worms on glass coverslips. A lateral incision was made along the length of the body, the flap of cuticle was glued down, and the intestine and gonad were removed. The preparation was then washed briefly (about 30 seconds) with a solution of 1 mg/mL collagenase (Sigma Type IV) in extracellular fluid. The extracellular fluid consisted of 150 mM NaCl, 5 mM KCl, 4 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 15 mM HEPES, and 10 M glucose (pH 7.4, osmolarity adjusted with 20 mM sucrose). A HEKA EPC-9 patch-clamp amplifier was used for whole-cell voltage-clamp recordings from muscle cells. Recording pipettes were fire polished to a resistance of 3-6 MΩ and filled with an intracellular fluid solution consisting of 115 mM K-gluconate, 25 mM KCl, 0.1 mM CaCl<sub>2</sub>, 50 mM HEPES, 5 mM Mg-ATP, 0.5 mM Na-GTP, 0.5 mM cGMP, 0.5 mM cAMP, and 1 mM BAPTA (pH 7.4, osmolarity adjusted with 10 mM sucrose).

Lyophilized peptide was diluted to a final concentration of 2 mM in diH<sub>2</sub>O, and this was diluted to a final concentration of 25 μM in intracellular fluid in the micropipette. About 1-2 minutes after breaking into the cells with the micropipette, 3 mM glutamate was applied via pressure application, and voltage-clamp recordings were made.



## Acknowledgements

I would like to thank Dr. Matthieu Sainlos, who came up with the initial design and synthesis of dimeric probes for tandem PDZ domains, which I then adapted and optimized for the *C. elegans* probes. I would especially like to thank Fred Hörndli (and his advisor Villu Maricq) for helpful discussions regarding the *C. elegans* system, as well as Fred Hörndli and Mike Jensen for spearheading the testing of these probes *in vivo*.

## References

1. Walker, C. S.; Brockie, P. J.; Madsen, D. M.; Francis, M. M.; Zheng, Y.; Koduri, S.; Mellem, J. E.; Strutz-Seebohm, N.; Maricq, A. V. "Reconstitution of invertebrate glutamate receptor function depends on stargazin-like proteins." *Proc. Natl. Acad. Sci. U S A* **2006**, *103*, 10781-10786.
2. Zheng, Y.; Brockie, P. J.; Mellem, J. E.; Madsen, D. M.; Walker, C. S.; Francis, M. M.; Maricq, A. V. "SOL-1 is an auxiliary subunit that modulates the gating of GLR-1 glutamate receptors in *Caenorhabditis elegans*." *Proc. Natl. Acad. Sci. U S A* **2006**, *103*, 1100-1105.
3. Zheng, Y.; Mellem, J. E.; Brockie, P. J.; Madsen, D. M.; Maricq, A. V. "SOL-1 is a CUB-domain protein required for GLR-1 glutamate receptor function in *C. elegans*." *Nature* **2004**, *427*, 451-457.
4. Wang, R.; Walker, C. S.; Brockie, P. J.; Francis, M. M.; Mellem, J. E.; Madsen, D. M.; Maricq, A. V. "TARP proteins have fundamental roles in the gating of glutamate receptors and the tuning of synaptic function." *Neuron* **2008**, *25*, 997-1008.
5. Sainlos, M.; Tigaret, C.; Poujol, C.; Olivier, N. B.; Bard, L.; Breillat, C.; Thiolon, K.; Choquet, D.; Imperiali, B. "Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization." *Nat. Chem. Biol.* **2011**, *in press*.
6. Feng, W.; Zhang, M. "Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density." *Nat. Rev. Neurosci.* **2009**, *10*, 87-99.
7. Fawell, S.; Seery, J.; Daikh, Y.; Moore, C.; Chen, L. L.; Pepinsky, B.; Barsoum, J. "Tat-mediated delivery of heterologous proteins into cells." *Proc. Natl. Acad. Sci.* **1994**, *91*, 664-668.
8. Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F. "In vivo protein transduction: delivery of a biologically active protein into the mouse." *Science* **1999**, *285*, 1569-1572.
9. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. "A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes." *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596-2599.
10. Dong, S.; Merkel, L.; Moroder, L.; Budisa, N. "Convenient syntheses of homopropargylglycine." *J. Pept. Sci.* **2008**, *14*, 1148-1150.
11. Weerapana, E.; Speers, A. E.; Cravatt, B. F. "Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes." *Nat. Protoc.* **2007**, *2*, 1414-1425.

12. Francis, M. M.; Evans, S. P.; Jensen, M.; Madsen, D. M.; Mancuso, J.; Norman, K. R.; Maricq, A. V. "The Ror receptor tyrosine kinase CAM-1 is required for ACR-16-mediated synaptic transmission at the *C. elegans* neuromuscular junction." *Neuron* **2005**, *46*, 581-594.
13. Weterings, J. J.; Khan, S.; van der Heden van Noort, G. J.; Melief, C. J.; Overkleeft, H. S.; van der Burg, S. H.; Ossendorp, F.; van der Marel, G. A.; Filippov, D. V. "2-Azidoalkoxy-7-hydro-8-oxoadenine derivatives as TLR7 agonists inducing dendritic cell maturation." *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2249-2251.

# Appendix 1

## Development of a Sox-Based Sensor for Calmodulin Kinase II (CaMKII) Activity

### Introduction

Calmodulin kinase II (CaMKII), a Ser/Thr protein kinase, comprises about 2% of the total protein in the hippocampus of rodents, and is thought to play crucial roles in long-term potentiation (LTP) and long-term depression (LTD), modulations in synaptic strength that govern learning and memory.<sup>1</sup> The current activity probe most widely used is an 88-residue sequence (vimentin, which is specifically phosphorylated by CaMKII) tagged with a GFP fluorophore.<sup>2</sup> Ratiometric analysis of phosphovimentin antibody staining and GFP fluorescence is needed to quantify kinase activity from the application of this probe. Work has been done toward the development of a more straightforward sensor that could eventually be used in neurons to detect CaMKII activity, and will be described here.

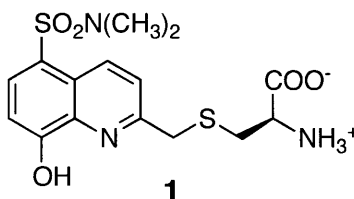
The sulfonamido-oxine (Sox) chromophore previously developed in the Imperiali laboratory has been successfully placed in kinase recognition motifs containing a Ser or Thr residue; upon phosphorylation of optimized sensors, the affinity of the chemosensor for Mg(II) increases, and chelation-enhanced fluorescence is observed.<sup>3,4</sup> To develop sensors for CaMKII based on this design, CaMKII recognition motifs from both “syntide” and NR2B were utilized, since these are two of the highest-affinity binding partners of CaMKII present in the PSD.<sup>1</sup> Short peptides of each of these recognition motifs were previously characterized via kinetic analysis as suitable substrates for CaMKII.<sup>1</sup>

### Results and Discussion

#### *A1-1. Sensor Development and Design*

Sensors based on the NR2B and syntide sequences were synthesized, with introduction of the Sox chromophore as the amino acid C-Sox 1 (see Figure A1.1), at the positions shown in Table A1.1.<sup>4</sup> The N- and C-termini of the peptides given in the literature were truncated, in order to make the sensors both easier to prepare and handle and also

hopefully less promiscuous with respect to recognition by off-target kinases. Both the Ser and phosphoserine (pSer) analogs of the peptides were synthesized, and the robust fluorescence increases between the phosphorylated and non-phosphorylated peptides are displayed in Table A1.1.



**Figure A1.1.** Sox-derived amino acid building block for peptide synthesis.

#	Origin	Peptide Sequence <sup>a</sup>	Formula	Mass expected	Mass <sup>b</sup> obtained [MH] <sup>+</sup>	RFI <sup>c</sup>
1	Syntide	AC-ARTLSV $\phi$ GLPGK-NR2	C <sub>63</sub> H <sub>104</sub> N <sub>18</sub> O <sub>18</sub> S <sub>2</sub>	1507.3	1507.2	-
2	Syntide	AC-ARTLpSV $\phi$ GLPGK-NR2	C <sub>63</sub> H <sub>104</sub> N <sub>18</sub> O <sub>21</sub> P <sub>1</sub> S <sub>2</sub>	1586.3	1586.2	5.4 (4.9)
3	NR2B	AC-KLRRQHSY $\phi$ TFV-NR2	C <sub>82</sub> H <sub>122</sub> N <sub>24</sub> O <sub>21</sub> S <sub>2</sub>	1843.7	1845.0	-
4	NR2B	AC-KLRRQHpSY $\phi$ TFV-NR2	C <sub>82</sub> H <sub>123</sub> N <sub>24</sub> O <sub>24</sub> P <sub>1</sub> S <sub>2</sub>	1922.7	1923.4	5.3 (4.5)

**Table A1.1.** Sensor sequences and their fluorescence increases at 485 nm. <sup>a</sup>  $\phi$  = C-Sox residue. <sup>b</sup> Masses were determined by MALDI-TOF spectroscopy. <sup>c</sup> RFI = Relative fluorescence increase of pSer sensor over Ser sensor at 485 nm, values in parentheses are the RFI in the presence of ATP.

The fluorescence increases of the Sox chromophore between phosphorylated and non-phosphorylated substrate sequences were promising, since they were in the range typically found for successful Sox-based kinases sensors. These fluorescence increases were also robust in the presence of ATP, which is present both in kinase assay conditions as well as the cellular environment. Therefore the sensors were both validated with recombinant CaMKII kinase and tested for kinetic competency.

## ***A1-2. Sensor Validation***

In order to test the initial sensors with CaMKII, it was first necessary to test kinase competency. CaMKII (truncated monomer of the  $\alpha$  subunit, Rat, Recombinant, *S. frugiperda* Sf9, NEB) was used, and multiple dilutions were tested. Table A1.2 below shows the competency tests (with 0.5  $\mu$ M substrate **3**) for the chosen dilution conditions (10-fold dilution of kinase prior to activation, followed by 10-fold dilution of kinase following activation). The activation procedure for the kinase is described in detail in the *Experimental Methods* section.

<b>Entry</b>	<b>ng CaMKII</b>	<b>Slope</b>	<b>Slope ratio<sup>a</sup></b>
<b>1</b>	<b>1</b>	<b>2307 <math>\pm</math> 163</b>	<b>1.00</b>
<b>2</b>	<b>0.5</b>	<b>1128 <math>\pm</math> 172</b>	<b>0.49</b>
<b>3</b>	<b>0.25</b>	<b>607 <math>\pm</math> 152</b>	<b>0.26</b>

**Table A1.2.** Recombinant kinase competency tests. Each entry is the average of three trials.  
<sup>a</sup>Ratio of slope with that of Entry 1.

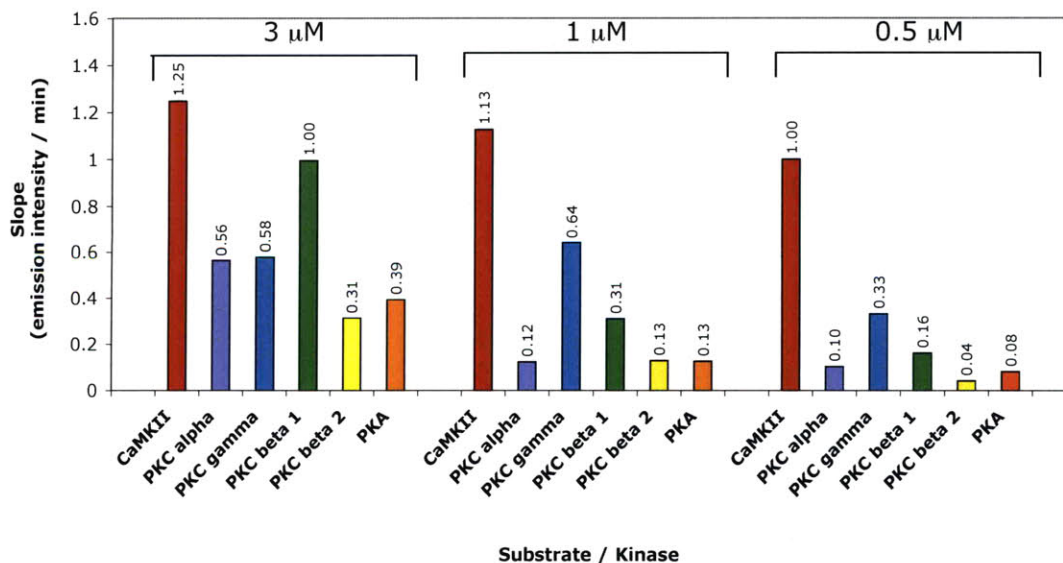
Sensors **1** and **3** were carried on in assays with the recombinant kinase, and the results of these assays are summarized in Table A1.3. Substrate **1** did not display favorable kinetic properties, with a  $K_m$  of greater than 70  $\mu$ M, much higher than the reported literature value for the parent syntide peptide. However, multiple assays of substrate **3**, the results of which are averaged and presented in entry 4 of Table A1.3, showed favorable kinetic properties comparable to those reported in the literature. Specifically, sensor **3** and the parent peptide reported in the literature displayed the same affinity ( $K_m$ ) for CaMKII, and the rates of catalytic efficiency ( $V_{max}$ ) were also comparable. Magnesium titrations yielded a dissociation constant of 7.17 mM  $\pm$  0.51 mM for sensor **3** and  $Mg^{2+}$ . This most likely indicates that supplementary  $Mg^{2+}$  would be necessary for this sensor to optimally detect CaMKII activity *in vivo*, since the typical concentration of this cation is about 1 mM in neurons. However, this is feasible, and due to the favorable kinetic properties observed, sensor **3** was further developed through analysis in cross-reactivity studies.

Entry	Origin	ng Kinase	$K_m$ ( $\mu\text{M}$ )	$V_{\text{max}}$ ( $k_{\text{cat}}$ ) ( $\mu\text{mol}/\text{min}/\text{mg}$ )	$k_{\text{cat}}/K_m$
1	Syntide	-	$13.4 \pm 1.8$	$2.0 \pm 0.25$	$15.4 \pm 4.1$
2	Syntide	10	71	1.9	0.03
3	NR2B	-	$0.29 \pm 0.1$	$0.62 \pm 0.16$	$2.25 \pm 0.29$
4	NR2B	0.9	$0.29 \pm 0.02$	$1.6 \pm 1.0$	$5.4 \pm 2.9$

**Table A1.3.** Kinetic assays of substrates. <sup>a</sup>Literature values for parent peptides are displayed in entries 1 and 3,<sup>1</sup> to be compared with entries 2 and 4, respectively.

### A1-3. Cross Reactivity with Other Kinases

Due to the favorable kinetic properties of sensor **3**, cross-reactivity studies with other kinases were performed to establish selectivity. Specifically, PKA and isoforms of PKC were chosen for the initial selectivity tests, as these kinases are expected to be prevalent in neurons.<sup>5</sup> Figure A1.2 shows the results of these studies, represented as reaction rates (slopes) normalized to that of  $0.5 \mu\text{M}$  sensor **3** + CaMKII. All of these studies were done using 1 ng kinase.



**Figure A1.2.** Normalized cross reactivity of PKC isoforms and PKA with sensor **3**.

At the lower 0.5  $\mu\text{M}$  concentration, near the  $K_m$  of sensor **3**, CaMKII is more than 3-fold selective for the sensor over PKA and the PKC isoforms. It is noteworthy that while PKC and PKA activities drop down as the sensor concentration is lowered, CaMKII activity basically remains steady. Introduction of the sensor at this low concentration in neurons should allow for low background phosphorylation by these PKC isoforms, provided that they are present at equal or lesser concentrations than CaMKII in the neurons. Unfortunately, the relative amount of these kinases present in neurons is unknown, and these values likely vary greatly depending on many biochemical processes and neuronal developmental stages. Additionally, there remain other off-target kinases to test against this sensor, including other isoforms of PKC as well as other kinases with preferences for basic substrates (ERK, AKT, and MK2).<sup>6,8</sup> However, thus far sensor **3** is promising for future application to neuronal studies.

## Conclusions

The design, synthesis, and *in vitro* analysis of an optimized Sox-based sensor for the detection of CaMKII activity in neurons has been performed. Further analysis has shown that the sensor is specific for CaMKII over multiple PKC isoforms, at concentrations close to the  $K_m$  of the sensor. Future work will involve additional cross-talk studies, as well as the testing of this probe in neuronal cells, where its utility is envisioned. Specifically, the sensor has been designed to be patched into neuronal cells, allowing for the imaging of CaMKII activity in these cells with spatial and temporal resolution.



## Experimental Methods

**Sensor synthesis and characterization.** All amino acids except Fmoc-C(Sox)-OH were added according to standard Fmoc SPPS procedures. Manual synthesis was performed on a 0.04 to 0.02 mmol scale with Fmoc-PAL-PEG-PS resin (0.2 mmol/g), using standard Fmoc-protected amino acid (6 equivalents), HBTU/HOBt as coupling reagents (6 equivalents each) and DIPEA (12 equivalents) in DMF or NMP. Coupling steps were conducted with a 50 mM solution of amino acid (~1 mL per 100 mg of resin) for 1 hour. Fmoc group removal was performed with a 20% solution of 4-methyl-piperidine in DMF (vol/vol) for 3 x 5 minutes. After removal of the N-terminal Fmoc group, the resulting free amine was capped with an acetyl group by using an acetic anhydride/pyridine solution (0.15 M each in DMF).

Fmoc-C(Sox)-OH was coupled overnight under the following conditions: amino acid / PyAOP / HOAt / DIPEA, 2.5:2.5:2.5:6.25, 0.15 M in DMF. Since the TNBS test for free amines can yield false positive results after coupling of Fmoc-C(Sox)-OH, test cleavage and MALDI-TOF-MS analysis after coupling was performed. Peptides were deprotected and cleaved from the resin with a TFA/H<sub>2</sub>O/TIPS (95:2.5:2.5) cleavage cocktail for 2.5 to 3 hours. TFA was evaporated until the peptides precipitated. The peptides were then triturated and precipitated in cold ether before purification by reverse phase HPLC on a semi-preparative column (YMC-Pack Pro C<sub>18</sub>, ODS-A 5/120, 250x20 mm) in water (0.1% TFA) using an acetonitrile (0.1% TFA) gradient and monitoring at 228 nm and 316 nm (Sox). Peptides were stored lyophilized at -80 °C until use.

**Peptide Characterization and Quantification.** Peptide identity was confirmed by MALDI-TOF mass spectroscopy (MALDI-TOF, PerSeptive Biosystems Voyager) using DHB as a matrix. Purity was assessed by analytical reverse phase HPLC (YMC C<sub>18</sub>, ODS-A 5/120, 250x4.6 mm) using a standard gradient (5% acetonitrile containing 0.1% TFA for 5 min followed by 5-95% acetonitrile containing 0.1% TFA over 35 min in water containing 0.1% TFA at a flow rate of 1 mL/min). The final peptides were quantified using the molar extinction coefficient of Sox ( $\epsilon_{355\text{nm}} = 8247 \text{ M}^{-1}\text{cm}^{-1}$  in water).

**Fluorescence Experiments.** Fluorescence scans were performed on a Fluoromax 3 from Jobin Yvon. When comparing fluorescence spectra, 3 nm excitation and emission slit widths were used. For enzyme assays, 5 nm emission and excitation slit widths were used. For all experiments, an excitation wavelength of 360 nm was used. Enzyme assays were performed by monitoring emission at 485 nm.

To determine the difference in fluorescence intensity of the Ser and corresponding pSer sensors at  $\lambda_{\text{max}}$ , these peptides were tested alone in buffer and in the presence 1.2  $\mu\text{L}$  100 mM ATP solution. The 10x stock buffer composition is as follows: 200 mM HEPES (pH 7.4) and 100 mM  $\text{MgCl}_2$ .

**Magnesium Titrations.** A 500  $\mu\text{L}$  solution of 1 or 10  $\mu\text{M}$  pSer NR2B sensor, 150 mM NaCl, and 20 mM pH 7.4 HEPES was made in a glass cuvette. Successive additions of  $\text{MgCl}_2$  were made, allowing for fluorescence emission spectra to be taken at intervals between 0.5  $\mu\text{M}$  and 300  $\mu\text{M}$   $\text{MgCl}_2$  concentrations. Data were fit using SPECFIT/32™ Global Analysis System for Windows (version 3.0.39).

**CaMKII Preparation.** According to the New England Biolabs (NEB) protocol and the literature, prior to substrate phosphorylation, CaMKII should be activated by autophosphorylation with ATP/ $\text{Mg}^{2+}$  in the presence of  $\text{CaCl}_2$  and calmodulin. Neither  $\text{CaCl}_2$  nor calmodulin are required for the subsequent phosphorylation of exogenous substrate. The CaMKII was activated according to the NEB protocol. Firstly, a 10x reaction buffer was made. The 1x reaction buffer composition is as follows: 50 mM TRIS-HCl, 10 mM  $\text{MgCl}_2$ , 2 mM dithiothreitol (DTT), 0.1 mM  $\text{Na}_2$  EDTA, pH 7.5 at 25 °C. Then 2  $\mu\text{L}$  CaMKII (truncated monomer of the  $\alpha$  subunit, Rat, Recombinant, *S. frugiperda* Sf9, NEB) (from a 2.5  $\mu\text{L}$  aliquot, stored at -80 °C) was mixed with 1x reaction buffer supplemented with 100  $\mu\text{M}$  ATP, 1.2  $\mu\text{M}$  calmodulin, and 2 mM  $\text{CaCl}_2$ , to yield the final recommended 10-fold dilution. After incubation for 10 minutes in a 30 °C water bath, the activated kinase was diluted an additional 10-fold before use.

***PKA Preparation.*** PKA does not need pre-activation. For the reactions, a 10x reaction buffer was made. The 1x buffer composition is the following: 20 mM HEPES (pH 7.4), 10 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mM EGTA. Then a 10x enzyme dilution buffer was made, in which the 1x buffer concentrations are: 50 mM Tris, 10 mM MgCl<sub>2</sub>, 1 mM DTT, and 0.15 mg/mL BSA. The enzyme dilution buffer was diluted to 1x, and used to dilute the enzyme solution 500-fold. Then 1 μL PKA (from an aliquot stored at -80 °C, after dilution) was used in the reaction.

***PKC Preparation.*** A 1 μL aliquot of PKCα (Human, Recombinant, Calbiochem) was diluted with enzyme dilution buffer to the final concentration of 1 ng / μL. A 10x dilution buffer was made, and the 1x buffer concentrations are: 20 mM HEPES (pH 7.4), 10 mM MgCl<sub>2</sub>, 0.3 mM CaCl<sub>2</sub>, 10 mg/mL BSA, 0.01% Brij-35 P. For the reaction, 120 μL of the following was made in a cuvette: 20 mM HEPES (pH 7.4), 10 mM MgCl<sub>2</sub>, 0.3 mM CaCl<sub>2</sub>, 0.1 mM EGTA, 100 μM ATP, 1 mM DTT, 0.5 μg/mL phosphatidylserine, 0.1 μg/mL diacylglycerol, 1 ng kinase. The reaction components were mixed and incubated for 4 min at 30 °C. Substrate was added last and mixed to begin the reaction.

***Assay Monitoring (Fluorimeter).*** Reaction solutions were mixed 5x in the cuvette before and after addition of the kinase, and were subsequently monitored for 10 minutes while fluorescence was recorded every second with a 1 sec integration time ( $\lambda_{ex} = 360$  nm,  $\lambda_{em} = 485$  nm, slit widths: Ex = 5 nm and Em = 5 nm). In order to correct for the intrinsic fluorescence of the substrate in the kinase reaction, separate samples of substrate and phosphopeptide were prepared (without enzyme) and the fluorescence was recorded under the same conditions. Note that, for CaMKII and PKA, kinase was then added to these same substrate samples for the reaction.

## Acknowledgements

I would like to thank Dr. Elvedin Lukovic for initializing me into the world of Sox-based sensors, and for his help as I learned the basics of Sox-based sensor design, synthesis, and assay procedures. Additionally, he provided me with the PKA and PKC aliquots, as well as multiple stock solutions.

## References

1. Praseeda, M.; Pradeep, K. K.; Krupa, A.; Krishna, S. S.; Leena, S.; Kumar, R. R.; Cheriyan, J.; Mayadevi, M.; Srinivasan, N.; Omkumar, R. V. "Influence of a mutation in the ATP-binding region of Ca<sup>2+</sup>/calmodulin-dependent protein kinase II on its interaction with peptide substrates." *Biochem J.* **2004**, *378*, 391-397.
2. Tsui, J.; Malenka, R. C. "Substrate localization creates specificity in calcium/calmodulin-dependent protein kinase II signaling at synapses." *J. Biol. Chem.* **2006**, *281*, 13794-13804.
3. Shults, M. D.; Carrico-Moniz, D.; Imperiali, B. "Optimal Sox-based fluorescent chemosensor design for serine/threonine protein kinases." *Anal. Biochem.* **2006**, *352*, 198-207.
4. Lukovic, E.; Gonzalez-Vera, J. A.; Imperiali, B. "Recognition-domain focused chemosensors: versatile and efficient reporters of protein kinase activity." *J. Am. Chem. Soc.* **2008**, *130*, 12821-12827.
5. Hu, H. J.; Glauner, K. S.; Gereau, R. W. t. "ERK integrates PKA and PKC signaling in superficial dorsal horn neurons. I. Modulation of A-type K<sup>+</sup> currents." *J. Neurophysiol.* **2003**, *90*, 1671-1679.
6. Gonzalez, F. A.; Raden, D. L.; Davis, R. J. "Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases." *J. Biol. Chem.* **1991**, *266*, 22159-22163.
7. Obata, T.; Yaffe, M. B.; Leparac, G. G.; Piro, E. T.; Maegawa, H.; Kashiwagi, A.; Kikkawa, R.; Cantley, L. C. "Peptide and protein library screening defines optimal substrate motifs for AKT/PKB." *J. Biol. Chem.* **2000**, *275*, 36108-36115.
8. Huang, C.; Jacobson, K.; Schaller, M. D. "MAP kinases and cell migration." *J. Cell Sci.* **2004**, *117*, 4619-4628.

