Novel Modulators of Rickets Receptor Mediated Signaling *In Vitro* and *In Vivo*

A thesis

submitted by

Benjamin N. Harwood

In partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Genetics

TUFTS UNIVERSITY

Sackler School of Graduate Biomedical Sciences

May 2014

ADVISER: Alan S. Kopin, MD.

Abstract

My initial efforts focused on extending the utility of membrane tethered ligands (MTLs) to enable exploration of the function of rickets (rk), a *Drosophila* melanogaster G protein-coupled receptor (GPCR). Rk mediates wing expansion, cuticle hardening, and melanization during the transition from the pupal to adult stages. The endogenous rk ligand is bursicon, a heterodimeric cystine-knot protein comprised of an α and a β subunit. As a probe of rk function, I developed a bursicon MTL by genetically fusing the α and β subunits. Subsequent experiments showed that the bursicon MTL activates rk in vitro. Tissue specific expression of the bursicon MTL using the Gal4/UAS system demonstrates that this construct can modulate rk signal transduction in vivo. Notably, in vivo expression of a bursicon MTL results in decreased viability and defects in wing expansion, phenotypes expected when bursicon/rk signaling is disrupted. Subsequent in vitro studies have shown that chronic exposure to a bursicon MTL results in rk desensitization. Expression of a bursicon MTL with a selection of tissue specific Gal4 drivers has shown that rk is critical for *Drosophila* development, specifically in muscle. Since our previous studies had shown an important role for bursicon/rk signaling during eclosion in conjunction with the fact that the pathway is conserved among insects; it became clear that rk is a putative target for a novel insecticide. To investigate this hypothesis, we conducted a high throughput screen for chemical modulators of rk; more than 350,000 compounds were assessed. The highest affinity ligands are currently being evaluated and further optimized. Finally, as another extension of MTL technology, we have developed a second type of membrane anchored ligand. These constructs, soluble

membrane anchored ligands (SMALs), are comprised of a peptide conjugated to a lipid. In vitro cell based studies have shown that SMALs recapitulate the desirable properties of MTLs (e.g. increased efficacy, prolonged activity). The combination of MTL and SMAL technologies creates a powerful system to develop and optimize membrane anchored ligands as molecular probes of GPCRs. In addition, SMALs hold promise as a new form of therapeutics.

Acknowledgements

First and foremost I would like to thank my thesis advisor and mentor Dr. Alan Kopin for allowing me to conduct my PhD research in his laboratory. He has been essential to my academic success. Alan has been incredibly supportive in allowing me to pursue my research interests. Throughout the course of my dissertation research, Alan has done an amazing job of creating a strong team of individuals and a laboratory environment which promotes science of the highest caliber. He has also been very generous with his time and efforts and my dissertation is better as a result.

Secondly, I would like to thank Dr. Martin Beinborn and Dr. Isabelle Draper for their role in my career development. Martin has always provided valuable scientific advice on pharmacology. Isabelle's role training me in *Drosophila* genetics has been crucial to my development as a well rounded scientist. Insights from both Martin and Isabelle have helped lead my research towards key areas that I have may have otherwise missed. Furthermore, their feedback on my dissertation and manuscripts is greatly appreciated.

Next, I would like to thank the various members of the Kopin laboratory that I have had the pleasure of working with throughout the years. Dr. Jean-Philippe (JP) Fortin helped tremendously in showing me many of the techniques used in the lab as well as keeping me on my toes. Much of the credit should go to JP for first demonstrating the incredible versatility of membrane tethered ligands. Many of the studies in this thesis rest upon the foundation that JP established. Dr. Jamie Doyle has also been a fantastic lab-mate throughout the years. I especially want to thank Jamie for her efforts with the

high throughput screening (HTS) projects. HTS is labor intensive, requires organization, excellent laboratory skills, and high performance standards; all attributes which Jamie possesses. Without her help, our HTS projects would not be where they are today. I also thank Jamie for her collaboration on the substance P and CCK4 work in this thesis, a second project where her efforts were integral. I would like to thank Ci Chen for her day to day help in the lab. Ci's work with cell culture is what makes all of the pharmacology we do possible. Ci also played an essential role in cloning the *Anopheles gambiae* rk ortholog. This was an extremely frustrating two year process which would not have been successful without Ci's resiliency. Finally, I thank a more recent member to the lab, Bina Julian. Bina has been a great addition to the lab dynamic and is more than capable of carrying the torch of membrane tethered ligands.

Last but not least, I would like to thank my friends and family who have been so supportive of me during my time in graduate school. I've made many great friends in school as well as in Boston/Cambridge during my time here. It has certainly become my home. My parents have always been incredibly supportive throughout my life including my decision to pursue my PhD. They have given me every opportunity to be successful, yet allowed me to follow my own trajectory. It has been a pleasure to live in a place where I can frequently visit them and can get away from the city whenever I need. I thank my brother Asch, who has also been on his own inspiring parallel trajectory of development. He is a fantastic big brother who always has words of encouragement when I need them. Finally, I thank my partner in life Kerin, who has lovingly been there for me for almost my entirety of graduate school. The task of supporting a person during

the ups and downs of graduate school is unenviable. Kerin has enriched my life so much and is always there when I need her most.

TABLE OF CONTENTS

GENERAL INTRODUCTION 2-27

CHAPTER 1 28-59

Membrane tethered bursicon constructs as heterodimeric modulators of the *Drosophila* GPCR rickets

CHAPTER 2 60-84

Membrane tethered bursicon as a tool to probe the role of the *Drosophila* GPCR rickets *in vivo*

CHAPTER 3 85-124

High throughput screen to identify novel modulators of the *Drosophila* GPCR rickets

CHAPTER 4 125-148

A systematic strategy expedites the conversion of low affinity peptides to high potency anchored ligands

GENERAL DISCUSSION 149-170

REFERENCES 171-179

LIST OF TABLES

GENERAL INTRODUCTION
Table i.1. ..... List of peptides converted into active MTLs and their cognate receptors. 25

CHAPTER 2
Table 2.1. ..... CFP-tBur- β-α screen with indicated Gal4 driver lines 76

LIST OF FIGURES

GENERAL INTRODUCTION

Figure i.1. .....Classification of the human GPCR superfamily. 7

Figure i.2. .... Figure i.2. Cartoon representing a GPCR signaling via either the $G\alpha_s$ or $G\alpha_i$ mediated pathways. 9

Figure i.3. Phylogenetic tree of select LGR subfamily members. 14

Figure i.4 Pie Chart showing the the percentage of FDA approved drugs by target class. 21

Figure i.5. ....Cartoon representing a GPCR and Type I TMD membrane tethered ligand at the cell surface 23

CHAPTER 1

Figure 1.1. ..... Schematic representation of heterodimeric membrane tethered bursicon subunits coexpressed with rk. 38

Figure 1.2. ..... Bursicon membrane tethered α and β subunits together activate the *Drosophila* rickets receptor..... 39

Figure 1.3. .....Rk is activated by co-expression of either two complementary soluble bursicon subunits or complementary combinations of soluble and tethered bursicon subunits. 42

Figure 1.4. ...Bursicon MTLs are active fusion proteins independent of C-terminal subunit positioning..... 43

Figure 1.5 .. Rk activation by a bursicon MTL fusion protein can be visually monitored by confocal microscopy. 45

Figure 1.6. ..... Development of a membrane tethered inhibitor of rk signaling. 47

Figure 1.7. ..... Nucleotide sequence encoding the TNFα type II transmembrane tethered ligand backbone. 53

Figure 1.8 .....Tethered bursicon subunits are expressed independent of complementary subunit. 54

Figure 1.9Active conditioned media cannot be made from expression of tethered bursicon constructs. 55

Figure 1.10 .....dLGR1and dLGR3 are not activated by soluble or tethered bursicon. 56
Figure 1.11 ..... Agonist of tBur-α-β is maintained despite serial C-terminal deletions 57

Figure 1.12. .....Soluble Chimeric Bursicon conditioned media activates rk. 58

Figure 1.13. .....Soluble Chimeric Bursicon deletion mutants are not potent rk antagonists. 59

CHAPTER 2

Figure 2.1. ....Ubiquitous overexpression of membrane tethered Bursicon recapitulates the phenotypes induced by downregulation of its cognate receptor, rk..... 67

Figure 2.2. ..... Chronic *in vitro* stimulation results in rk desensitization. 69

Figure 2.3. ..... Expression of membrane tethered bursicon in muscle and motor neurons results in wing expansion defects. 71

Figure 2.4. ..... CFP-tBur-β-α expression or rk knockdown in the mesoderm causes developmental lethality and wing expansion defects..... 73

Figure 2.5….. Rk knockdown and CFP-tBur-β-α expression in adult indirect flight muscle causes lethality and wing expansion defects…….80

Figure 2.6. Expression of membrane tethered bursicon with Rk^{PAN}-Gal4 results in reduced survival and in wing expansion defects. 84

CHAPTER 3

Figure 3.1 ..... Experimental protocol for 384-well plate natural product screen. 94

Figure 3.2. ..... Follow up analysis for 27 fractions. 95

Figure 3.3 ..... Final four fractions were retested on rk and a second counterscreen receptor MC4R..... 96

Figure 3.4. Subfractionation of fraction C3 derived from the dried aerial portion of *Pogostemon cablin*..... 99

Figure 3.5. ..... Second fractionation process also identified fractions 46 and 47 as differentially active between rk and MC4R. 100

Figure 3.6. .....Structure of Verbascoside, the most prominent component identified in subfractions 46 and 47. 101

Figure 3.7. ..... Powders of Verbascoside and 2 structurally similar compounds are not receptor specific.. 102

Figure 3.8. ..... Flow chart of high throughput screen at the Broad Institute. 104

Figure 3.9. ..... Experimental protocol for 1536-well plate natural product screen 105

Figure 3.10. ..... High throughput screening for an rk antagonist screen validation. 106

Figure 3.11. ..... HTS screening plate layout and heatmap based read out. 107

Figure 3.12. ..... Screen shot of image representing luciferase from all plates for rk pilot screen done at the broad in 384 well plates. 108

Figure 3.13 ..... Z factors for all plates in Broad screen. 109

Figure 3.14. .....DOS Hit structures. 112

Figure 3.15. ..... The *Anopheles gambiae* (Ag) rk homolog GPR-rk is activated by Ag Bursicon conditioned media 113

Figure 3.16. ..... Inhibition by the four most potent small molecules identified during HTS. 114

CHAPTER 4

Figure 4.1 .....Cartoon models depicting recombinant membrane tethered ligand and corresponding soluble membrane anchored ligand..... 134

Figure 4.2 ..... Chemical structure, purity, and molecular weight of synthesized peptides 135

Figure 4.3. ..... Recombinant SubP MTL mediated signaling predicts activity of a corresponding SMAL on neurokinin receptors..... 137

Figure 4.4. ..... CP 99994 inhibits NK1R signaling induced by either a recombinant SubP MTL, soluble SubP with a C-terminal free acid (s-SubP-COOH), or the corresponding SMAL (l-SubP-COOH). 139

Figure 4.5 ..... Tethered glycine extended CCK4 (tCCK4-Gly) anticipates activity of the corresponding soluble membrane anchored ligand..... 141

Figure 4.6. .....YM022 inhibits CCK2R signaling induced by tethered CCK4-Gly, soluble CCK4-Gly, or lipidated CCK4-Gly. 142

Note: All experiments were performed by B. Harwood unless otherwise indicated in figure legend.

LIST OF ABRREVIATIONS

α-MSH, alpha melanocyte stimulating hormone
Ag, *Anopheles gambiae*
AMP, adult muscle precursor
Ap-Gal4, Apterous-Gal4
ATCHEC , Authenticated Traditional Chinese Herbal Extract Collection
β2AR, beta adrenergic receptor 2
Bur, bursicon
cAMP, cyclic adenosine monophosphate
CCAP, crustacean cardioactive peptide
CCK, cholecystokinin
CCL2, chemokine (C-C motif) ligand 2
CCL20, chemokine (C-C motif) ligand 20
CCR6, chemokine receptor 6
cDNA, complementary DNA
CHE, monomeric cherry fluorescent protein
CM, conditioned media
CRF, corticotropin-releasing hormone
CPS, counts per second
CRE, cAMP response element
Cy, Curly
DDT, dichlorodiphenyltrichloroethane
dLGR1/2/3. *Drosophila* Leucine-rich repeat receptors 1, 2 and 3.
DMEM, Dulbecco's modified eagle medium
DopR2, dopamine receptor 2
DOS, diversity oriented synthesis
Exendin-4, Exe-4
FBS, fetal bovine serum
GIP, gastric inhibitory polypeptide
GIPR, gastric inhibitory polypeptide receptor
GDP, guanine diphosphate
GFP, green fluorescent protein
GLP-1, glucagon-like peptide-1
GLP1R, glucagon-like peptide-2 receptor
GLP-2, glucagon-like peptide-2
GLP2R, glucagon-like peptide-2 receptor,
GN, glycine-asparagine
GPCR, G protein-coupled receptor
GPR-rk, Ag glycoprotein hormone rk-like receptor
GS, glycine-serine
GTP, guanine triphosphate
HEK, human embryonic kidney

hFSHR, human follicular stimulating hormone receptor
hLHR, human luteinizing hormone receptor
HOW-Gal4, Held Out Wing-Gal4
HPLC, high pressure liquid chromatography
HTS, high throughput screen
IFM, indirect flight muscle
LGR, leucine rich-repeat containing G protein-coupled receptor
MC4R, Melanocortin 4 receptor
MEF-2, myocyte enhancer factor-2
MLPCN, molecular libraries probe production centers network
MTL, membrane tethered ligand
NMJ, neuromuscular junction
NMR, nuclear magnetic resonance
PTH, parathyroid hormone
PTHR, parathyroid hormone receptor
PDF, pigment dispersion factor
PEG, polyethylene glycol
PEI, Polyethylenimine
PCAP, Pituitary adenylate cyclase-activating peptide
rk, rickets
SAR, structure activity relationship
sBur, soluble bursicon
SEG, subesophageal ganglion
SMAL, soluble membrane anchored ligand
SubP, substance P
Tb, Tubby^{1}
tBur, membrane tethered bursicon
TMD, transmembrane domain
TNFα ,tumor necrosis factor alpha
TSHR, thyrotropin stimulating hormone receptor
UAS, upstream activation sequence
VIP, vasoactive intestinal peptide
VIPR, vasoactive intestinal peptide receptor

Novel Modulators of Rickets Receptor Mediated Signaling *In Vitro* and *In Vivo*

General Introduction

Thesis Overview

This thesis will focus on rk/bursicon, a *Drosophila* GPCR and its corresponding endogenous ligand, which modulates melanization, sclerotization, and wing expansion. The sections presented in the introduction are selected to conceptualize rk within the broader context of the GPCR superfamily. We have included information regarding evolutionary conservation, signal transduction, and therapeutic relevance of GPCRs. In addition, information is presented to provide more in depth background regarding the molecular, pharmacological and physiological role of bursicon/ rk in insects. Prior to the background information we have also outlined the objective for each thesis chapter.

Chapter Objectives

The sections that follow summarize the objectives and rationale for each chapter. As a whole body of research, we have made significant contributions to extending the utility of membrane tethered ligand (MTL) technology and translating this approach into a practical tool for generating novel modulators of GPCRs. Utilizing a second generation of MTLs that we developed, we have constructed a bursicon heterodimeric MTL and used this to explore structure function relationships with its cognate receptor, rk. A corresponding bursicon MTL was also utilized *in vivo*. These studies led to a better understanding of the tissue specific requirements of rk and bursicon in *Drosophila* development. Finally, we have taken significant steps toward the identification of a small molecule rk antagonist, setting the stage to develop a new class of insecticides.

Chapter 1: Develop novel tools and study bursicon/rk in vitro

A reoccurring objective throughout this thesis was to develop novels tools that could be used to study G protein-coupled receptors *in vitro*. Specifically we sought to focus on a technology developed in the Kopin lab known as membrane tethered ligands (MTLs). As a model system to study MTLs, we chose the arthropod specific GPCR rickets (rk) and its cognate ligand bursicon. Our first goal was to develop a membrane tethered ligand that could be used to better understand the structure function relationships of bursicon and rk. This was particularly challenging in light of the fact that bursicon is a complex heterodimeric ligand comprised of two relatively large cysteine knot proteins.

Chapter 2: Utilize novel molecular tools to probe the tissue specific mechanism underlying rk activation in vivo

We then sought to take the technology and information garnered from *in vitro* structure function studies with MTLs and apply it to an *in vivo* system that could help better understand the tissue specific requirements of rk activation. Bursicon/rk signal transduction in *Drosophila* as a model system was appealing due to the fact that clear morphological readouts such as wing expansion defects can be easily monitored. While these studies were conducted in *Drosophila* they also have broad implications in many other insect species. Specifically, the results of inhibiting rk mediated signaling provide insight into whether rk might be a good insecticide target.

Chapter 3: Identify small molecule rk antagonists that can be used as complementary probes for understanding rk function and as a template for the development of a new class of insecticides.

The next goal was to identify chemical antagonists to better understand rk function *in vivo*; such molecules would complement the use of MTLs. Since our previous work with MTLs had extensively utilized an *in vitro* signal transduction assay, the ability to adapt this cell based assay into one that could be used to screen for chemical antagonists of rk was opportune. Rk is both physiologically important and uniquely found in arthropods. At the same time, based on the track record of other GPCRs, it was likely that rk would be amenable to chemical modulation. Taken together, screening of rk presented an exciting opportunity to identify an antagonist that would have promise as a putative insecticide

Chapter 4: Explore the conversion of recombinant membrane tethered ligands to a format that can be synthesized and delivered as a potential therapeutic

Finally given the broad applicability of membrane tethered ligand technology we wanted to develop a comparable synthetic anchored ligand equivalent that could be delivered without relying on recombinant DNA technology. For proof of process studies, we began by utilizing two well studied relatively small peptide hormones, substance P and cholecystokinin. We utilized these ligands to investigate whether the properties of membrane tethered ligands would be recapitulated as soluble membrane anchored peptides.

Background information

The following sections provide background information and additional context for the research in this thesis. An overview of relevant aspects of G protein-coupled receptors, bursicon/rickets mediated physiology, and membrane tethered ligand technology are highlighted below. These themes are important throughout the thesis.

GPCRs: structural definition of a receptor superfamily

G protein-coupled receptors (GPCRs) are a diverse class of cell surface proteins found in all eukaryotes. These proteins are characterized by a series of 7 transmembrane domains that integrate into the plasma membrane. The first cDNA sequence encoding a GPCR was cloned and sequenced 30 years ago. It was identified as the now prototypical GPCR rhodopsin (Nathans and Hogness 1983). Following the cloning of rhodopsin, a wide range of GPCRs were identified. It gradually became clear that GPCRs comprised a large family of receptors that controlled a diverse array of functions. The advent of genome sequencing has now allowed for a thorough examination of GPCRs in many species and has facilitated a comprehensive categorization of GPCRs. The GPCR superfamily is a diverse but evolutionarily conserved group of proteins. GPCRs are found in all eukaryotic organisms. For example the single cell organism *Saccharomyces cerevisiae* has a single GPCR encoded in its genome (Fredriksson and Schioth 2005). In contrast there are an estimated 791 GPCR encoding genes predicted in the human genome (Bjarnadottir, Gloriam et al. 2006). Efforts to sequence a diverse array of genomes over the past decade, has also uncovered a large number of GPCRs in the genomes of insects. For example there are 210 predicted GPCRs in the *Drosophila*

melanogaster genome and 268 in the *Anopheles gambiae* genome (Fredriksson and Schioth 2005).

Organization of the GPCR superfamily is conserved and can be broken down into 4 groups. Classification is based on amino acid sequence conservation. As seen in Figure i.1 the classes of receptors are as follows: rhodopsin-like (Class A), secretin-like (Class B), glutamate receptor-like (Class C), and others which include receptors such as the Frizzled receptors (Bjarnadottir, Gloriam et al. 2006).

The Class A rhodopsin-like receptors are the largest group of GPCR. This class includes rk, a major focus of this thesis. Rk falls within the leucine rich-repeat containing family of GPCRs (LGR) (Figure i.3). The LGR receptors all have an extracellular N-terminal leucine rich-repeat containing domain. LGR receptors are evolutionarily conserved in many phyla including nematodes, arthropods, and chordates (Van Hiel, Vandersmissen et al. 2011) .

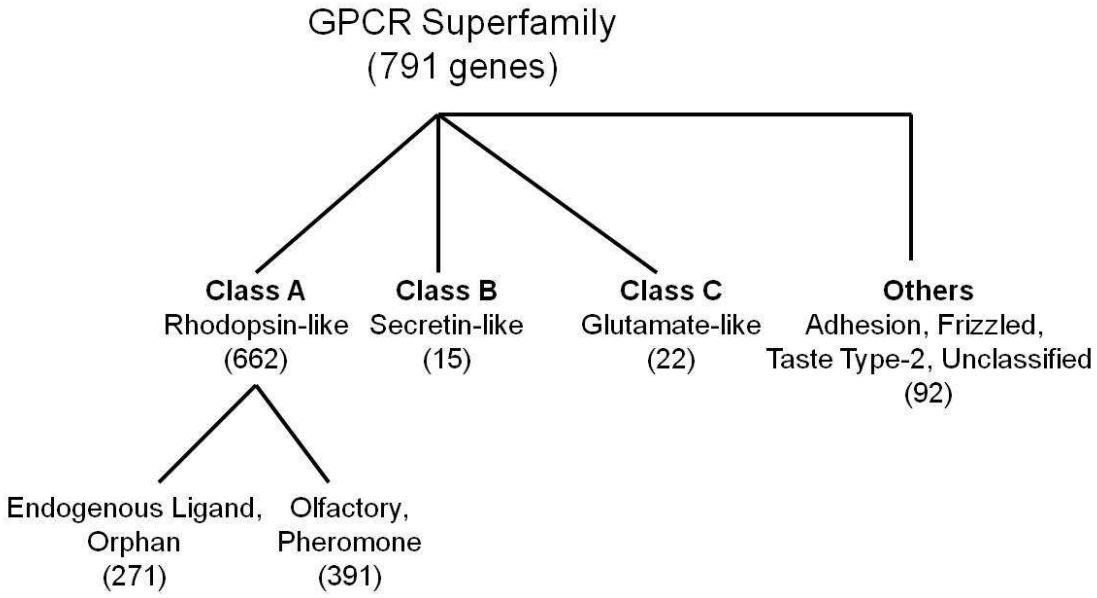


Figure i.1. Classification of the human GPCR superfamily. Adapted from (Bjarnadottir, Gloriam et al. 2006).

G Protein-coupled receptor signaling

Canonical GPCR signal transduction is triggered by dissociation of the heterotrimeric G protein (consisting of Gα, Gβ, and Gγ subunits) associated with the intracellular portion of the GPCR. Upon receptor activation, a conformational shift in the receptor results in the exchange of guanine diphosphate (GDP) for guanine triphosphate (GTP). The receptor shifts from an inactive to active conformation which thereby triggers the Gα subunit to dissociate from the Gβγ subunit complex. While this description gives an overview of canonical GPCR signaling, the details of this process are in reality substantially more complex. As shown in Figure i.2, there are in fact multiple subtypes for each Gα, Gβ, and Gγ subunit encoded in the human genome. Each G protein subunit subtype is used in combination with other subunit subtypes, thereby vastly increasing the complexity of signal transduction (Robishaw and Berlot 2004).

For the purposes of this thesis, we have a particular interest in the modulation of adenylate cyclase, the pathway through which rk signals. Two common pathways that result in opposing effects on adenylate cyclase are the $G\alpha_s$ and $G\alpha_i$ pathways. $G\alpha_s$, when activated results in an increase cAMP levels, while $G\alpha_i$ activation results in a decrease in intracellular cAMP levels (Figure i.2). Both $G\alpha_s$ and $G\alpha_i$ upon dissociation from Gβγ subunits respectively activate or inhibit the enzyme adenylate cyclase which in turn modulates intracellular cAMP levels (Pierce, Premont et al. 2002).

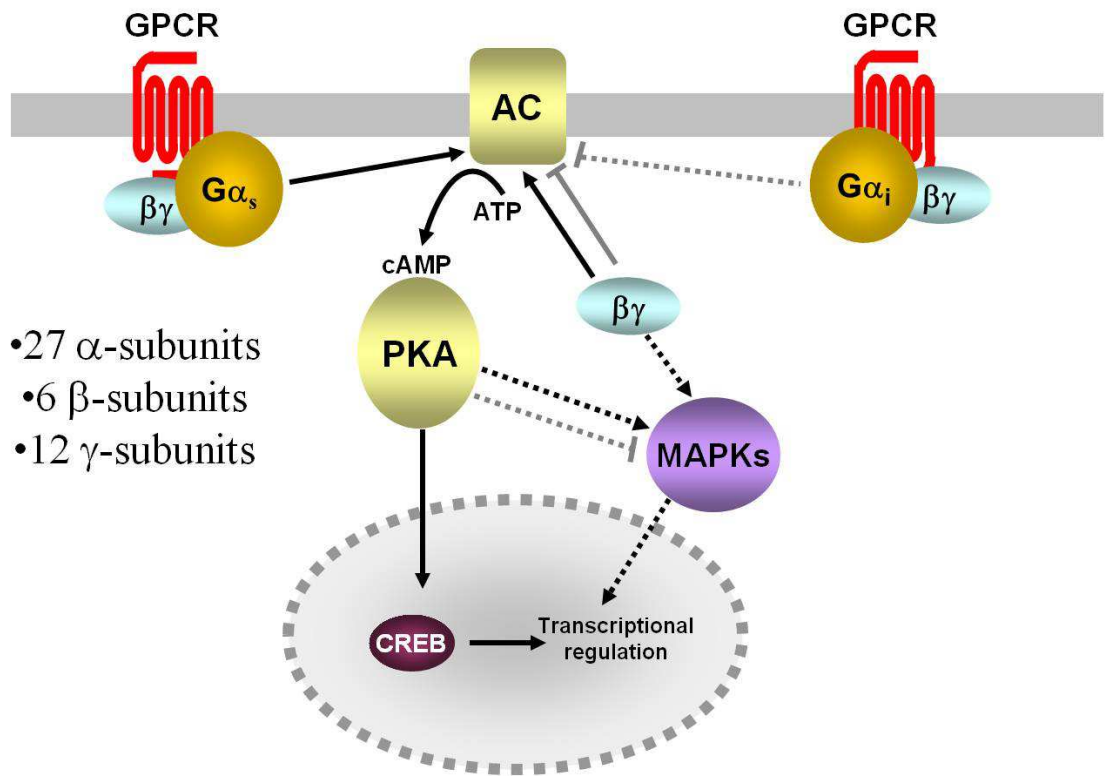


Figure i.2. Cartoon representing a GPCR signaling via either the $G\alpha_s$ or $G\alpha_i$ mediated pathways. The number of different G protein subtypes for each subunit highlights the complexity of signaling. Adapted from (New and Wong 2007).

In the context of this thesis, many of the studies are performed on the GPCR rk. The rickets receptor is $G\alpha_s$ coupled and increases intracellular cAMP levels when activated (Luo, Dewey et al. 2005). While these two mechanisms are classic examples of canonical GPCR signaling, it has become clear that GPCR activation potentially involves many additional pathways. For example the previously mentioned Gβγ heterodimer is also able to elicit a series of downstream signaling pathways upon dissociation from the Gα subunit (Khan, Sleno et al. 2013). Further adding another layer of complexity to signaling, receptors are silenced following ligand activation through multiple intersecting mechanisms (e.g. receptor phosphorylation, arrestin binding), ultimately leading to receptor internalization and/or desensitization (Pierce, Premont et al. 2002). While beyond the scope of this introduction, the complex mechanisms of GPCR signal transduction remains an active area of research and are specific to cellular context.

Constitutive signaling

The canonical Gα subunit signal transduction of GPCRs as previously described can occur either as ligand-independent (constitutive) or ligand-dependent. Constitutive activity is defined as GPCR signaling that is detectable in the absence of ligand. Different wild type GPCRs that signal using a common Gα subunit may differ in their levels of constitutive activity (Seifert and Wenzel-Seifert 2002). For example a comparison of constitutive activity of rk and the related *Drosophila* GPCR dLGR1 would show that intracellular cAMP triggered by dLGR1 expression are measurably higher than corresponding values with rk.

Alternatively constitutive receptor activity may be the result of point mutations in GPCRs. There are a variety of pathological conditions which have been linked to mutations that increase constitutive basal receptor activity. Relevant examples of these mutations have been identified in human LGR receptors. Activating mutations in both the luteinizing hormone receptor (hLHR) and the follicular stimulating hormone receptor (hFSHR) have been identified and linked to precocious puberty and ovarian hyperstimulation syndrome, respectively (Chan 1998; Zhang, Tao et al. 2007).

It should also be noted that our group and others (Nishi, Hsu et al. 2000) have demonstrated that both dLGR1 and rk show elevated basal levels of signaling. The constitutive activity of rk can be increased further though the introduction of point mutations (data not shown). However, the physiological consequence of wild type rk receptor basal signaling remains speculative and will be further discussed in Chapter 2.

Ligand dependent signaling and peptide ligand processing

Although receptors can possess varying degrees of constitutive activity, the vast majority of GPCR modulation occurs as a result of ligand-induced signaling. Ligands can either activate (agonists) or block (antagonists) receptor activity. There is a wide spectrum of endogenous GPCR ligands including biogenic amines, lipids, and peptides (Lin, Sassano et al.). In addition to these naturally occurring ligands, exogenous synthetic small molecules and peptides have been developed that can also modulate GPCR signal transduction (Wootten, Christopoulos et al. 2013).

The focus in this thesis is on a number of peptide hormone receptors. Peptide ligands are typically derived from larger pro-peptides that are proteolytically cleaved and

processed to form active peptides. Many peptides require further modification following cleavage of the mature peptide, including chemical modifications to the amino or carboxy termini of the peptides. Examples of such post-translational modifications include amidation (e.g. CCK, substance P), octanoylation (e.g. ghrelin), or glycosylation (e.g. LH, FSH)(Cuttitta 1993; Fox, Dias et al. 2001; Delporte 2013). For larger hormones proper folding is essential to promote the formation of disulfide bridges which in turn confer stability and activity under physiological conditions. Among the hormones where precise hormone folding is essential for function are the cystine-knot proteins (Darling, Ruddon et al. 2000). This group of hormones is of particular interest given that bursicon, a member of this family, is the ligand for rickets.

GPCR mediated physiology

G protein-coupled receptors mediate a wide array of physiological functions. This is true within each of the major classes of receptors. The range of functions regulated by GPCRs includes but is not limited to locomotion, satiation, anxiety, acid secretion, pain sensation, blood pressure, and heart rate (Jacoby, Bouhelal et al. 2006). Some GPCR mediated physiologies are conserved between mammals and lower organisms such as insects. For example the neurotransmitter dopamine and its cognate receptors are important for controlling locomotor activities in both humans and *Drosophila* (Draper, Kurshan et al. 2007). In contrast some GPCRs have also significantly diverged or have been lost throughout evolution. For example octopamine receptors are unique to insects and are involved in learning and memory. Currently there is no known GPCR activated by octopamine in vertebrates (Farooqui 2007). As outlined in the next section, rk has

some degree of structural conservation when compared with the mammalian LGR family however functionally this arthropod receptor has completely diverged.

Divergence of physiology: rk and LGR5

The amino acid sequences of some GPCRs have been conserved between insects and vertebrates; however their functional roles have diverged. An example of this scenario is with the LGR family of receptors. The most well studied receptors in this family are the mammalian LGR homologs, the glycohormone receptors; thyrotropin stimulating hormone receptor (TSHR), luteinizing hormone receptor (LHR), and follicular stimulating hormone receptor (FSHR). In mammals these receptors respond to heterodimeric ligands (glycohormones) and control a variety of physiological functions including metabolism, ovulation, and puberty. Insect genomes also encode LGR receptors. For example the *Drosophila melanogaster* genome encodes 4 family members, dLGR1-4 (Van Loy, Vandersmissen et al. 2008). In *Drosophila*, ligands for both dLGR1 and dLGR2 (rk) have been identified while dLGR 3 and 4 remain orphans (Luo, Dewey et al. 2005; Sudo, Kuwabara et al. 2005). A clear example of GPCR functional divergence is between the rk receptor its corresponding mammalian homologs. While rk, like the glycohormone receptors (mammalian LGRs) is activated by a heterodimeric ligand, and is $G\alpha_s$ coupled, there is no interspecies physiological overlap between this insect receptor and these mammalian homologs. Among the human LGR receptors, rk is most closely related to LGR5 (Figure i.3).

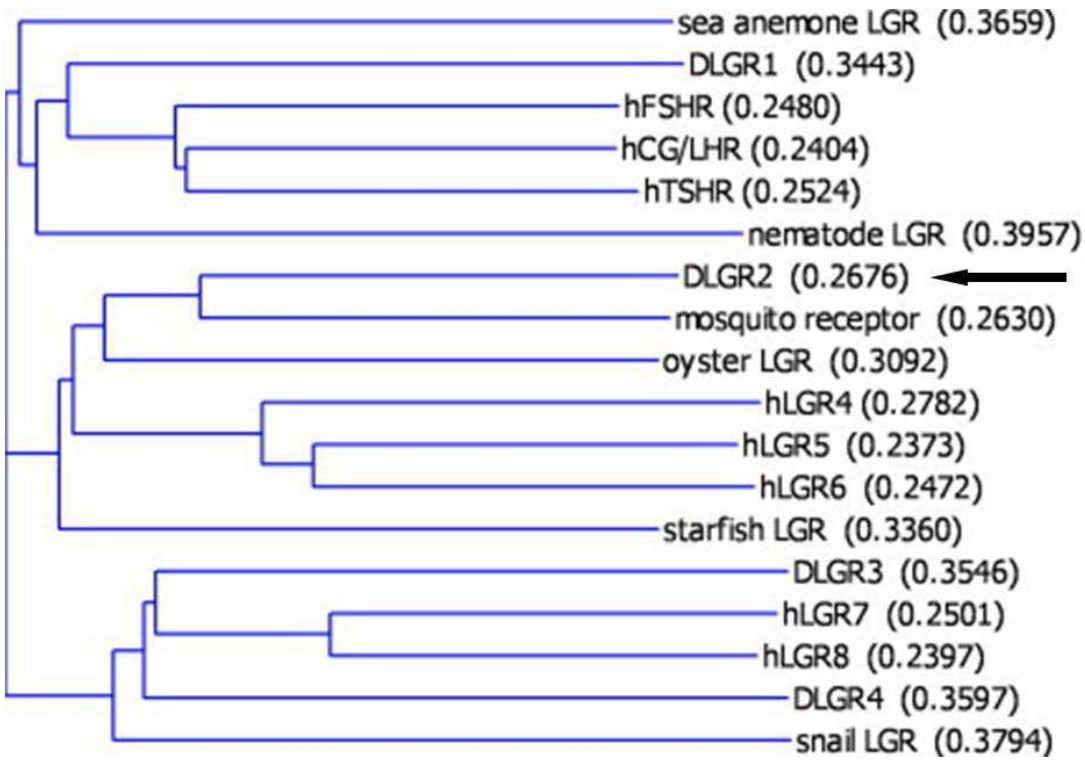


Figure i.3. Phylogenetic tree of select LGR subfamily members.
Arrow indicates DLR2, the *Drosophila* rk receptor. Adapted from (Van Loy, Vandersmissen et al. 2008)

LGR5 has been identified as a gastrointestinal stem cell marker and an upstream regulator of Wnt signaling (Garcia, Ghiani et al. 2009; Sato, Vries et al. 2009). Recently mammalian LGR4 and 5 have been shown to activate the Wnt/β Catenin pathway through G-protein independent mechanisms using a class of secreted ligands known as R-spondins (Carmon, Gong et al. 2011). As will be discussed in depth throughout this thesis, rk is important for wing expansion and cuticle formation in insects. As with the other mammalian LGRs there is currently no functional link between LGR5 and rk. Thus although rk is most closely related to human LGR5, this is solely based on amino acid sequence conservation. Current research suggests that both their ligands and functions have diverged.

Conservation of bursicon and other LGR receptor ligands

As previously mentioned, rickets is activated by the heterodimeric cystine-knot protein bursicon which is comprised of two related cystine-knot proteins conserved among invertebrates (Van Loy, Van Hiel et al. 2007). Active bursicon requires two unique subunits, bursicon α (Bur α) and partner of bursicon β (Bur β). Based on a characteristic structural conformation created by a motif of cysteine residues in key positions, bursicon falls into a protein family with little sequence homology, but a similar knotted conformation. Critical to the characteristic knot structure are the aforementioned cysteine residues that form disulfide bridges required to maintain its shape. Specifically both bursicon alpha and bursicon beta fall into the CAN family of eight membered ring cystine-knot proteins. This family also includes the TGFβ bone morphogenetic protein

antagonists. These proteins are known to be required for development and organogenesis (Avsian-Kretchmer and Hsueh 2004).

In addition to the BMP antagonists, bursicon is structurally related to the family of glycohormone cystine-knot proteins that activate LGR GPCRs in humans. These include hLHR, hFSHR, and hTSHR (Figure i.3). Glycohormones for these receptors share a common α subunit and each have a unique β subunit that is proposed to confer receptor specificity(Van Loy, Vandersmissen et al. 2008).

Unlike bursicon, glycohormones, like their name suggests are post-translationally glycosylated which is an important feature for receptor activation (Wheatley and Hawtin 1999). Interestingly the mammalian glycohormone receptors also appear to have an additional more ancient ligand. GPA2/GPB5 is also a heterodimeric cystine-knot protein encoded from two genes in the human genome. Interestingly a homolog for this heterodimeric ligand is also found in the *Drosophila* genome and activates the LGR receptor dLGR1, further suggesting a common ancestry (Sudo, Kuwabara et al. 2005). While comparable dLGR1 ligands have been identified in vertebrates, no bursicon like sequences have yet to have been identified in any chordates. Bursicon like sequences have only been identified in arthropods and echinoderms (Van Loy, Van Hiel et al. 2007).

While bursicon and rk appears to be been lost in vertebrates, each is conserved among insects and crustaceans. The recent use of whole genome sequencing and transcriptome analysis has rapidly increased the number of arthropods where rk and bursicon like sequences have been identified. Bursicon and rk like sequences were recently analyzed in 17 species of insects to better understand how polymorphisms at

specific amino acid residues in receptors and ligands co-evolve (Hughes 2012). While multiple *Drosophila* species were used in the analysis, so was *Anopheles Gambiae*, the honey bee *Apis Mellifera*, and the deer tick *Ixodes scapularis*. Using transcriptome analysis, bursicon like sequences have been identified in the crustacean *Calanus finmarchicus*, the most highly abundant zooplankton in the North Atlantic(Christie, Roncalli et al. 2013). Additionally another abundant crustacean in the Southern Ocean, the Ice Krill, *Euphausia crystallorophias* also expresses bursicon like sequences (Toullec, Corre et al. 2013). Widespread evolutionary conservation suggests that bursicon and rk are critical in arthropod physiology.

Rickets/bursicon in vivo: known function and beyond

Bursicon and rickets were identified for their roles in cuticle melanization and wing expansion processes which occur immediately after emergence of the adult insect following metamorphosis (eclosion). In *Drosophila*, bursicon is sequentially secreted from two distinct clusters of neuroendocrine cells in newly emerged flies. An initial wave of hormone is released from neurons in the subesophageal ganglion (SEG), in turn inducing secondary release of bursicon from a distinct subset of neurons in the abdominal ganglion. Neurons in both the SEG and abdominal ganglion that release bursicon have been identified as a subset of neurons that also release crustacean cardioactive peptide (CCAP) (Park, Schroeder et al. 2003; Luan, Lemon et al. 2006; Peabody, Diao et al. 2008). In fact, it is thought that bursicon and CCAP are packaged in the same vesicle and are co-released into the circulatory fluid (hemolyph) of the newly emerged adult fly (Woodruff, Broadie et al. 2008). CCAP is known to target the heart to increase cardiac

contractility (Dulcis, Levine et al. 2005). Concurrently, bursicon while circulating in the hemolyph, regulates the process of wing expansion, which requires a choreographed set of behaviors including grooming, perch selection, abdominal contractions, air gulping, and hemolyph pumping (Peabody, Pohl et al. 2009).

The first physiological processes shown to be modulated by bursicon included melanization and hardening of the cuticle (tanning). The role of bursicon was elegantly demonstrated in a series of relatively simple experiments. It was shown that if the neck of a blow fly is ligated during eclosion but prior to cuticle tanning; melanization and sclerotization will not occur. This is due to the fact that bursicon is not released into the hemolymph and thus cannot activate rickets. However, if a downstream signaling molecule cAMP analog (8-bromo-cAMP) is injected into the body of a neck ligated fly, it will subsequently tan (Honegger, Dewey et al. 2008).

Additional support for the functional role of rk/bursicon mediated signaling comes from genetic studies in *Drosophila*. There are two classic mutant fly lines that fail to tan when injected with bursicon. Each of these mutant flies has a single base pair substitution that leads to a premature stop codon in the rickets coding sequence. If 8-br-cAMP, an analog of cAMP (which is downstream of rk) is administered to either of these decapitated flies, they will tan supporting that the pathway is mediated by the $G\alpha_s$ activity of rk (Baker and Truman 2002).

Functionally bursicon and rk are thought to control a similar process in other insect species as well. Immunohistochemical studies in *Anopheles Gambiae* suggest that bursicon is released from a similar group of CCAP positive neurons comparable to those in *Drosophila* (Honegger, Estevez-Lao et al. 2011). This similar pattern of release

suggests that bursicon also controls wing expansion, melanization, and cuticle hardening in mosquitoes. Treatment of the tobacco hornworm *Manduca sexta* with recombinant bursicon promotes wing expansion and tanning, also suggesting a conserved function in this moth species (Dai, Dewey et al. 2008). It is postulated that burs/rk is responsible for these functions in a wide spectrum of insects where this ligand and receptor pair have been shown to be present in the genome (e.g. ticks, lice, red flour beetles, honey bees, pea aphids, and mosquitoes) (Hughes 2012).

In addition to localization studies, experiments using RNAi also reinforce that bursicon function is conserved. RNAi knockdown of bursicon or rickets in both *Drosophila* and the red flour beetle, *Tribolium castaneum* led to defects in wing expansion and tanning. In *Drosophila* and Tribolium it has been noted that both bursicon and rk knockdown also result in some developmental lethality (Arakane, Li et al. 2008; Loveall and Deitcher 2010). What remains less clear is (i) what tissues require rk activation and (ii) what the developmental consequence are if signaling is perturbed.

Rk as a potential insecticide target.

The evolutionary conservation and requirements of rk and bursicon for insect development, makes rickets an appealing putative insecticide target. The amino acid sequence variability between rk orthologs in different species and the specificity of GPCR ligands provides an opportunity to develop species selective insecticides. However, prior to this possibility, lead compounds that can modulate rk activity need to be developed as none currently exist. Precedent has been set for the identification of chemical modulators of LGR receptors. A high throughput screen of small molecules

was able to identify an agonist for the human TSH receptor. Upon finding a receptor agonist, they were also able to modify this compound to make a receptor antagonist. The fact that a screen has been successful for another LGR receptor suggests that a similar strategy could also be successful for rk.

GPCRs and therapeutics.

The fact that rk is a GPCR, enhances the likelihood of finding novel chemical modulators. GPCRs are involved in a diverse array of physiologies and have long been targets for therapeutic development. Their plasma membrane location and structural configuration makes GPCRs highly amenable to pharmacological modulation. Currently, as seen in Figure i.4, approximately 30% of all pharmaceuticals on the market target GPCRs (Lundstrom 2009). These include a wide variety of both peptides and synthetic small molecules. Given the track record and clinical importance of GPCRs, they remain a high priority in the continuing search for novel therapeutics by pharmaceutical and biotechnology companies. However, few efforts have attempted to leverage these features in the context of developing an insecticide.

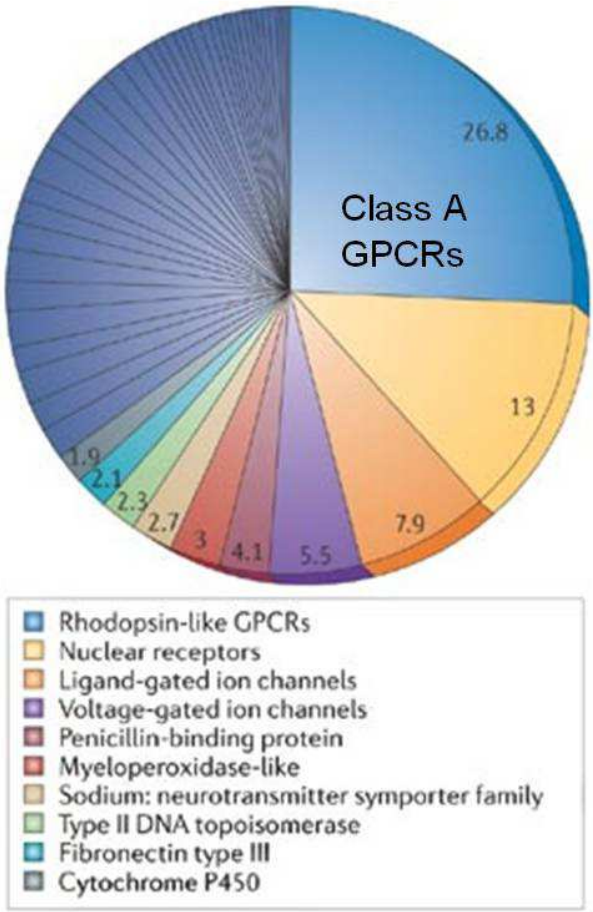


Figure i.4 Pie Chart showing the percentage of FDA approved drugs by target class. Class A GPCRs represent the largest percentage of targets for FDA approved drugs. Adapted from (Overington, Al-Lazikani et al. 2006)

History of membrane tethered ligands

Membrane tethered ligands (MTLs) are recombinantly expressed extracellular ligands that are anchored to the cell membrane using a transmembrane domain. The predecessor to membrane tethered ligands targeting GPCRs, were several membrane tethered toxins generated as modulators of ion channels(Auer and Ibanez-Tallon 2010). A prototype construct, tethered α-bungarotoxin was anchored to the cell membrane using a GPI linker to modulate nicotinic receptors in zebrafish (Ibanez-Tallon, Wen et al. 2004). Given that membrane anchored ligands were capable of blocking ion channels, it was postulated that parallel membrane anchored ligands could modulate GPCRs. Confirming this hypothesis, the Nitabach and Kopin laboratories demonstrated in a collaborative study that a GPI linked pigment dispersion factor (PDF) (i.e. tethered PDF) acted as an agonist at the *Drosophila* PDF receptor (PDFR). Membrane tethered PDF (tPDF) was the first membrane tethered ligand of a GPCR and was functional using both *in vitro* and *in vivo* assays. *In vivo* expression of tPDF in a PDF null background was able to partially rescue abnormal circadian rhythms (Choi, Fortin et al. 2009). Following the encouraging results with tethered PDF, the Kopin lab subsequently made a series of peptide encoding MTLs cable of activating a wide range of class B GPCRs (Figure i.5) (Fortin, Zhu et al. 2009). Follow up structure function studies on MTLs demonstrated that these constructs could be genetically engineered to have unique properties. For example introduction of two amino acid changes into the peptide sequence of membrane tethered Exendin, a glucagon-like peptide 1 receptor agonist, could convert the construct from a membrane tethered agonist to a membrane tethered antagonist (Fortin, Zhu et al. 2009).

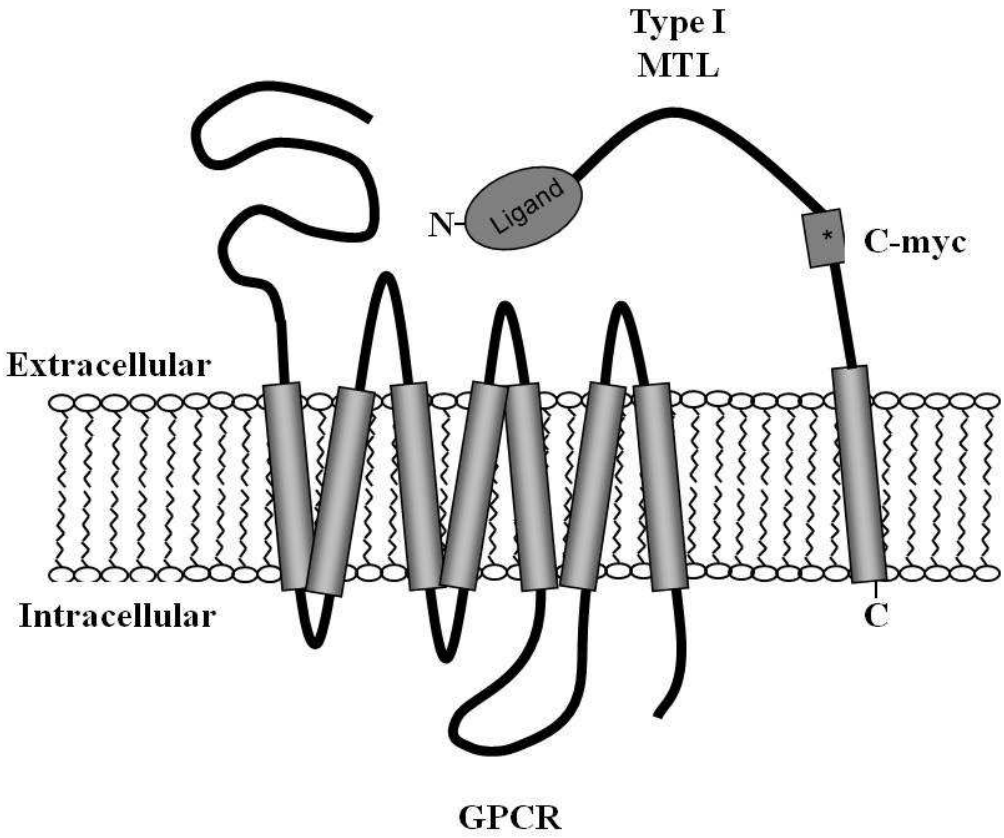


Figure i.5. Cartoon representing a GPCR and Type I TMD membrane tethered ligand at the cell surface. Adapted from (Fortin, Zhu et al. 2009)

The list of peptides and GPCRs which have been successfully targeted using the MTL system has grown rapidly. Given the large range of peptides in Table i.1, it has become readily apparent that membrane tethered ligand technology is a widely applicable technology for modulating GPCRs and for understanding the interaction between peptides and their cognate receptors.

Membrane tethered bursicon

In the context of previously studied MTLs, we set out to generate a membrane tethered bursicon construct to better understand both the structure function relationship of the receptor/ligand pair and the physiology linked to signaling. Even in theoretical terms, generating membrane tethered bursicon posed some unique challenges. Since bursicon is normally a heterodimeric ligand, making two membrane tethered ligands that need to interact and subsequently activate a receptor pushed the boundaries of MTL technology. In addition each of the bursicon subunits was a cysteine knot protein thus greatly increasing the complexity of the ligands incorporated into the MTL. It should also be noted that all of the previously mentioned membrane tethered ligands utilized a type I transmembrane domain for cellular anchoring (results in a free N terminus of the tethered peptide). Since bursicon membrane tethered ligands were being developed de novo, we also wanted to determine whether a type II transmembrane domain could be utilized instead (resulting in a free C terminus of the tethered peptide).

Table i.1. List of peptides converted into active MTLs and their cognate receptors.

Membrane Tethered Ligand	**Receptor**
gastric inhibitory polypeptide (GIP)	gastric inhibitory polypeptide receptor (GIPR)
glucagon-like peptide-1 (GLP-1)	glucagon-like peptide-1 receptor (GLP1R)
glucagon-like peptide-2 (GLP-2)	glucagon-like peptide-2 receptor (GLP2R)
Exendin-4 (Exe-4)	GLP1R
vasoactive intestinal peptide (VIP)	vasoactive intestinal peptide (VIPR)
pituitary adenylate cyclase-activating peptide (PCAP)	pituitary adenylate cyclase-activating peptide receptors
parathyroid hormone (PTH)	Parathyroid hormone PTH receptor (PTHR)
corticotropin-releasing hormone (CRF)	corticotropin-releasing hormone receptors
calcitonin	calcitonin receptor
substance P	Neurokinin receptors
cholecystokinin-8 (CCK-8)	cholecystokinin 1 &2 receptors (CCK2R), (CCK1R)
cholecystokinin-4 (CCK-4)	cholecystokinin 2 receptor (CCK2R)
galanin	galanin receptor 1 and 2
met-enkephalin	μ opiod receptor
pigment dispersion factor (PDF)	pigment dispersion factor receptor (PDFR)
amylin	calcitonin receptor
bursicon	rickets receptor
chemerin	chemerin receptor
chemokine (C-C motif) ligand 2 (CCL2)	chemokine receptor 2 (CCR2)
chemokine (C-C motif) ligand 20 (CCL20)	chemokine receptor 6 (CCR6)

Utility of membrane tethered ligands

One feature of membrane tethered ligands that makes them useful tools for *in vitro structure function* studies is their recombinant nature. These DNA constructs can be efficiently modified using standard molecular approaches. In contrast, for standard soluble peptides, alteration of protein sequence or length would require labor intensive efforts including synthesis and purification. Given the size and complexity of bursicon, the synthetic option is impractical. However being able to generate cDNAs that encode a series of bursicon MTLs that can be recombinantly expressed, is a feasible way to create a unique series of readily available tools for the study of rk mediated function.

When expressed in transgenic models, MTLs offer an effective means to probe receptor specific function in selected tissues. Given that soluble GPCR ligands diffuse after peptide administration, it is difficult to determine if observed effects result from targeted tissue modulation or from ligand that has diffused and acted elsewhere. Membrane tethered ligands offer a novel approach that enables the tissue specific study of GPCRs. This is especially relevant in the genetically tractable *Drosophila* model organism. Utilizing the UAS/Gal4 system in conjunction with MTLs enables the tethered ligands to be expressed and to act only in selected tissues. Tethering ensures that the ligand does not diffuse into surrounding tissues resulting in off target confounding effects. All of these advantages of MTLs can be leveraged with bursicon to help dissect the tissue specific requirements of bursicon/rk signal transduction *in vitro* and *in vivo*.

Extensions of membrane tethered ligands: lipidated peptides

While the recombinant nature of MTLs is useful for *in vivo* studies in model organisms, this mode of delivery is not conducive to the development of therapeutics.

Studies of MTLs have shown membrane anchoring confers unique properties when compared to the corresponding soluble ligands (e.g. enhanced potency and duration of activity). To extend the utility of tethered ligands, it would be desirable to have ligands that mimic the favorable features of an MTL without requiring expression of DNA. Toward this objective we have developed soluble membrane anchored ligands (SMALs). Once a peptide has been optimized as an MTL, it can then be synthesized as a SMAL which includes three domains: peptide, linker (e.g. PEG) and anchor (palmitate). SMALs can be directed delivered to a target tissue (e.g. lung, skin) where they anchor and act on target cell receptors. In Chapter 4 we explore the methodology to convert membrane tethered ligands to SMALs.

Chapter 1:

Membrane tethered bursicon constructs as heterodimeric modulators of the *Drosophila* GPCR rickets

Abstract

The study of complex heterodimeric peptide ligands has been hampered by a paucity of pharmacological tools. To facilitate such investigations we have explored the utility of membrane tethered ligands (MTL). Feasibility of this recombinant approach was explored with a focus on *Drosophila* bursicon, a heterodimeric cystine-knot protein that activates the G protein-coupled receptor rickets (rk). Rk/bursicon signaling is an evolutionarily conserved pathway in insects required for wing expansion, cuticle hardening, and melanization during development. We initially engineered two distinct MTL constructs each comprised of a type II transmembrane domain, a peptide linker, and a C-terminal extracellular ligand that corresponded to either the α or β bursicon subunit. Co-expression of the two complementary bursicon MTLs triggered rk mediated signaling *in vitro*. We were then able to generate functionally active bursicon MTLs in which the two subunits were fused into a single heterodimeric peptide, oriented as either α-β or β-α. Carboxy-terminal deletion of 32 amino acids in the β-α MTL construct resulted in loss of agonist activity. Co-expression of this construct with rk inhibited receptor-mediated signaling by soluble bursicon. We have thus generated membrane-anchored bursicon constructs that can activate or inhibit rk signaling. These probes can be used in future studies to explore the tissue and/or developmental stage-dependent effects of bursicon in the genetically tractable *Drosophila* model organism. In addition, our success in generating functionally diverse bursicon MTLs offers promise that such technology can be broadly applied to other complex ligands including the family of mammalian cystine-knot proteins.

Introduction

The *Drosophila* receptor rickets (rk, dLGR2) is a member of the leucine-rich repeat subfamily of G protein-coupled receptors (GPCRs) (Van Loy, Vandersmissen et al. 2008). Rk activation is required for wing expansion, cuticle sclerotization, and melanization. The endogenous rk agonist, bursicon, is a heterodimeric cystine-knot protein. Bursicon has been known as the insect tanning hormone for more than four decades, however it was only in 2005 that it was discovered that the active ligand is comprised of two unique subunits, BURS (Bur α) and Partner of bursicon (Bur β) (Luo, Dewey et al. 2005; Mendive, Van Loy et al. 2005).

Rk/bursicon signaling is highly conserved among insects and has been shown to play an important role in development (Van Loy, Van Hiel et al. 2007; Bai and Palli 2010; Loveall and Deitcher 2010). In *Drosophila,* bursicon is sequentially secreted from two distinct clusters of neuroendocrine cells shortly following eclosion. An initial wave of hormone is released from neurons in the subesophageal ganglion, which in turn induces secondary release of bursicon from another subset of neurons in the abdominal ganglion. This sequence ultimately triggers wing expansion, cuticle hardening and pigmentation (Luan, Lemon et al. 2006; Peabody, Diao et al. 2008). RNAi studies in *Drosophila* revealed that down regulation of rk during development compromises insect survival (Dietzl, Chen et al. 2007; Loveall and Deitcher 2010). Although bursicon and rk signaling have been most extensively investigated in *Drosophila*, other studies have shown that this pathway is essential for viability of other insect species including *T. castaneum* (Bai and Palli 2010). Rk and bursicon like sequences have been identified in a

wide variety of insects, suggesting that the physiological significance of this signaling cascade has been highly conserved (Van Loy, Van Hiel et al. 2007; Honegger, Dewey et al. 2008; An, Wang et al. 2009; Honegger, Estevez-Lao et al. 2011). The vast majority of research on rk/bursicon has focused on the functional role of this regulatory system during development. One limitation of these efforts stems from the paucity of pharmacologic modulators of this GPCR which can be used as experimental tools.

Both bursicon subunits (Bur α and Bur β) are members of the eight membered ring cystine-knot proteins. This family also includes the TGFβ bone morphogenetic protein antagonists known to be required for development and organogenesis (Avsian-Kretchmer and Hsueh 2004). Bursicon is also structurally related to the family of glycohormone cystine-knot proteins that activate leucine-rich repeat GPCRs. Corresponding mammalian GPCRs include the luteinizing hormone (LHR), follicular stimulating hormone (FSHR), and thyroid stimulating hormone receptors (TSHR). The glycohormone ligands share a common α subunit and each have a unique β subunit that confers receptor specificity (Hearn and Gomme 2000). Comparison of bursicon/rk, with LH, FSH, and TSH ligand/receptor pairs suggest that these structurally related hormones and GPCRs arose from common ancestors (Van Loy, Vandersmissen et al. 2008).

Generating pharmacological tools to probe the physiology of rk/bursicon *in vivo* presents considerable practical hurdles. Like mammalian glycohormones bursicon is comprised of two large and complex molecules. As a result, conventional peptide synthesis is impractical for making functionally active ligand. In addition, introduction of mutations into corresponding recombinant DNA constructs aimed at expressing variant

cystine-knot proteins in heterologous cell lines may be hampered by impaired processing and/or secretion of the peptide (Darling, Ruddon et al. 2000; Galet, Guillou et al. 2009) .

To circumvent these challenges we have extended a strategy that our lab has previously utilized to study relatively short GPCR peptide ligands. Membrane tethered ligand (MTL) technology uses recombinant DNA to encode an extracellular peptide hormone fused to a linker sequence and a transmembrane domain. To date, a variety of short peptide MTLs have been generated that selectively activate cognate class B GPCRs (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Fortin, Chinnapen et al. 2011). Previous investigations also demonstrated the utility of membrane tethered toxins as ion channel blockers (Auer and Ibanez-Tallon 2010).

In the current report we demonstrate that large, complex cystine-knot proteins which require a heterodimeric partner can be generated as functionally active MTLs. Furthermore, we utilize this extended MTL technology to identify a ligand domain that is required for rk receptor activation and to generate an inhibitor of rk signaling. In addition to providing insights into the structure function relationships underlying bursicon activity, respective constructs provide novel tools for further analysis of associated physiology *in vivo*. Extending from our current investigation, the approach developed for bursicon can be utilized to study related cystine-knot proteins (e.g. glycohormones, bone morphogenetic protein antagonists) as well as other complex peptide ligands.

Materials and Methods

Cell Culture

Human embryonic kidney cells (HEK293) were cultured in Dulbecco's modified eagle medium (DMEM, Life Technologies, Grand Island, NY) with 10% Fetal Bovine Serum (FBS, Atlanta Biologicals, Lawrenceville GA), 100U/mL penicillin, and 100µg/mL streptomycin (Life Technologies, Grand Island, NY). Cells were maintained at 37 °C in a humidified 5% CO_2 atmosphere.

Plasmids

DLGR2 (rk), GenBank: AF142343.1 was generously provided by Dr. Cornelis Grimmelikhuijzen and subcloned into pcDNA1.1 using the restriction enzymes HindIII and XbaI (Eriksen, Hauser et al. 2000). Bursicon α and β cDNAs in pcDNA3.1 were generously provided by Dr. J. Vanden Broeck (Mendive, Van Loy et al. 2005). The type II MTL backbone was generated by PCR amplification of the transmembrane domain (amino acids 10-56) of Tumor Necrosis Factor alpha (TNFα) from a cDNA template (NCBI accession # BC028148) (Marmenout, Fransen et al. 1985). The nucleotide and corresponding amino acid sequence of the Type II construct is presented in Figure 1.7. The bursicon α and β subunits were subcloned by PCR into the Type II MTL backbone. The bursicon MTL constructs included nucleotide sequences corresponding to amino acids 30-143 for Bur α and 21-141 for Bur β (Mendive, Van Loy et al. 2005). For the CHE-tBur-β-α construct, a cDNA encoding cherry fluorescent protein was ligated in

frame 5' of the TNFα transmembrane domain coding sequence to create an intracellular fluorescent tag.

The negative control MTL, CHE-tCCL2 contains the same fluorescent tag and backbone as CHE-tBur-β-α with the alpha and beta subunits replaced by amino acids 25-99 of human chemokine ligand 2 (NCBI accession # NP_002973.1). All cAMP response element (6X-CRE) reporter genes and β-galactosidase plasmids were as previously described (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009).

Transfections

Polyethylenimine (PEI) transfection reagent was prepared as previously described (Zaric, Weltin et al. 2004). All transfections were done using a final PEI concentration of 2 μg/mL. Transfections were performed in serum free DMEM with antibiotics at 37 °C. Cells were incubated with transfection mix for 20-48 hours as indicated in the figure legends prior to initiating functional or MTL expression assays.

Bursicon Conditioned Media

HEK293 cells were seeded in 75cm^2 flasks at 1.2 million cells/flask. Twenty four hours later cells were co-transfected with 4μg each of bursicon α and bursicon β cDNAs (or 4μg of a chimeric α-β cDNA construct as indicated) with PEI as previously described. Following a 24 hour incubation, the transfection media was aspirated and 12ml serum free DMEM with antibiotics was added. Medium was conditioned for 48 hours, then collected, centrifuged at 1600 g for 5 minutes to remove cellular debris, aliquoted, and stored at -80 °C.

Luciferase Assays

Luciferase assays were done as previously described (Hearn, Ren et al. 2002; Al-Fulaij, Ren et al. 2007; Fortin, Zhu et al. 2009) with minor modifications. HEK293 cells at ~80% confluence in 96-well plates were transiently transfected using PEI. To assess rk activity, each well was co-transfected with cDNAs encoding; rk (0.25ng), a luciferase reporter gene under the control of a cAMP response element (6X-CRE-Luc, 5ng), bursicon constructs as indicated in the figure legends, and β-galactosidase as a transfection control (5ng). To assess the function of tethered ligands, luciferase levels were quantified 24 hours after transfection using Steady-Light (PerkinElmer, Waltham, MA**)** and normalized relative to β-galactosidase as previously described (Fortin, Zhu et al. 2009). To assess the function of soluble ligands, bursicon conditioned media was added twenty hours after transfection for an additional 4 hour incubation. Luciferase and β-galactosidase levels were then measured as indicated above.

Confocal Microscopy

HEK293 cells were transfected in 35mm glass bottom dishes (Mattek, Ashland, MA). Cells were transfected with cDNAs encoding rk, a cherry fluorescent protein tagged tBur-β-α construct, and a GFP reporter gene under the control of a cAMP response element (6X-CRE-GFP) reporter gene (Fortin, Chinnapen et al. 2011). After 48 hours, the cells were fixed for 10 minutes using 4% paraformaldehyde in PBS. The cells were subsequently washed with PBS containing 100mM glycine and then kept covered with

PBS to prevent drying. Microscopy was performed on a Leica TCS SP2 confocal microscope with an inverted 40x oil objective. Two channels were used to simultaneously monitor MTL expression (mCherry fluorescent protein) and rk activation (GFP).

ELISA

ELISA was performed to quantify the expression of membrane tethered ligands. MTL encoding cDNAs were transfected into HEK293 cells grown in 96-well plates. Twenty four hours after transfection, the media was replaced with 50 μL of DMEM +10% FBS with antibiotics; the cells were then grown for an additional 24 hours. Following this period, ELISA was performed as previously described (Fortin, Zhu et al. 2009; Doyle, Fortin et al. 2012) using a rabbit polyclonal c-myc conjugated HRP antibody at 1:2500 dilution (Abcam, Cambridge, MA).

Programs and Statistics

All luciferase and expression data were graphed and analyzed using GraphPad Prism 5 (GraphPad Software Inc, La Jolla, CA). All cDNA sequences were designed and analyzed using Vector NTI Advanced 9 software (Life Technologies, Grand Island, NY).

Results

In this study we explored the utility of MTLs as pharmacological tools for studying complex heterodimeric protein ligands. We applied this technology to bursicon and rickets as a prototypical ligand-receptor pair. As a first step we generated two independent tethered constructs; one encoding the alpha subunit and the other encoding the beta subunit of bursicon (Figure 1.1).

The design of these membrane tethered subunits included a type II transmembrane domain (TMD) with the intent of expressing the peptide ligand at the extracellular C-terminus. Type II transmembrane domains specifically orient within the plasma membrane such that the N-terminus is intracellular and the C terminus is extracellular. To verify the predicted orientation of the MTL, an ELISA directed at the extracellular c-myc epitope included in the construct was performed. In unpermeabilized HEK293 cells, both individually and co-expressed α and β subunit constructs were readily detected at the cell surface (Figure 1.2A and Figure 1.8).

We next examined ligand induced signaling. Co-expression of cDNAs encoding both bursicon subunit MTLs together with rk and a 6X-CRE-Luc reporter gene led to concentration dependent receptor activation (Figure 1.2B). Co-expression of rk with each MTL alone (α or β) and a 6X-CRE-Luc reporter gene did not trigger signaling.

To quantify the magnitude of the signal obtained with co-expression of both bursicon MTLs, comparison was made relative to the soluble bursicon (sBur). To enable these studies, we generated bursicon conditioned media by co-expressing both α (sBur-α) and β (sBur-β) subunit cDNAs in HEK293 cells and collecting the supernatant as detailed

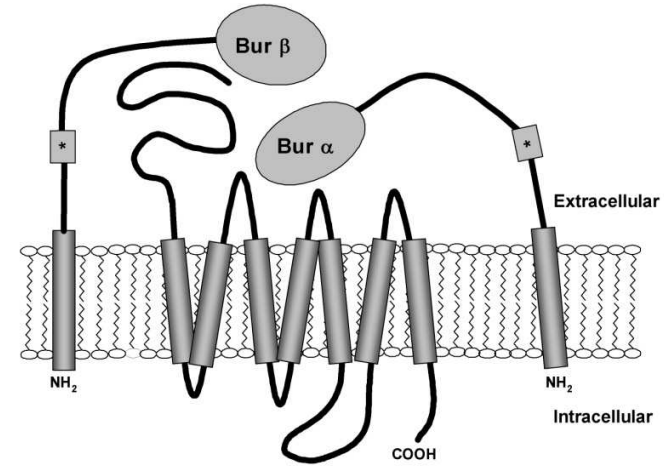


Figure 1

Figure 1.1. Schematic representation of heterodimeric membrane tethered bursicon subunits coexpressed with rk. Abbreviations: *=c-myc epitope tag.

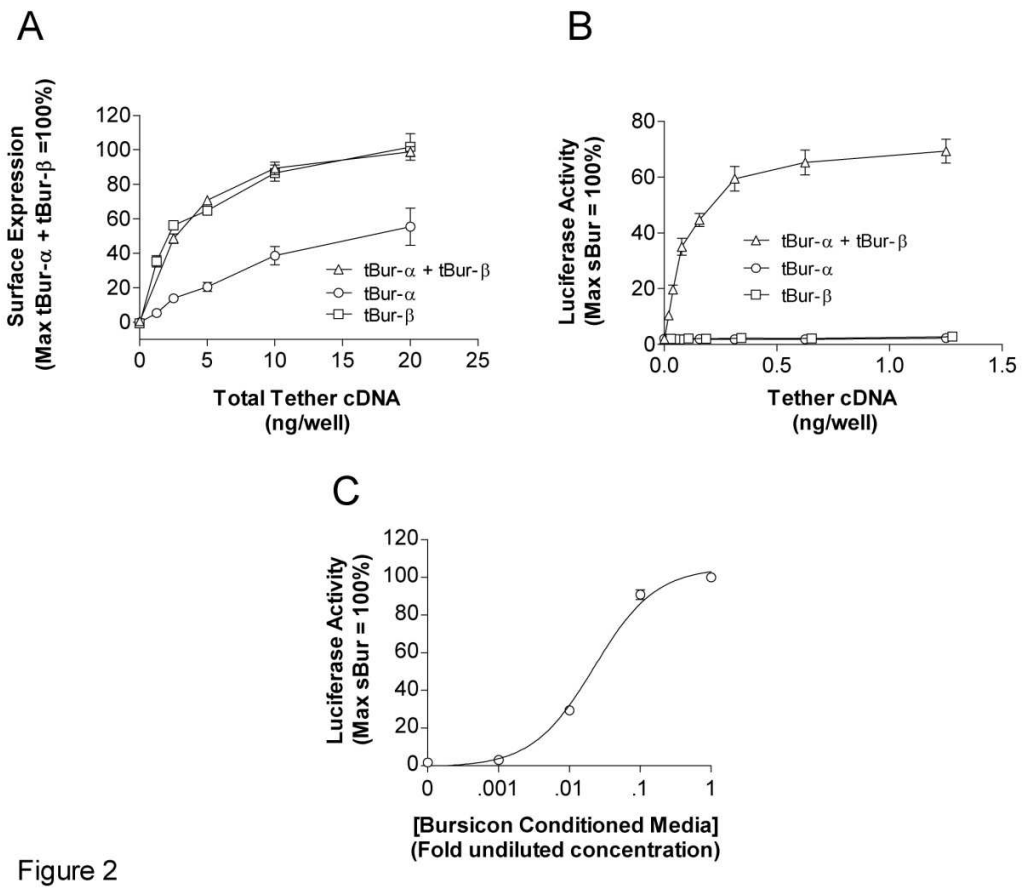


Figure 2

Figure 1.2. Bursicon membrane tethered α and β subunits together activate the *Drosophila* rickets receptor.
A) Quantification of cell surface expression of bursicon MTLs. Forty eight hours after transfection ELISA was performed using an antibody directed against a c-myc epitope. The x-axis denotes the total amount of cDNA transfected. B) Tethered ligand-induced activation of rk mediated signaling. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, one or both bursicon MTL subunit(s), and a β-galactosidase gene to control for transfection variability. For tethered ligand activity, twenty four hours after transfection luciferase activity was quantified and normalized relative to a four hour maximal stimulation of rk with bursicon conditioned media. The x-axis denotes the amount transfected for each cDNA subunit. C) Concentration dependent activation of rk with bursicon conditioned media. A series of tenfold dilutions of conditioned media (1=undiluted conditioned media) was added to cells 20 hours after transfection; the duration of ligand stimulation was 4 hours. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tBur= bursicon MTL subunit cDNA

in methods (Luo, Dewey et al. 2005; Mendive, Van Loy et al. 2005). When conditioned media was applied to rk expressing cells, the resulting heterodimeric bursicon ligand triggered concentration dependent luciferase reporter gene activity (Figure 1.2C).

In parallel studies we demonstrated that co-expression of freely soluble bursicon subunits in HEK293 cells together with rk and a 6X-CRE-luc reporter gene, also led to receptor mediated signaling (Figure 1.3A). Consistent with the known requirement of bursicon to form a heterodimer, expression of either subunit alone did not trigger receptor mediated signaling.

The above studies set the stage to examine whether co-expression of one soluble (α or β) and one tethered ligand (β or α) would trigger receptor mediated signaling. As shown in Figure 1.3B, the soluble and tether combinations were active. In contrast, when conditioned media was generated from a single subunit cDNA (α or β) and added to cells expressing the complementary tethered subunit (β or α), no rk activation resulted (data not shown).

To further simplify a system for studying complex heterodimeric ligands, we explored the potential of membrane tethered fusion proteins as functional ligands (Figure 1.4A). The initial MTL β-α fusion protein that was generated positioned the α subunit at the construct's free extracellular C-terminus. When co-expressed with rk and the luciferase reporter gene, this MTL triggered GPCR mediated signaling. A parallel construct with the opposite order of subunits (α-β, where the carboxy-terminus of the β subunit was at the free extracellular end of the protein) also activated rk. Activity of each bursicon fusion protein MTL was similar regardless of orientation of the subunits.

In addition, signaling by these MTLs was comparable to co-expression of tBur-α and tBur-β (Figure 1.4B). Notably, expression levels of the fusion MTLs as assessed by

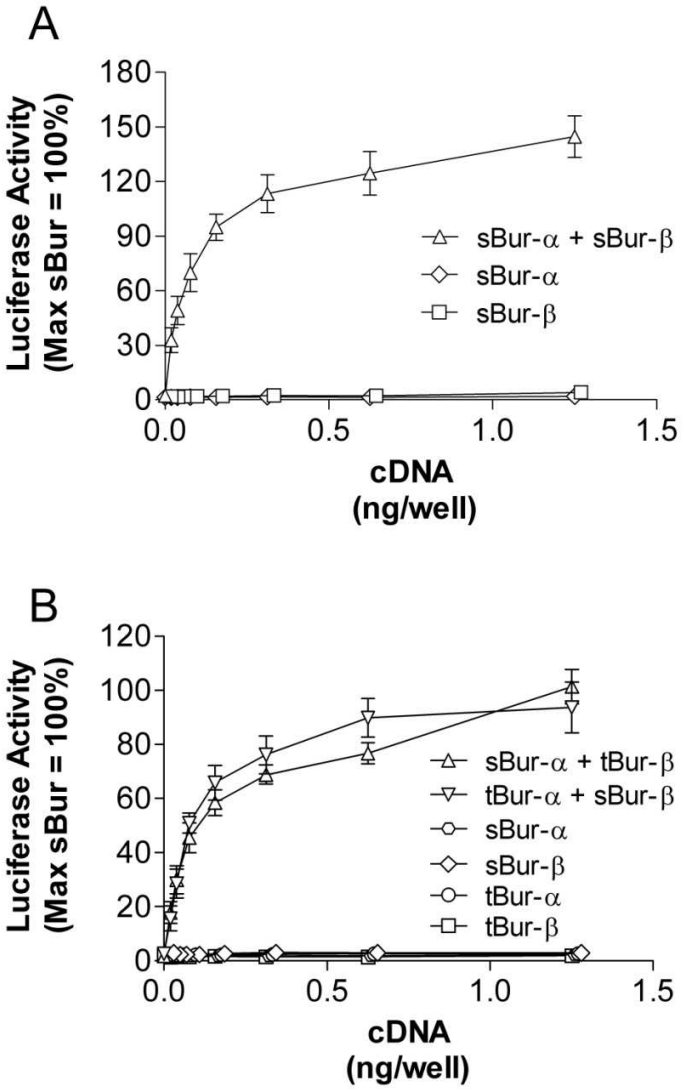


Figure 3

Figure 1.3. Rk is activated by co-expression of either two complementary soluble bursicon subunits or complementary combinations of soluble and tethered bursicon subunits.
HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, and either soluble bursicon subunits (A) or combinations of soluble and tethered bursicon subunits (B). The x-axes denote the amount transfected for each cDNA subunit. Twenty four hours following transfection, luciferase activity was quantified and activity values were normalized relative to maximal stimulation of rk with the addition of independently prepared bursicon conditioned media. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations. sBur = soluble Bursicon subunit cDNA, tBur=bursicon MTL subunit cDNA.

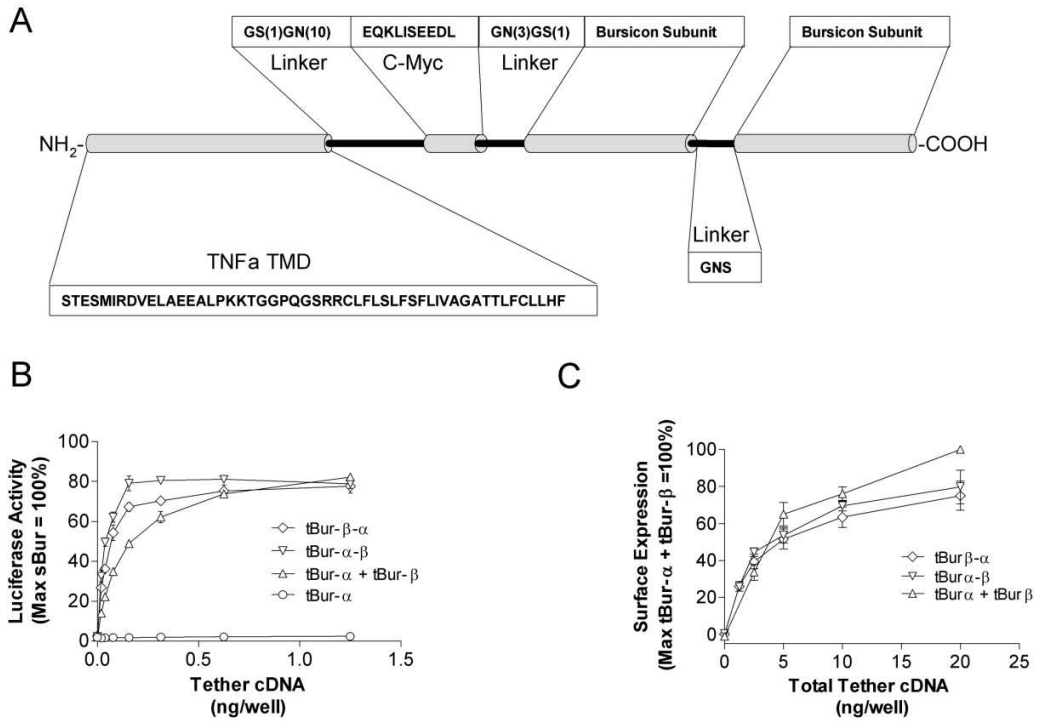


Figure 4

Figure 1.4. Bursicon MTLs are active fusion proteins independent of C-terminal subunit positioning

A) Cartoon illustrating the protein structure of membrane tethered constructs that include the two complementary bursicon subunits. B) Tethered ligand induced activation of rk mediated signaling. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, the indicated bursicon MTL encoding construct(s), and a β-galactosidase control gene. Twenty four hours after transfection, luciferase activity was quantified and normalized relative to maximal stimulation of rk with addition of bursicon conditioned media. The x-axis denotes the amount transfected for each cDNA subunit. C) Quantification of cell surface expression of bursicon MTLs. Forty eight hours after transfection ELISA was performed using an antibody directed against a c-myc epitope. The x-axis denotes the total amount of cDNA transfected. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tBur= bursicon MTL subunit cDNA, TNFα=Tumor necrosis factor α.

64ELISA was also comparable to the levels observed with single subunit constructs (Figure 1.4C). These latter experiments confirmed, as observed with MTLs including a single bursicon subunit, that the ligand domains of tethered β-α and α-β are localized in the extracellular space As an additional control we demonstrated that conditioned media cannot be made from cells expressing either a heterodimeric fusion MTL or co-expressing individual subunit MTLs (Figure 1.9). This observation indicates that bursicon MTLs are not secreted.

As a complementary index of tBur-β-α function (in addition to luciferase activity) we visually monitored ligand expression as well as MTL induced signaling using a 6X-CRE-GFP reporter gene. To enable these studies, a tBur-β-α construct was generated that included a cherry fluorescent protein at the intracellular amino-terminus, (CHE-tBur-β-α). After co-transfection of cDNAs encoding: CHE-tBur-β-α, rk and a 6X-CRE-GFP reporter gene, MTL expression and receptor mediated signaling could be simultaneously observed by confocal imaging. As shown in Figure 1.5, CHE-tBur-β-α expression results in rk activation, triggering GFP production. In contrast, a non-specific MTL CHE-tCCL2 (designed to activate the CCR2 receptor) can be visualized at the cell surface, but does not trigger rk mediated signaling (no 6X-CRE-GFP expression is induced).

In summary, our results with recombinant bursicon demonstrate that co-expression of both α and β subunits, either as two soluble peptides, or as two independent membrane tethered constructs is sufficient to generate active hormone. In addition, a single heterodimeric MTL with the peptide ligand in either the β-α or α-β configuration results in active bursicon. All bursicon MTLs appear to specifically activate rk (dLGR2).

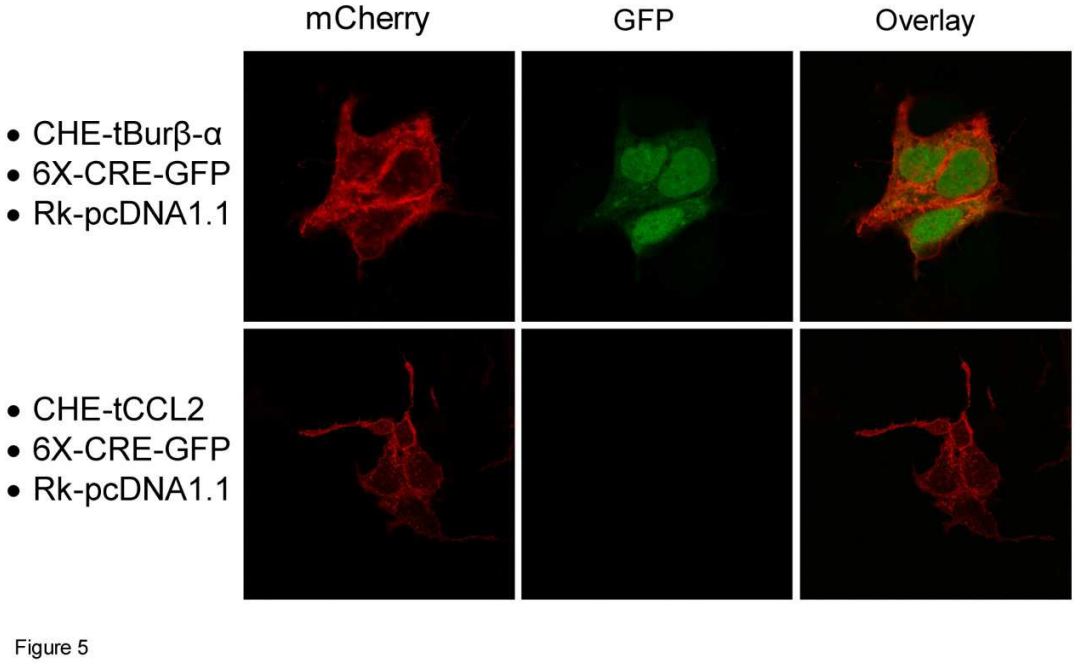


Figure 1.5. Rk activation by a bursicon MTL fusion protein can be visually monitored by confocal microscopy.
Representative images showing bursicon mCherry-fluorescent protein (CHE) MTL triggering rk mediated GFP expression. HEK293 cells were transiently co-transfected with cDNAs encoding: CHE-tBurβ-α or CHE-tCCL2 (negative control), rk, and a 6X-CRE-GFP reporter gene. Confocal images were obtained 48 hours after transfection. Data represent 3 independent experiments. Abbreviations: CRE-GFP=cAMP response element green fluorescent protein reporter gene.

When tested on related *Drosophila* LGR receptors (dLGR1, dLGR3), no activation could be detected (Figure 1.10.).

The ability to express recombinant functionally active bursicon heterodimers as a single MTL fusion protein enabled an expedited strategy for *structure function* studies. As a first step, we examined the effect of serial deletions at the C-terminus of the tBur-β-α construct (Figure 1.6A). Deletion of 10, 21, 32, or 35 amino acids from the C-terminus of tBur-β-α led to a progressive loss of MTL activity. The Δ10 and Δ21 constructs were partial agonists compared to full length tBur-β-α. In contrast, little if any activation of rk was detected with expression of Δ32 and Δ35 MTLs (Figure 1.6B). Deletion of the C-terminus had little effect on cell surface expression levels (Figure 1.6C). In contrast to the tBur-β-α constructs, corresponding deletions of the C-terminus of tBur-α-β (up to or including the final cysteine residue) did not result in loss of agonist activity (Figure 1.11).

Further analysis of tBur-β-α Δ32 and Δ35 MTLs revealed these tethered constructs markedly inhibited receptor stimulation by soluble bursicon (conditioned media). Functional antagonism of tBur-β-α Δ35 was suggested by a significant rightward shift of the conditioned media concentration response curve when this construct was expressed (Figure 1.6D). With the tBur-β-α Δ32 an even more pronounced inhibition resulted, essentially eliminating agonist induced signaling.

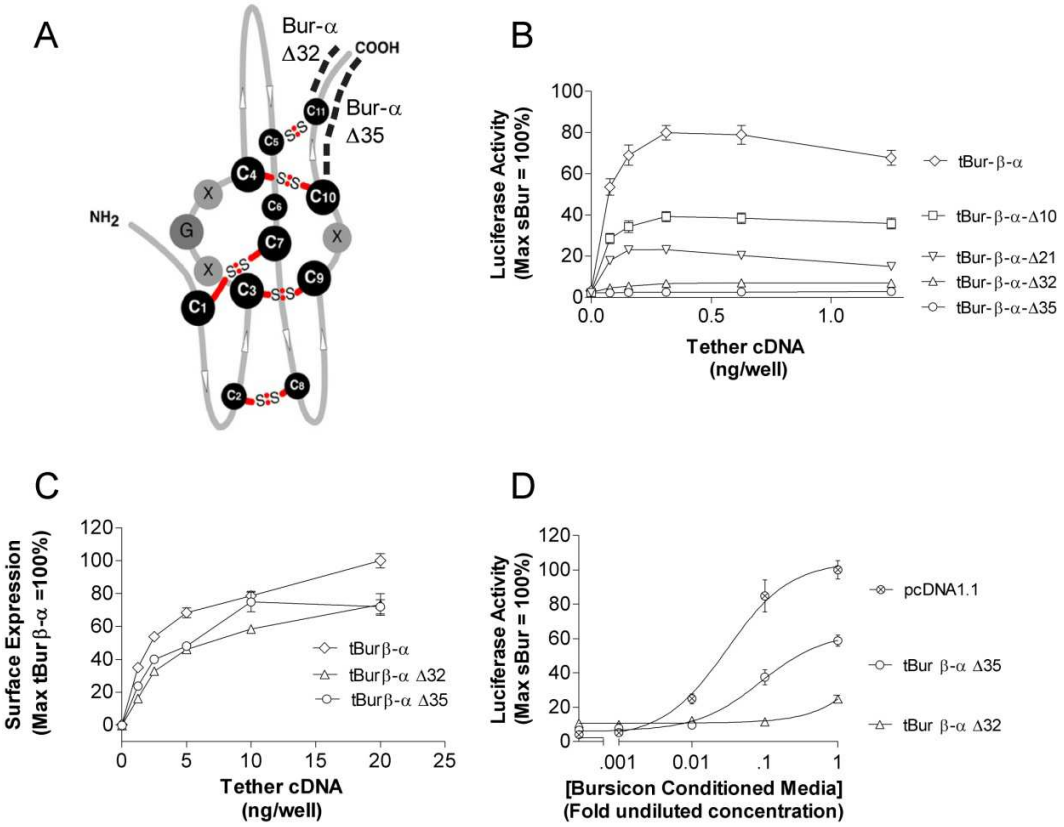


Figure 6

Figure 1.6. Development of a membrane tethered inhibitor of rk signaling.
A) Cartoon illustrating the secondary structure of the bursicon α subunit, a cystine-knot protein, highlighting the relative positions of deleted domains (dotted lines) (adapted from Honegger et al., 2008, Figure 2). B) Activity screen of bursicon MTL serial deletions. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, the indicated bursicon MTL, and a β-galactosidase control gene. Twenty four hours after transfection, luciferase activity was quantified and normalized relative to maximal stimulation of rk after addition of bursicon conditioned media. C) Quantification of cell surface expression of full length vs. C-terminally truncated bursicon MTLs. Forty eight hours after transfection ELISA was performed using an antibody directed against a c-myc epitope. D) Expression of rk bursicon MTL C-terminal deletion constructs disrupts receptor activation by soluble bursicon. HEK293 cells were transiently transfected with 2ng of the indicated MTL construct, rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase control gene. Twenty hours following transfection bursicon conditioned media was added at a series of tenfold dilution (1=undiluted conditioned media). Following a four hour incubation with bursicon conditioned media, luciferase activity was quantified and normalized relative to maximal stimulation of rk by bursicon conditioned media in the absence of a tethered inhibitor. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate.
Abbreviations: tBur= bursicon MTL subunit cDNA.

Discussion

We have developed novel recombinant constructs that enable membrane-anchored expression of bioactive bursicon. Our study demonstrates that MTL technology can be applied to larger and considerably more complex GPCR ligands than those described in prior reports. Previously, only MTLs that included short peptide ligands (up to 39 amino acids) have been described (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Auer and Ibanez-Tallon 2010; Fortin, Chinnapen et al. 2011; Ibanez-Tallon and Nitabach 2012). In contrast, the mature bursicon subunits, α and β, are 141 and 121 amino acids, respectively. Furthermore, each of these subunits is a cystine-knot protein that includes a series of intramolecular disulfide bridges which confer tertiary structure. As an additional prerequisite of agonist activity, the α and β subunits must interact to form a structurally integrated heterodimer (Mendive, Van Loy et al. 2005).

Given the stringent requirements underlying the formation of active soluble bursicon, including cellular co-expression, co-processing and co-secretion, the success in generating corresponding functional membrane tethered ligands could not have been anticipated. Initially, we demonstrated that expression of both single tethered bursicon subunits (alpha and beta) in the same cell was sufficient to generate an active ligand. Follow-up studies revealed that co-expression of soluble and tethered complementary subunits also enabled the formation of active ligand. In contrast when a single soluble subunit was added as conditioned media to cells expressing a tethered complementary subunit, no agonist activity was detectable (data not shown). This finding suggests that intracellular assembly of the alpha-beta heterodimer is a critical step in the formation of active hormone. These observations are consistent with reports on the heterodimerization requirements of soluble bursicon and other cystine-knot proteins that are known to undergo

intracellular assembly prior to secretion as an active ligand (Xing, Myers et al. 2004). Remarkably, both membrane tethered and soluble bursicon subunits, despite the complexity of processing, appear to be fully compatible with each other in forming active heterodimers.

In an attempt to further understand the structural requirements underlying tethered bursicon function, we generated constructs in which both the α and β subunits were included in a single MTL. Since an active tethered ligand can be generated as either a β-α or α-β fusion construct, neither a free N nor a free C-terminus is a requirement for agonist activity (Figure 1.4B). It is of note that conditioned media containing a soluble form of the bursicon fusion protein tested in the β-α arrangement also shows agonist activity (Figure 1.12). Whether tethered or soluble, the bursicon fusions are active ligands. Our observations with bursicon reveal another parallel with mammalian heterodimeric cystine-knot proteins. Fusion of the α and β subunits of mammalian glycohormones including TSH, LH, and FSH as single soluble peptides, also results in ligands that can activate their corresponding mammalian GPCR (Sugahara, Grootenhuis et al. 1996; Fares, Yamabe et al. 1998; Sen Gupta and Dighe 2000; Park, Semyonov et al. 2005; Setlur and Dighe 2007).

The generation of tethered bursicon fusion proteins provided a simplified model system to define domains of the dimer that are important for agonist activity (Figure 1.6B). These experiments were guided by prior observations that the β subunit of mammalian glycohormones provides specificity and affinity for cognate receptors while the α subunit is required for receptor activation (Park, Semyonov et al. 2005). Furthermore, the literature suggests that the C-terminal domain of the glycohormone alpha subunit is an important determinant for ligand activity (Sato, Perlas et al. 1997; Sen Gupta and Dighe 2000; Butnev,

Singh et al. 2002). Based on this knowledge, we generated a series of deletions in the C-terminus of Burs α in the context of the tBur-β-α heterodimer. These experiments demonstrated that the C-terminal domain in tethered bursicon was essential for rk activation. One of the deletion mutants in which 32 C-terminal residues were truncated (designated as Δ32) not only led to loss of agonist activity, but also markedly inhibited the function of soluble bursicon (Figure 1.6D). This observation suggests that once a domain essential for agonist activity was removed in the corresponding MTL, the remaining truncated peptide can inhibit soluble agonist induced signaling. An MTL with a larger C-terminal deletion (Δ35), while also lacking agonist activity, was much less effective (vs. Δ32) in blocking soluble bursicon induced signaling. The difference between Δ32 and Δ35 is that three additional highly conserved residues including a critical cysteine are truncated in Δ35. The loss of these 3 residues may have compromised the tertiary structure of the tethered ligand, in turn explaining the functional difference in constructs. Soluble versions of the Δ32 and Δ35 constructs did not confer the same ability to block ligand induced signaling (Figure 1.13). Thus it is possible that membrane anchoring is required to generate a functional antagonist.

It is of note that the GPCR targeted MTLs that had been reported prior to this study all shared a common orientation, in which the peptide ligand was expressed with a free extracellular N-terminus. In contrast, the bursicon MTLs were engineered with the opposite orientation (i.e., with a free extracellular carboxy-terminus). This was achieved by incorporating a different transmembrane domain anchor (a type II TMD) into the construct. The ability to generate membrane tethered ligands in either orientation markedly enhances the potential utility of MTL technology. For many peptides orientation may be a critical factor in generating an active MTL. It is well established that for peptide hormones recognizing class B GPCRs (e.g. Secretin, PTH, CRF, GLP-1,

GIP), the critical determinants of ligand efficacy reside in the N terminal domain of the hormone (Hoare 2005). We have previously shown that each of these peptides remains active when incorporated into an MTL that includes a type I TMD, i.e. the extracellular free end of the peptide is the N terminus (Fortin, Zhu et al. 2009). In contrast, peptide ligands recognizing class A GPCRs are more diverse. As examples, the amino termini of chemokines are generally considered critical for ligand activity whereas for neuropeptides, functional determinants are often localized at the carboxyl terminus (Eipper and Mains 1988; Mayer and Stone 2001). In the latter case, it is anticipated that MTLs including a type II TMD will preserve biological activity when corresponding peptides are anchored to the cell membrane.

In summary, we have developed a strategy that can be widely applied to the study of peptide ligands. More specifically, we have identified bursicon MTLs that either activate or block rk mediated signaling. These findings set the stage for future *in vivo* studies. In the investigations to follow, we intend to selectively express tethered constructs in targeted tissues of *Drosophila* thus exploring the utility of the approach for defining corresponding rk mediated pathways/physiologies. Precedent with these bursicon MTLs will set the stage for parallel studies using other tethered cystine-knot proteins as tissue selective molecular probes. Candidate MTLs include mammalian glycohormones as well as non-GPCR regulators such as bone morphogenetic protein antagonists. The efficiency and flexibility of recombinant MTL technology will enable generation of a wide range of unique tools to complement the use of soluble ligands in understanding corresponding receptor mediated physiologies.

Acknowledgements

We would like to thank Dr. Isabelle Draper and Dr. Jamie Doyle for their constructive suggestions throughout the course of this research.

Footnotes

This work was supported by the National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases [Grant 5R01DK070155]. This work was also supported by the synapse neurobiology training program, National Institute of Health National Institute of Neurological Disorders and Stroke [Grant T32-NS061764].

Type II Membrane Tethered Ligand backbone sequence

```
atgagcactgaaagcatgatccgggacgtggagctggccgaggaggcgctccccaagaag
 M  S  T  E  S  M  I  R  D  V  E  L  A  E  E  A  L  P  K  K
acagggggccccagggctccaggcggtgcttgttcctcagcctcttctccttcctgatc
 T  G  G  P  Q  G  S  R  R  C  L  F  L  S  L  F  S  F  L  I
gtggcaggcgccaccacgctcttctgcctgctgcactttggagtgatcggatcaggtaac
 V  A  G  A  T  T  L  F  C  L  L  H  F  G  V  I  G  S  G  N
ggaaacggtaatggtaacggtaatggaaacggtaacggaaatggtaatggtaacgaacag
 G  N  G  N  G  N  G  N  G  N  G  N  G  N  G  N  G  N  E  Q
aagctcatttcagaggaagacctgggaaatggtaacggaaacggatccgggaattct
 K  L  I  S  E  E  D  L  G  N  G  N  G  N  G  S  G  N  S
```

Figure 1.7. Nucleotide sequence encoding the TNFα type II transmembrane tethered ligand backbone.

All bursicon MTL constructs were generated using the same transmembrane domain and linker sequence. All backbone sequences included a TNFα transmembrane domain (blue shading), a repetitive glycine-asparagine linker (red shading) and a myc epitope tag (green shading). The expression plasmid pcDNA1.1 was used as the vector. Bursicon subunits were cloned in frame into the construct using BamHI and EcoRI restrictions sites (gray shading).

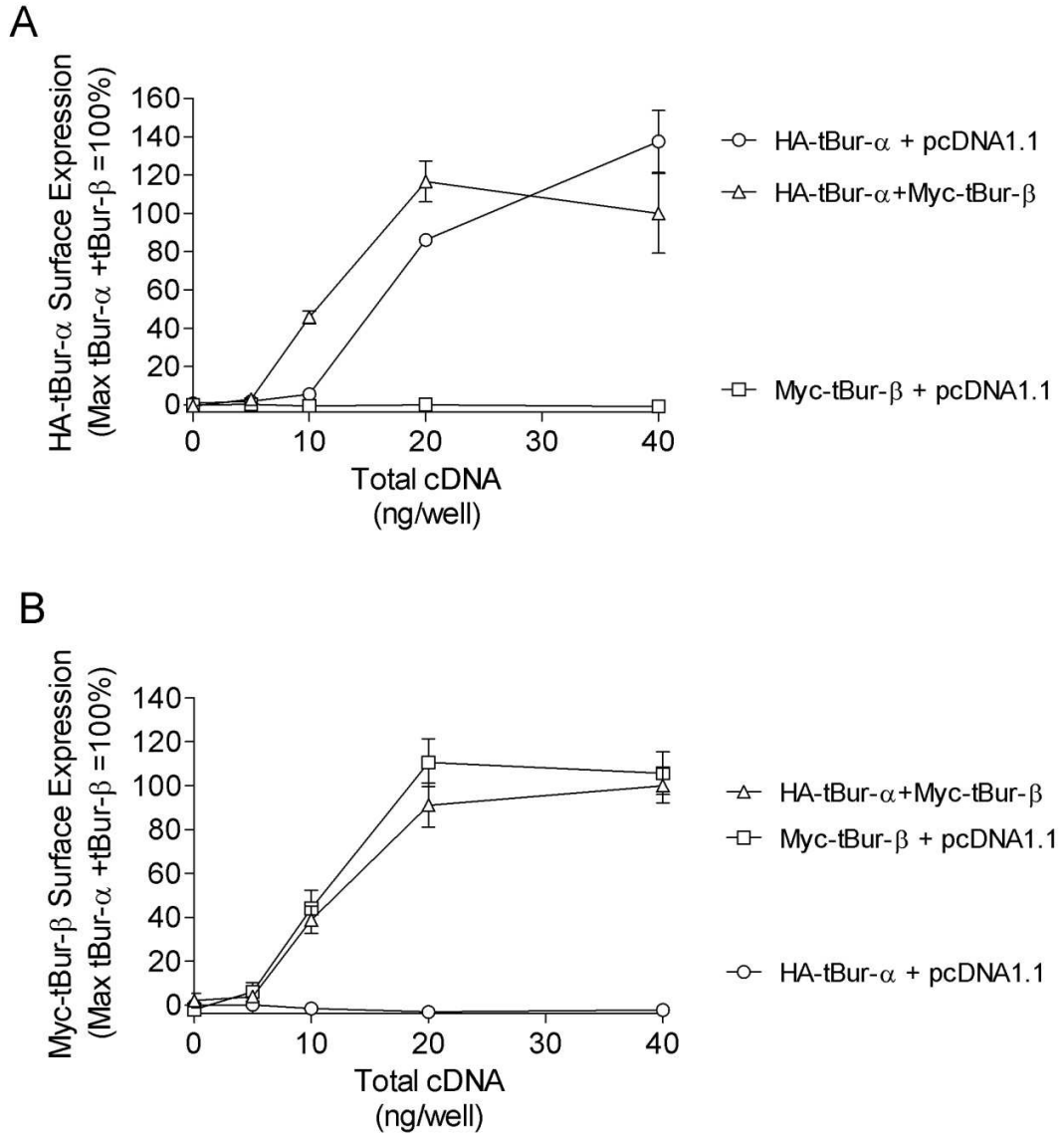


Figure 1.8. Tethered bursicon subunits are expressed independent of complementary subunit.

A) Quantification of cell surface expression of tBur-α in the presence or absence of tBur-β. Forty eight hours after transfection ELISA was performed using an antibody directed against an HA epitope. B) Quantification of cell surface expression of tBur-β in the presence or absence of tBur-α. Forty eight hours after transfection ELISA was performed using an antibody directed against a c-myc epitope. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tBur= bursicon MTL subunit cDNA

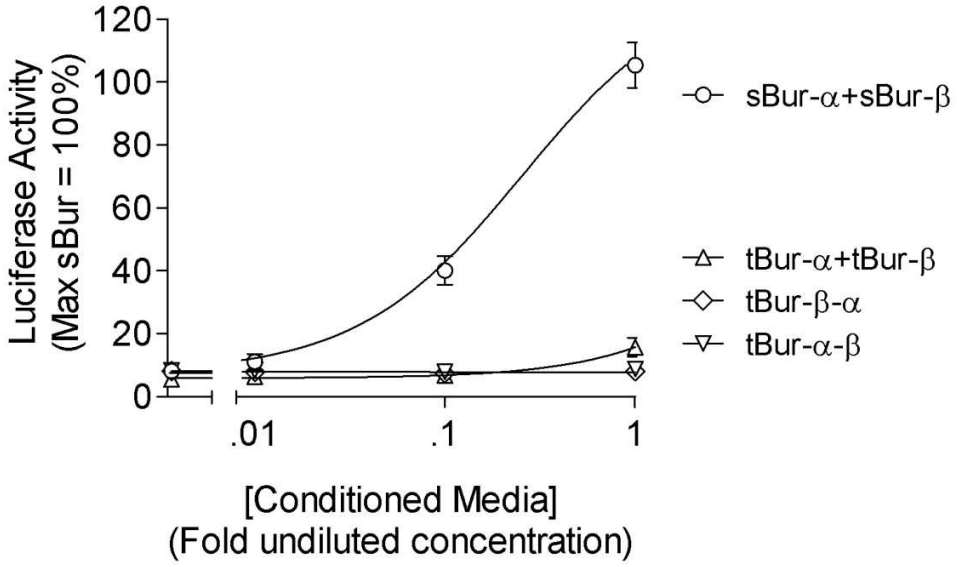


Figure 1.9. Active conditioned media cannot be made from expression of tethered bursicon constructs.

Activity of media isolated from cells expressing indicated bursicon MTLs. Method: Conditioned media was made from cells transfected in 6-well dishes with 75ng/well of each indicated tethered or soluble bursicon cDNA construct. Twenty four hours after transfection, the dishes were aspirated and 1.2mL of serum free DME was added. Media was then conditioned for 48 hours and prepared for signaling assays as described in Materials and Methods. The conditioned media was transferred to cells transfected with cDNAs encoding rk, a 6X-CRE-luc reporter gene, and a β-galactosidase control gene. After 4 hours, luciferase activity was quantified. The activity values were normalized relative to maximal stimulation of rk with bursicon conditioned media made by co-expression of soluble α and β subunits. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations. sBur = soluble Bursicon subunit cDNA, tBur= bursicon MTL subunit cDNA, CM=conditioned media.

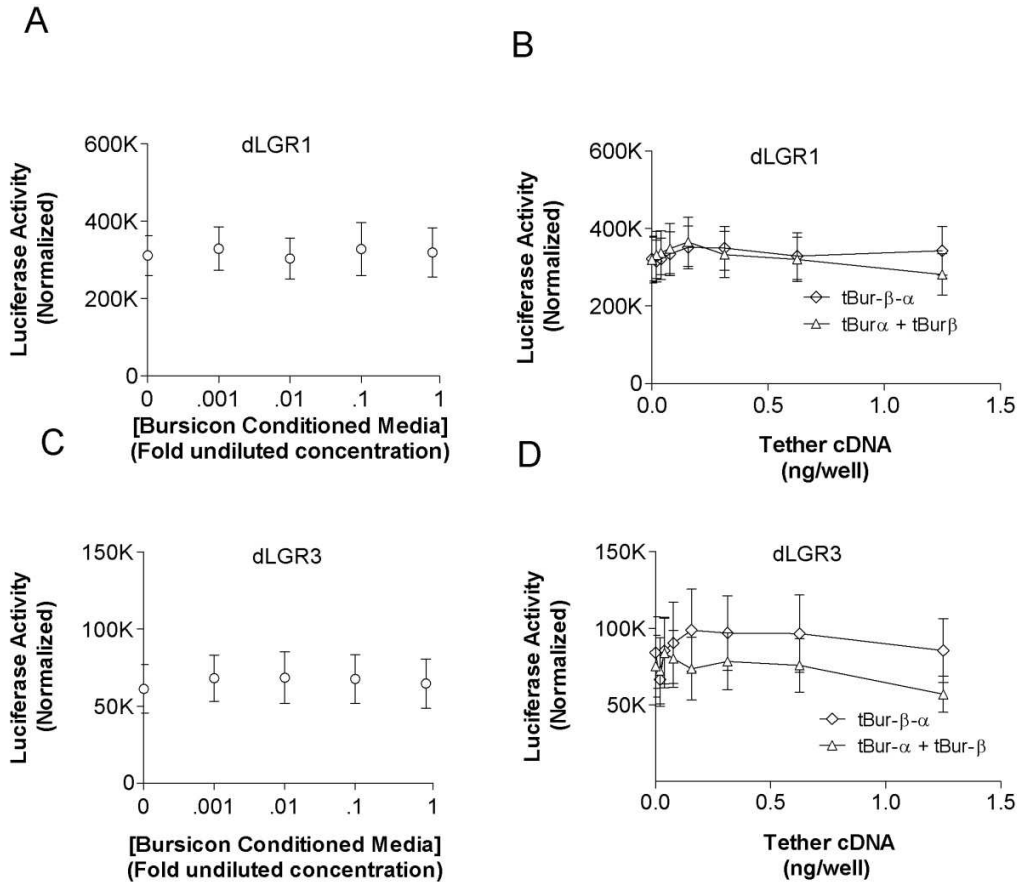


Figure 1.10. dLGR1and dLGR3 are not activated by soluble or tethered bursicon.

A) dLGR1 is not activated by bursicon conditioned media B) Bursicon MTLs do not activate dLGR1. C) dLGR3 is not activated by bursicon conditioned media. D) Bursicon MTLs do not activate dLGR3. Method: HEK293 cells were transiently co-transfected with cDNAs encoding dLGR1 or dLGR3, a 6X-CRE-Luc reporter gene, and a β-galactosidase control gene. For assessment of tethered construct activity, cDNAs encoding bursicon MTL constructs were also co-transfected as indicated. The x-axes denote the amount transfected for each cDNA subunit (B, D). The cells were transfected for twenty hours, followed by stimulation with bursicon conditioned media for 4 hours if no MTLs were co-expressed (A, C). Luciferase activity was then determined and corrected for variability using the β-galactosidase control gene. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations. tBur= bursicon MTL subunit cDNA, K=1000

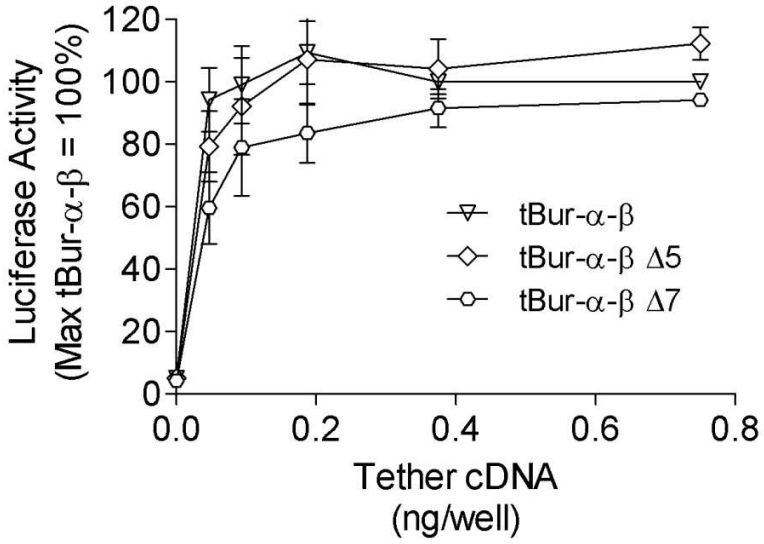


Figure 1.11. Agonist of tBur-α-β is maintained despite serial C-terminal deletions .

Assessment of tBur-α-β C-terminal deletion agonist activity. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, the indicated bursicon MTL, and a β-galactosidase control gene. Twenty four hours after transfection, luciferase activity was quantified and normalized relative to maximal stimulation of rk by full length tBur-α-β. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tBur= bursicon MTL subunit cDNA.

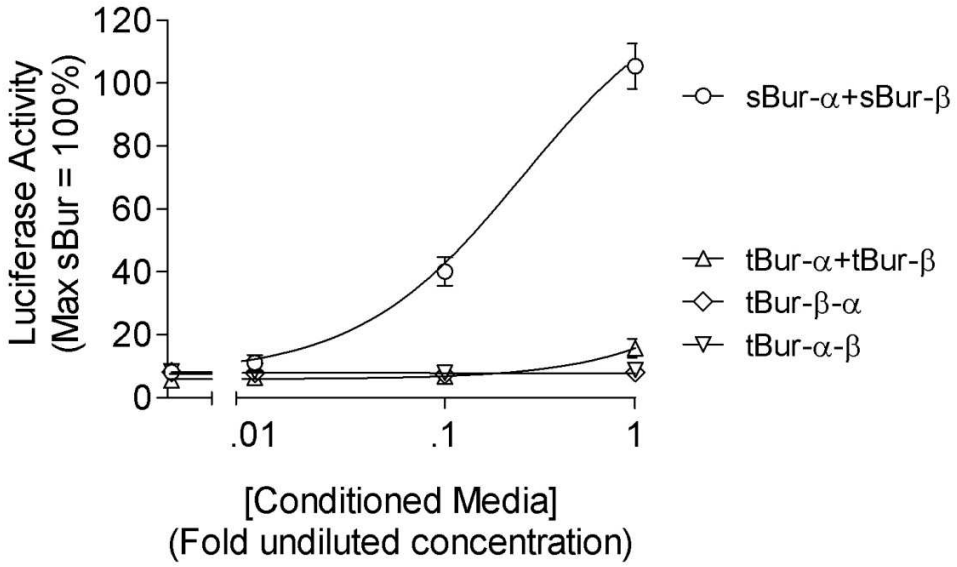


Figure 1.12. Soluble Chimeric Bursicon conditioned media activates rk.

Concentration dependent activation of rk with bursicon and soluble chimeric bursicon conditioned media. A series of tenfold dilutions of conditioned media (1=undiluted conditioned media) was added to cells 20 hours after transfection; the duration of ligand stimulation was 4 hours. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: sBurβ-α= genetically fused α and β bursicon subunit cDNA

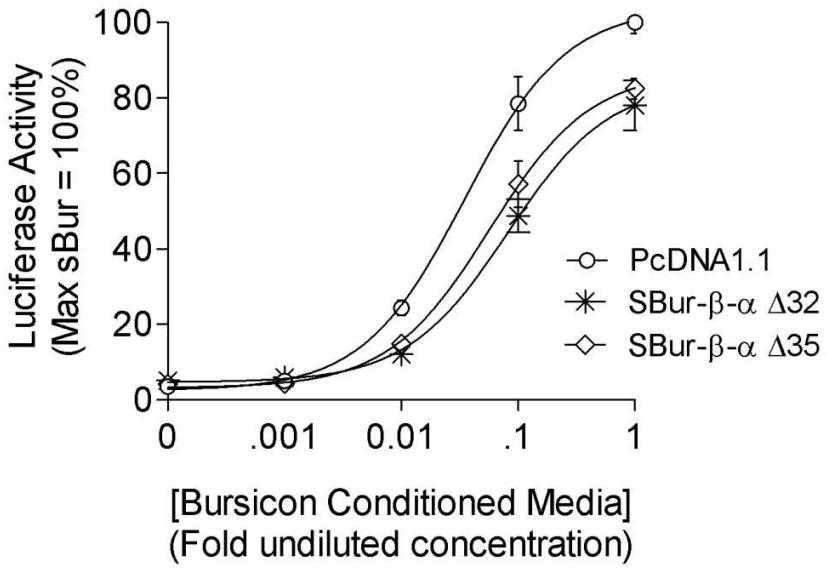


Figure 1.13. Soluble Chimeric Bursicon deletion mutants are not potent rk antagonists.

Co-expression of rk and indicated soluble chimeric bursicon C-terminal deletion constructs do not block receptor activation by soluble bursicon. HEK293 cells were transiently transfected with 2ng of the indicated deletion construct, rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase control gene. Twenty hours following transfection bursicon conditioned media was added at ten-fold serial dilutions (1=undiluted conditioned media). Following a four-hour incubation with bursicon conditioned media, luciferase activity was quantified and normalized relative to maximal stimulation of rk by bursicon conditioned media in the absence of a potential inhibitor. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: sBurβ-α= genetically fused α and β bursicon subunit cDNA

Chapter 2:

Membrane tethered bursicon as a tool to probe the role of the *Drosophila* GPCR rickets *in vivo*

Abstract

Rickets (rk) is a G protein-coupled receptor (GPCR) that modulates wing expansion, cuticle hardening, and melanization in *Drosophila melanogaster*. The endogenous ligand for rk is the heterodimeric cystine-knot protein bursicon comprised of an alpha and beta subunit. We have developed a membrane tethered bursicon (CFP-tBur-β-α) construct that recognizes its cognate receptor and enables temporal/spatial modulation of rk mediated physiology in transgenic flies. Ubiquitous expression of CFP-tBur-β-α in flies results in developmental arrest at the pupal stage. Organisms that escape lethality do not expand their wings. These features suggest that expression of the tethered construct results in a decrease in rk mediated signaling. Consistent with this hypothesis, we show in complementary *in vitro* studies, that sustained stimulation of rk by bursicon, soluble or tethered, leads to receptor desensitization. Targeted expression of CFP-tBur-β-α, in specific tissues enabled us to unravel an important role for rk in muscle, specifically at the postsynaptic membrane of the neuromuscular junction. In addition, we show that rk, expressed in a subset of adult muscles, facilitates the emergence of pharate adults from the pupal case. In this study, we have developed a novel membrane tethered ligand (CFP-tBur-β-α) that negatively regulates rk function when expressed *in vivo*. Using this novel tool, we have identified adult muscle as a tissue that requires rk expression for survival through eclosion and wing expansion in *Drosophila*.

Introduction

Bursicon is a heterodimeric cystine-knot protein required for wing expansion and cuticle hardening in a variety of insects. Many of the initial studies on bursicon as well as the discovery of its receptor have been conducted in *Drosophila melanogaster*. Its cognate receptor is a leucine rich repeat-containing G protein-coupled receptor (GPCR) known as rickets (rk, dLGR2) (Luo, Dewey et al. 2005). Previously we have shown that a membrane anchored single subunit fusion construct of the bursicon heterodimer (CFP-tBur-β-α) can activate rk *in vitro* in a concentration dependent manner (Harwood, Fortin et al. 2013). Membrane tethered ligands (MTLs) are cDNA constructs that express genetically encoded peptide hormones anchored to the cell membrane via a transmembrane domain. We have developed a series of MTLs capable of selectively activating either insect or mammalian GPCRs (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Fortin, Chinnapen et al. 2011).

A major advantage of using MTL technology *in vivo* is that it enables selective activation of receptors in a given tissue without the confounding effects of soluble ligand diffusion. The transgenic Gal4/UAS system in conjunction with membrane tethered bursicon offer an excellent model to better understand the underlying tissue dependence of rk mediated signaling. In *Drosophila*, wing expansion occurs within 1 hour following eclosion. A tightly choreographed motor program is required for wing expansion and cuticle hardening. This series of events coincides with a biphasic release of bursicon, first from the subesphogeal ganglion, followed by release from the abdominal ganglion. This allows dispersion of bursicon into the insect circulatory system via the hemolyph

(Peabody, Diao et al. 2008; Peabody, Pohl et al. 2009). While most studies have focused on the location of bursicon release, much less research has addressed the importance of tissue selective rk activation (Honegger, Dewey et al. 2008).

Previously it was thought that rk was only required following eclosion, as a trigger for wing expansion and cuticle hardening. This was based on two fly stocks known as Rk^1 and Rk^4. Originally identified in a mutagenesis screen, these fly lines encode point mutations that introduce a premature stop codon in the rk receptor coding sequence resulting in what was thought to be a truncated and nonfunctional receptor. Both rk mutants are homozygous viable but do not expand their wings and have defects in cuticle hardening and tanning. However these mutants have been somewhat misleading due to the fact that they are actually hypomorphs rather than true nulls (Baker and Truman 2002; Loveall and Deitcher 2010). Studies using rk RNAi *in vivo*, suggest that global knockdown of rk results in developmental arrest, rather than just impairing wing expansion, melanization, and cuticle hardening. A recent study also showed that deletion of the bursicon β subunit resulted in significant lethality during pupariation, specifically throughout ecdysis (Loveall and Deitcher 2010; Lahr, Dean et al. 2012). Very few studies have specifically looked at what tissues require rk expression and activation for proper development.

In normal *Drosophila*, rk is expressed in the epidermis. Activation by bursicon promotes cuticle hardening (Diao and White 2012). More detailed analysis of rk expression as revealed by FlyAtlas and by multiple microarray studies have shown that this receptor is expressed at low levels throughout development (Chintapalli, Wang et al.

2007). However, data about which tissues or cells require rk expression and subsequent activation for proper development remain unclear.

Use of the bursicon MTL, CFP-tBur-β-α, has offered a novel approach to investigate the tissue specific requirements of rk. Expression of the fused heterodimer offers a way to selectively modulate rk without the confounding effects of soluble ligand diffusion. Using tissue specific Gal4 drivers to express CFP-tBur-β-α provides a tool to examine the tissue specific effects of rk. Parallel studies using rk RNAi transgenic flies enables a complementary approach to confirm conclusions drawn through the use of membrane tethered ligands. Rk is an excellent receptor to use as a model system for the study of MTLs as multiple phenotypes (e.g, survival, wing expansion, and melanization) can be easily be monitored.

Prior studies by our group have shown that a membrane tethered agonist can trigger long term receptor activation in flies (Choi, Fortin et al. 2009). In the current study, we illustrate that there is an alternative potential consequence resulting from membrane tethered ligand expression *in vivo*, i.e. receptor desensitization. Chronic GPCR stimulation with soluble agonists often results in receptor desensitization and can occur through a variety of cellular mechanisms (Kang, Tian et al. 2013). We have shown in this work that bursicon either as a soluble ligand or an MTL triggers similar desensitization. As a result, long term membrane tethered bursicon expression *in vivo* functionally inactivates the receptor.

In this manuscript we show that CFP-tBur-β-α expression *in vivo* results in developmental lethality in a pattern that is similar to rk RNAi mediated receptor knockdown. An *in vitro* model of chronic MTL stimulation suggests that rk is

desensitized with long term stimulation leading to functional blockade of the receptor. Finally by expressing membrane tethered bursicon and knockdown of rk with a collection of increasingly focused tissue specific Gal4 drivers, we demonstrate that rk expression is required in adult muscle.

Results

Previously, using a membrane tethered bursicon construct (CFP-tBur-β-α) we have shown that a single subunit fusion of both bursicon subunits can activate rk *in vitro*. As a follow-up to this project we generated transgenic flies expressing CFP-tBur-β-α under the control of an upstream activating sequence (UAS). These flies were used in combination with selected Gal4 *Drosophila* driver lines to express the bursicon MTL in a tissue specific manner. CFP-tBur-β-α were first crossed to the ubiquitous Gal4 driver Actin-5C. Unexpectedly, ubiquitous expression of CFP-tBur-β-α resulted in a significant decrease in survival through eclosion. Developmental lethality was also observed upon ubiquitous knockdown of the rk receptor, but not from knock down of the bursicon ligand (bursicon α subunit) as seen in Figure 2.1A. In addition to developmental lethality, flies that ubiquitously expressed CFP-tBur-β-α that survived through eclosion (escapers) failed to expand their wings 100% of the time. In contrast, with Bur-α knockdown, approximately 50% of flies failed to properly expand their wings. We observed no escapers with ubiquitous rk knockdown (Figure 2.1B). As a confirmation, CFP-tBur-β-α flies were crossed with a novel rk specific driver, Rk^{PAN}-Gal4 (Diao and White 2012). With the rk specific driver significant amounts of developmental lethality occurred and escapers were unable to expand their wings (Figure 2.6). Unfortunately given the way in which this Gal4 line was created, RNAi mediated knockdown of the receptor is not possible. The Rk^{PAN}-Gal4 line overexpresses rk transcript complementary to the RNAi sequence expressed by each of the available rk RNAi fly lines.

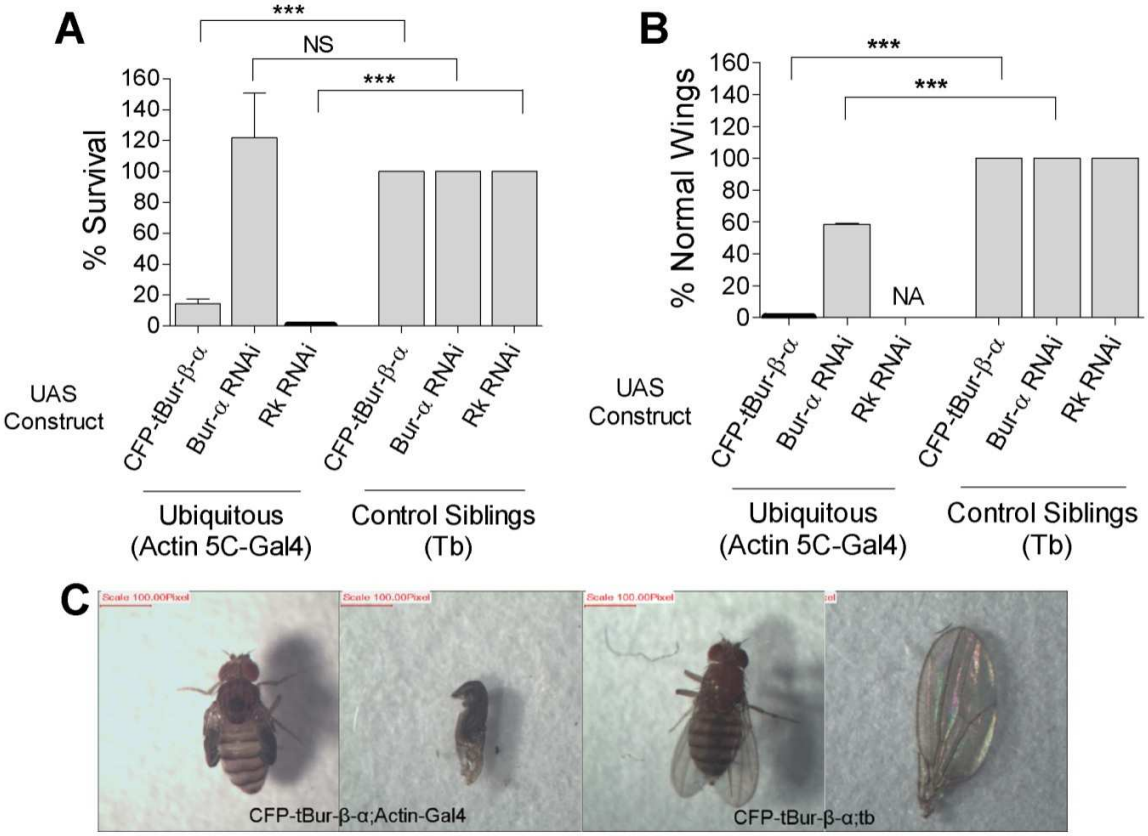


Figure 2.1. Ubiquitous overexpression of membrane tethered bursicon recapitulates the phenotypes induced by downregulation of its cognate receptor, rk.
A) Survival through eclosion of indicated progeny. B) Percentage of flies with normal expanded wings for indicated progeny C) Representative images of wings of eclosed flies and corresponding wings from indicated genotypes. Abbreviations: NA= Wing expansion data not available due to 100% developmental lethality, UAS= upstream activation sequence, Tb=Tubby[1], CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon. Statistics: Data represent mean ± SEM of experiments done in triplicate. N≥50-134 flies were counted for each independent cross. ***=P<.0001, NS=P>0.05,

We were somewhat surprised that the observed phenotype with ubiquitous expression of CFP-tBur-β-α (an agonist) recapitulated results of the rk knockdown. One possible explanation was that chronic activation of rk during development led to receptor desensitization, the receptor was thus functionally inactivated. To determine whether rk was desensitized, a previously described *in vitro* model of activation (Harwood, Fortin et al. 2013) was adapted to look at the effects of chronic stimulation of rk in cell culture. Interestingly, rk is efficiently desensitized following chronic stimulation with bursicon conditioned media. After an overnight incubation, rk is no longer able to respond to a second 4 hour pulse of bursicon conditioned media (Figure 2.2A). Desensitization with bursicon appears to be rk specific. Following an overnight stimulation with bursicon conditioned media, isoproterenol was added on to the cells to stimulate β2 Adrenergic receptors (β2AR) which are endogenously expressed on HEK293 cells. β2AR activation was observed suggesting that rk desensitization is homologous rather than heterologous (Figure 2.2A).

In addition to rk desensitization by soluble bursicon, expression of CFP-tBur-β-α triggers a similar pharmacological consequence. This effect is dependent on the amount of cDNA transfected. Desensitization increases in parallel with tether expression. Importantly, over-expression of the bursicon α cDNA subunit alone does not result in receptor desensitization, suggesting that the receptor must be activated by the full heterodimer (α plus β) for desensitization to occur (Figure 2.2B).

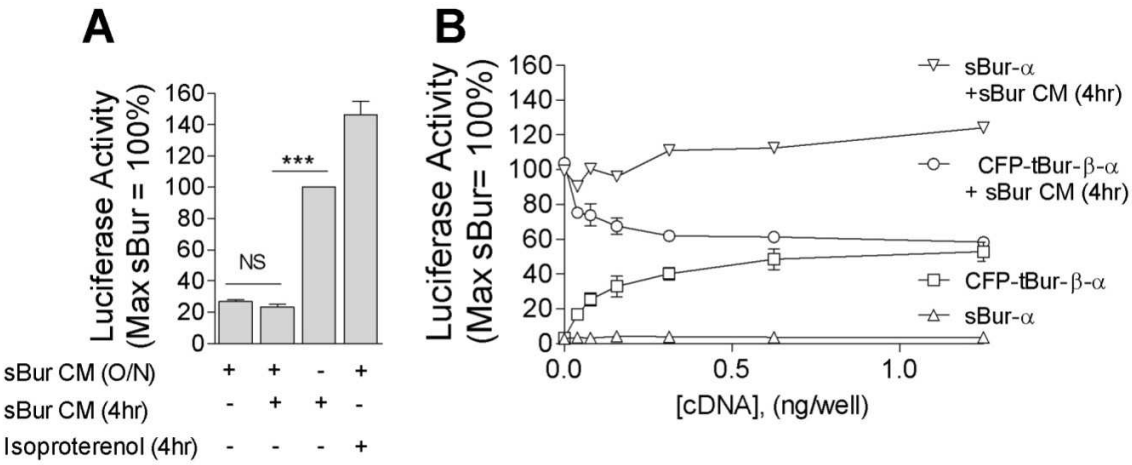


Figure 2.2. Chronic *in vitro* stimulation results in rk desensitization. A) Overnight stimulation with bursicon conditioned media desensitizes rk receptors *in vitro*. HEK293 cells were transfected with rk, Pgl4.22 (a CRE luciferase reporter construct), and a β-galactosidase control gene. Twenty four hours after transfection, cells were treated overnight (O/N) with bursicon conditioned media. The next day, the cells were again stimulated with bursicon conditioned media or isoproterenol for an additional 4 hours. Activity was compared to that of a 4 hour stimulation of rk expressing cells without an overnight bursicon conditioned media pre-treatment. B). CFP-tBur-β-α desensitizes rk receptor activation in a cDNA concentration dependent manner. HEK293 cells were transfected as described for panel A, but with the addition of increasing amounts of the indicated cDNAs. Forty four hours following transfection, cells were further stimulated for 4 hours with a maximal concentration of bursicon conditioned media or serum-free DMEM. Luciferase activity was then quantified. Abbreviations: sBur CM = soluble bursicon conditioned media, sBur-α = soluble bursicon α subunit cDNA, CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon. Statistics: Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. ***= $P<0.0001$, NS= P value>0.05

Based on the above data, it became evident that expression of CFP-tBur-β-α could be used as a tool to selectively inactivate rk in potential tissues of interest. A general screen of Gal4 drivers was conducted to determine what tissues are critical for rk dependent survival (Table 2.1). We found that the muscle specific driver, Held out wings (HOW-Gal4), led to a significant decrease in survival. HOW-Gal4 is expressed in both larval and adult muscle. Notably, expression of either CFP-tBur-β-α or rk RNAi resulted in significant amounts of developmental lethality. CFP-tBur-β-α expression with HOW-Gal4 resulted in a 100% penetrant wing expansion defect for escapers. In contrast, knockdown of Bur-α with How-Gal4 did not cause any lethality or wing expansion defects (Figure 2.3A-B). This suggests that the bursicon tethered ligand is acting on muscle cells that express rk but do not release bursicon.

These observation raised the question of whether bursicon was locally released from cells adjacent to rk expressing muscle cell (e.g. at the neuromuscular junction, NMJ). To address the role of the NMJ, CFP-tBur-β-α, rk RNAi, and Bur-α RNAi constructs were expressed in motor neurons using the D42-Gal4 driver. With CFP-tBur-β-α expression in motor neurons, a small decrease in survival was observed. In addition wing expansion defects were seen when CFP-tBur-β-α and Bur-α RNAi constructs were expressed in motor neurons. However no wing expansion defects or lethality was observed when rk was knocked down in motor neurons (Figure 2.3C-E). The combined results of D42-Gal4 (motor neuron driver) and HOW-Gal4 (muscle driver) suggests that rk is likely expressed in muscle and at least some bursicon is released from the NMJ.

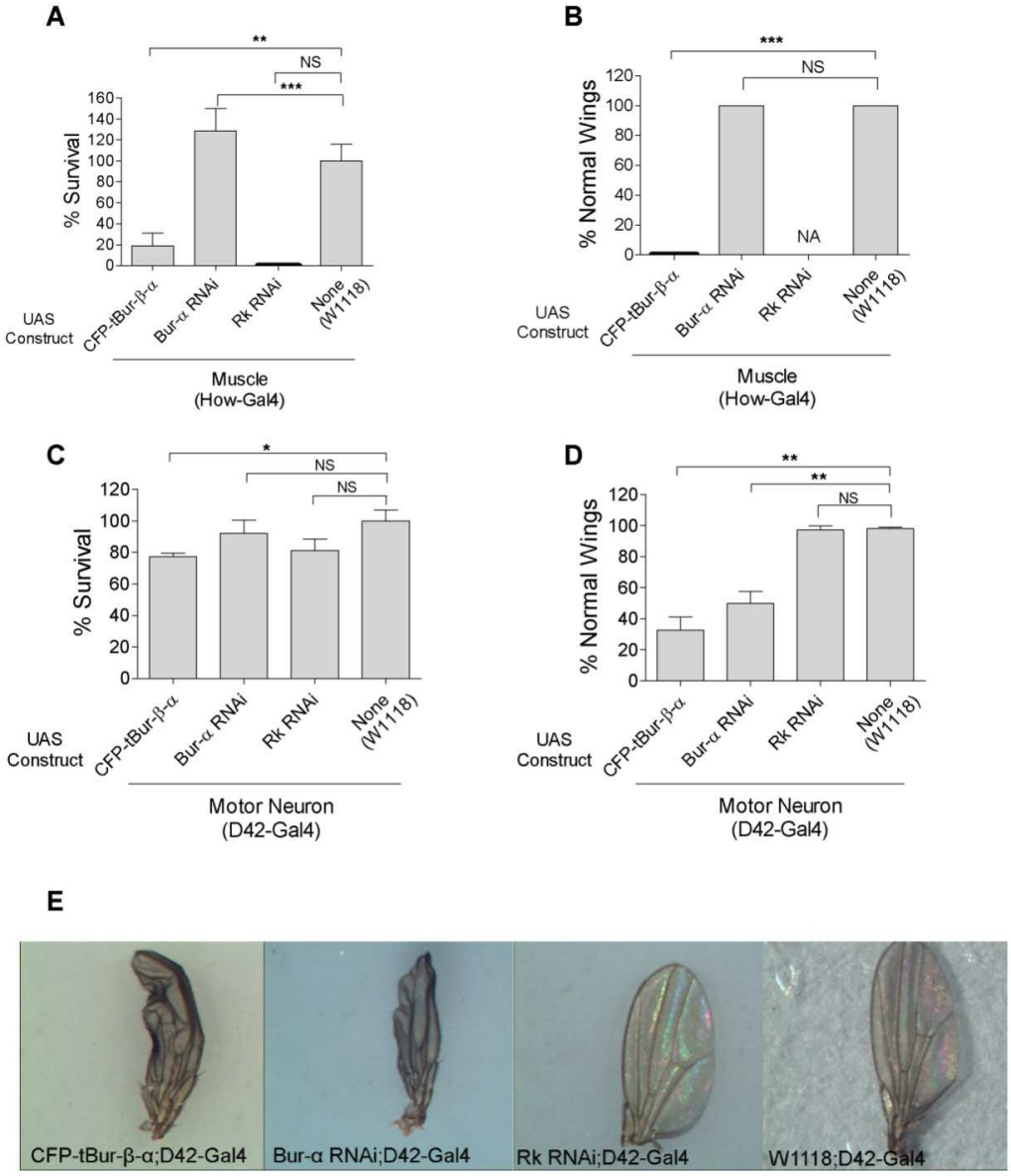


Figure 2.3. Expression of membrane tethered bursicon in muscle and motor neurons results in wing expansion defects.

Expression of CFP-tBur-β-α and rk RNAi, but not Bur-α RNAi in muscle results in a reduction in survival through eclosion (A) and wing expansion defects (B). Expression of CFP-tBur-β-α and Bur-α RNAi, but not rk RNAi in motor neurons results in a reduction in survival through eclosion (C) and wing expansion defects (D). E) Representative images of wings following eclosion from indicated genotypes. Abbreviations: NA= Wing expansion data not available due to 100% developmental lethality, UAS= upstream activation sequence, CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon. How = Held out wing. Statistics: Data represent mean ± SEM of experiments done in triplicate. 85-172 progeny were counted for each independent control cross. ***=P<0.0001, **=P<0.001, *=P<.01, NS=P>0.05,

As a next step we investigated at what stage of muscle development, rk activation is required for both survival and wing expansion. To address this question, a series of mesoderm specific Gal4 lines were tested. The first used was Apterous Gal4 (Ap-Gal4). Corresponding fly lines showed significant developmental lethality at the pupal stage with either CFP-tBur-β-α expression or with RNAi mediated rk knockdown. Notably, all flies that survived through development had wing expansion defects (Figure 2.4A). A second mesoderm specific driver Mef2-Gal4 also showed decreased survival and wing expansion defects both with expression of CFP-tBur-β-α or rk RNAi. Although significant, the phenotypes were less severe than with the Ap-Gal4 and How-Gal4 drivers (Figure 2.4B-D).

Finally we wanted to determine whether rk plays a demonstrable role in adult muscle. As the focus of these efforts, we targeted adult muscle, using the Actin88F-Gal4 driver line. Using this driver line, with expression of either CFP-tBur-β-α or rk RNAi, significant lethality was observed prior to eclosion. We also observed wing expansion defects in all escapers. Again no decrease in survival or wing expansion defects were observed in parallel experiments expressing a Bur-α subunit RNAi construct (Figure 2.5). Taken together these results suggest that rk activation is required in developing muscle and formed adult muscles. Rk mediated activity on muscle thus appears to be important both for proper eclosion and wing expansion.

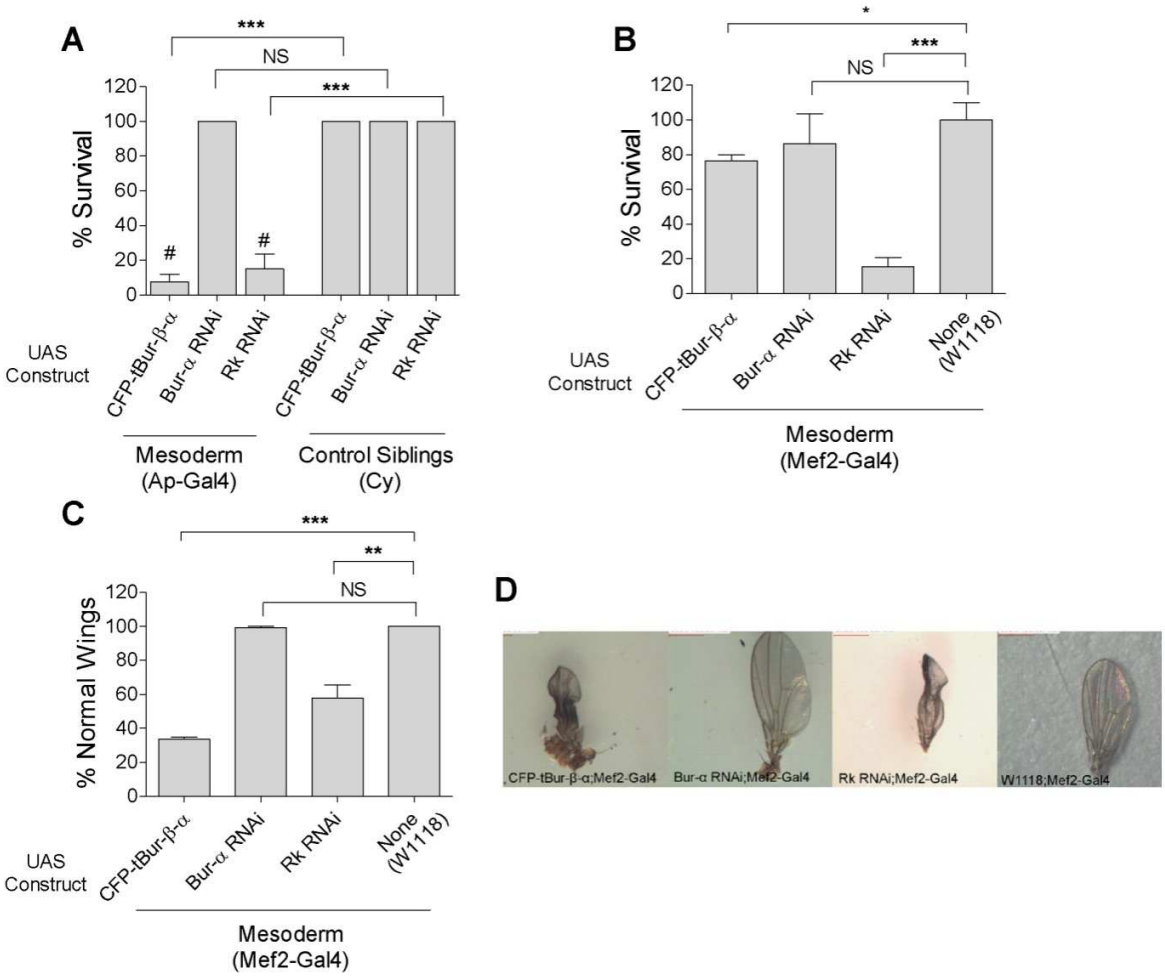


Figure 2.4. CFP-tBur-β-α expression or rk knockdown in the mesoderm causes developmental lethality and wing expansion defects A) CFP-tBur-β-α and rk RNAi expression with Apteris-Gal4 results in reduced survival through eclosion. Escapers are unable to expand wings. B) Expression of CFP-tBur-β-α and rk RNAi, but not Bur-α RNAi in Mef2 positive mesoderm results in a reduction in survival through eclosion (B) and wing expansion defects of escapers(C). D) Representative images of wings following eclosion from indicated genotypes. Abbreviations: UAS= upstream activation sequence, CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon, Ap=Apterous, Mef2= Myocyte enhancer factor-2, Cy= Curly, # = 0% normal wing expansion. Statistics: Data represent mean ± SEM of experiments done in triplicate. 40-112 progeny were counted for each independent control cross. ***=P<0.0001, **=P<0.001, *=P<.01, NS=P>0.05.

Discussion

In this study we have utilized a membrane tethered bursicon construct (CFP-tBur-β-α) to probe the role of rk in development. Previous studies from our lab have shown that when assessed *in vitro*; CFP-tBur-β-α is an rk agonist (Harwood, Fortin et al. 2013). Initial *in vivo* investigations revealed that expression of CFP-tBur-β-α in otherwise rk/bursicon wildtype *Drosophila* (W^{1118}) led to lethality and wing expansion defects, both unanticipated phenotypes. Originally, we had predicted that the agonist activity of CFP-tBur-β-α would provide a gain of function and would be useful to rescue bursicon mutant fly lines. Contrary to this expectation, the phenotypes resulting from ubiquitous expression of CFP-tBur-β-α were loss of function (i.e. lethality and wing expansion defects) and resembled those observed with knockdown of rk using RNAi (Figure 2.1). Flies that expressed CFP-tBur-β-α or rk RNAi ubiquitously generally reached pupal development however failed to eclose. Dead flies were fully developed but did not emerge from their pupal cases (data not shown). Based on the parallel phenotypes with the RNAi flies, we hypothesized that expression of CFP-tBur-β-α resulted in a decrease of rk mediated signaling.

We then investigated how expression of CFP-tBur-β-α constructs reduces/eliminates rk mediated signaling, thus mimicking the rk RNAi phenotypes. We postulated that decreased or absent rk mediated signaling could result from chronic stimulation leading to receptor desensitization. This pharmacological consequence may alter the ability of rk to elicit an appropriate response to endogenous bursicon at key developmental stages. To examine this possibility, desensitization of rk was tested in an *in vitro* model comparing CFP-tBur-β-α and soluble bursicon conditioned media. The data suggests that activation

of rk with long term (overnight) stimulation renders the receptor unable to further respond to bursicon (Figure 2.2). The desensitization of rk is receptor specific. HEK293 cells expressing recombinant rk that are desensitized to bursicon, still signal in response to treatment with a β2AR agonist. The β2AR is endogenously expressed and also signals through Gs_{α}. These data suggest that the observed desensitization was not heterologous and specific to rk. Desensitization has previously been documented among other LGR receptors. The human intestinal GPCR LGR5, which is the most closely related human receptor to rk, is known to be constitutively internalized (Snyder, Rochelle et al. 2013). Notably several C-terminal serine residues in LGR5 have been postulated to play an important role in desensitization (Snyder, Rochelle et al. 2013). The C-terminus of rk is also highly enriched in serine residues and could potentially be important through a conserved mechanism. Taken together, we consider tethered bursicon mediated desensitization of rk a likely explanation for the observed loss of function phenotypes *in vivo*. The ability to selectively desensitize rk in a tissue and/or temporal specific manner provided an important complementary tool for studying the rickets bursicon system.

We have shown that ubiquitous expression of CFP-tBur- β-α leads to developmental arrest at the pupal stage. Escapers that survive to adulthood show characteristic features of rk knockdown. CFP-tBur- β-α induces phenotypes that are readily monitored, and provides a useful tool to further dissect the role of rk in the developing fly (complementing rk RNAi constructs). We subsequently used a series of Gal4 driver lines that direct expression in selected tissues to identify the cell type(s) that underlies the lethality/wing expansion phenotypes (Table 2.1).

Table 2.1. CFP-tBur- β-α screen with indicated Gal4 driver lines

Gal4 Driver	Reduced Survival	Wing Expansion Defect	Expression Pattern
Actin5C-Gal4	+	+	Ubiquitous
CCAP-Gal4	-	+	CCAP neurons
D42-Gal4	-	+	Motor neurons
Twist-Gal4	-	-	Stem cells, Adult muscle precursors
HOW-Gal4	+	+	Pan mesoderm, tendon cells
Ap-Gal4	+	+	Direct flight muscle, tendon cells, leg and abdominal muscle, neurons
Mef2-Gal4	+	+	Differentiating adult muscle, neurons
Actin88F-Gal4	+	+	Indirect Flight Muscle, leg muscle, abdominal muscle
Tinman-Gal4	-	-	Cardiac muscle
Striped-Gal4	-	-	Tendon cells

The vast majority of drivers resulted in normal fly development. In contrast, expression of CFP-tBur- β-α restricted to muscle under the control of the pan-mesodermal driver HOW-Gal4 (Zaffran, Astier et al. 1997), resulted in defects that phenocopied those seen with ubiquitous expression of the tethered ligand (i.e. lethality, wing phenotypes). Consistent with results from our studies on receptor desensitization (detailed above), similar phenotypes were also observed when rk was downregulated in the same tissue (using rk RNAi). Although many studies have focused on the CNS to probe the spatial requirements of bursicon-mediated effects (Peabody, Diao et al. 2008; Peabody, Pohl et al. 2009; Lahr, Dean et al. 2012), the localization of the bursicon receptors that are essential for survival, has remained elusive. Our study strongly supports that peripheral rk in muscle plays an important role in *Drosophila* development. This observation is consistent with a previous report indicating that bursicon alpha is released at the neuromuscular junction, prior to ecdysis (Loveall and Deitcher 2010).

By design, CFP-tBur- β-α anchors in the membrane and projects into the extracellular space (Harwood, Fortin et al. 2013). Its function is defined by four structural elements: transmembrane domain, linker, epitope tag, and peptide. Analysis using both CFP-tBur-β-α and relevant RNAi constructs suggests the receptor is localized only on the postsynaptic (muscle side). Expression of either CFP-tBur- β-α, or rk RNAi, in muscle leads to comparable lethality/wing phenotypes. In contrast, expression of rk RNAi in motor neurons (using the D42-Gal4 driver) has no effect, suggesting the absence of a presynaptic receptor. Expression of CFP-tBur- β-α using the same driver leads to a wing defect phenotype. We postulate that tethered bursicon anchored in the motor neuron acts on muscle in a trans effect which leads to desensitization of rk and results in the wing

expansion phenotype. Taken together these results suggest an important role for rk at the postsynaptic membrane of the NMJ.

In addition to defining a role for rk at the postsynaptic compartment of muscles, we began to define the role of rk during myogenesis. We used selected Gal4 driver lines to explore when, during muscle formation, rk signaling is required for normal adult fly development (Table 2.1). In agreement with previous rk RNAi studies (Loveall and Deitcher 2010) and Bur-β null fly studies (Lahr, Dean et al. 2012), we observe developmental arrest at the time of pupal ecdysis when using CFP-tBur- β-α. It is well-established that during metamorphosis, most of the adult muscles are formed de novo from progenitor cells, i.e. adult muscle precursors (AMPs) located on the larval wing imaginal disc and leg imaginal disc (Raghavan, Gendre et al. 1996). Undifferentiated AMPs express high levels of the helix-loop-helix transcription factor, twist (Bate, Rushton et al. 1991; Currie and Bate 1991). Myoblast differentiation /commitment to the muscle lineage require downregulation of twist, and the interplay of other transcription factors. This core regulatory network includes Tinman (cardiac muscle specification), Myocyte Enhancer Factor-2, Mef2 (which plays a key role in myoblast fusion/ formation of somatic muscles) and Apterous (which specifies selected subtypes of skeletal muscles) (Olson, Perry et al. 1995; Ranganayakulu, Zhao et al. 1995; Yin, Xu et al. 1997; Cripps, Black et al. 1998). The results of our genetic analysis using the corresponding drivers with CFP-tBur- β-α and rk RNAi is summarized in Table 2.1. We show that expression of rk in differentiating myofibers, but not in adult muscle precursors (i.e. twist positive cells), is important to ensure normal fly development. In addition we show that the heart is not likely involved in the burs/rk mediated control of eclosion/wing expansion.

We then investigated whether expression of rk in selected subtypes of pharate adult muscles could account for the observed CFP-tBur- β-α/rk RNAi induced phenotypes. To carry out this analysis, the Act88F-Gal4 driver line proved useful. Expression of CFP-tBur- β-α and knockdown of rk with Act88F-Gal4 results in developmental lethality, prior to eclosion (Figure 2.5A). Escaper flies fail to expand their wings (Figure 2.5B). Comparison of the tissues that are targeted by the Apterous Gal4 (as outlined above) and Act88F Gal4 enabled us to exclude selected muscle subtypes potentially underlying the rk mediated phenotypes. Actin 88F is a muscle actin predominantly expressed in the thoracic indirect flight muscles (IFMs). Notably, Apterous is absent in the IFMs. Actin88F is also expressed to a lesser extent in the mesothoracic leg (tibial depressor muscles), as well as in the abdominal muscles (both ventral and dorsal) (Ghazi, Anant et al. 2000; Nongthomba, Pasalodos-Sanchez et al. 2001). Although Apterous expression in the leg and/or abdominal muscles needs to be confirmed, these two muscle types emerge as potentially important for rk signaling-regulation of eclosion.

Both leg and abdominal muscles are important during metamorphosis. At 12 hours after puparium formation, contraction of the abdominal muscles forces an air bubble forward which in turn, triggers head eversion (Fernandes, Bate et al. 1991). In addition, at the end of the pupal stage, the newly formed adult uses its legs to free itself from the case. The potential role of rk in leg muscle could also explain the namesake kinked leg phenotype observed with rk classical mutants (Baker, Tanaka et al. 2005; Loveall and Deitcher 2010).

In conclusion, we have developed a novel membrane tethered ligand (CFP-tBur-β-α) that negatively regulates rk function when expressed *in vivo*. Using this novel tool in

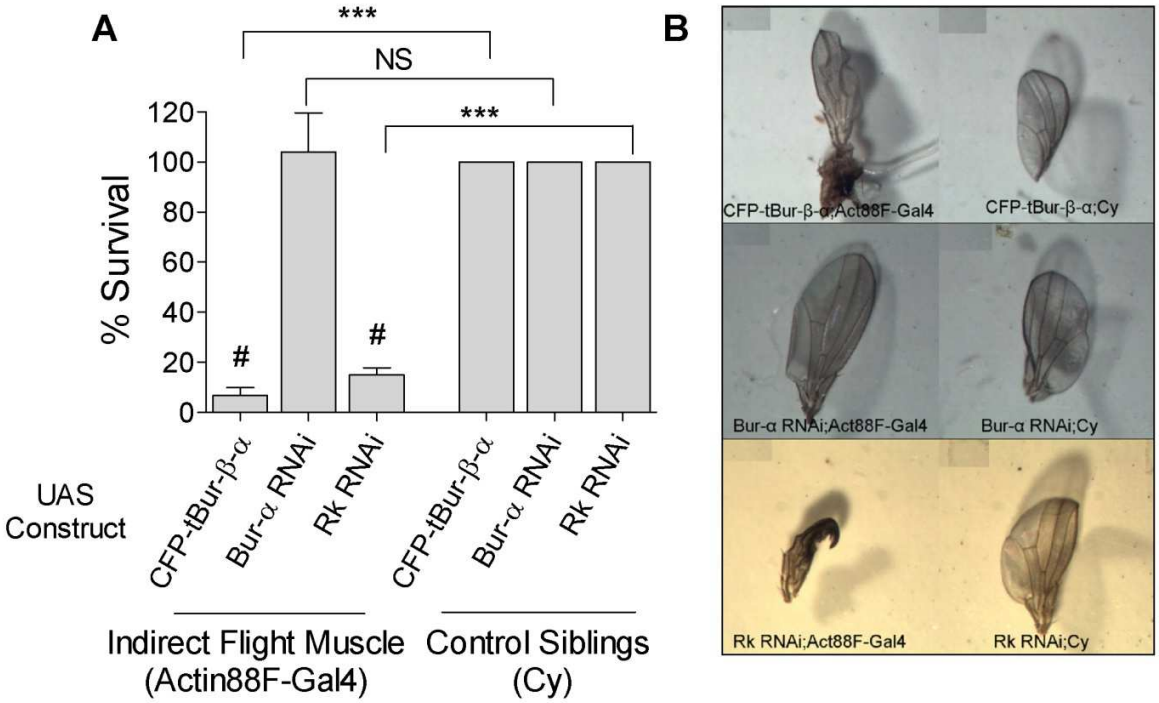


Figure 2.5. Rk knockdown and CFP-tBur-β-α expression in adult indirect flight muscle causes lethality and wing expansion defects.
A) Survival through eclosion is reduced for progeny expressing CFP-tBur-β-α and rk RNAi, but not Bur-α RNAi in adult indirect flight muscle. B) Representative images of wings from eclosed flies for indicated genotypes. Abbreviations: UAS= upstream activation sequence, CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon, Cy= Curly, #= 0% normal wing expansion. Statistics: Data represent mean ± SEM of experiments done in triplicate. N≥58-191 flies were counted for each independent cross. ***=P<.0001, NS=P>0.05.

conjunction with existing RNAi fly lines we have identified adult muscle as a tissue that requires rk expression for survival and wing expansion in *Drosophila*. In particular, leg and abdominal muscles or a subset of these could be important. Future studies will be required to further understand the role of muscle rk in regulating fly development.

Materials and Methods

Cell Culture

Human embryonic kidney cells (HEK293) were cultured in Dulbecco's modified eagle medium (DMEM, Life Technologies, Grand Island, NY) with 10% Fetal Bovine Serum (FBS, Atlanta Biologicals, Lawrenceville GA), 100U/mL penicillin, and 100µg/mL streptomycin (Life Technologies, Grand Island, NY). Cells were maintained at 37 °C in a humidified 5% CO2 atmosphere.

Luciferase Assays

Luciferase assays were done as previously described with minor modifications (Harwood, Fortin et al. 2013). In brief HEK293 cells in 96 well plates were transfected in serum free DME with antibiotics and PEI (1µg/ml) for 48 hours. Cells were transfected with 1ng/well rk receptor, 5ng/well CRE-LUC-HCL-PEST luciferase reporter, and 5ng/well β-galactosidase plasmid as a transfection control. In addition cells were transfected with various membrane tethered or soluble bursicon cDNAs as indicated.

Desensitization Assays

In brief cells were transfected as described for luciferase assays. For soluble ligand based desensitization, 24 hours following transfection, bursicon conditioned media or serum

free DME was added onto cells and incubated overnight. The following day bursicon conditioned media or isoproterenol was then added onto cells for an additional 4 hours. Luciferase levels were then quantified. For membrane tethered ligand based desensitization, cells were transfected as previously described by with the addition of increasing amounts of membrane tethered ligand. Approximately 44 hours after transfection, cells were then stimulated with bursicon conditioned media for 4 hours. Luciferase levels were then quantified. Bursicon conditioned media was prepared as previously described (Harwood, Fortin et al. 2013).

Plasmids

Rk pcDNA1.1, β-galactosidase pcDNA1.1, membrane tethered ligands cDNAs, and bursicon plasmids were generated as previously described (Harwood et al. 2013). For the luciferase reporter gene plasmid (CRE-LUC-HCL-PEST), a 6x cAMP response element was cloned into the promoter region pGL4.22, a destabilized luciferase gene (Promega, Catalog # E6771).

Transgenic fly generation

The coding sequence of CFP-tBur-β-α pcDNA1.1 (Harwood et al. 2013), was subcloned into pUAST vector using EcoRI and XbaI restrictions sites. CFP-tBur- β-α pUAST plasmid was sequence verified and then sent to Duke University Model System Genomics *Drosophila* embryo injection service. Positive transformants on W1118 genetic background were then selected by the presence of red eye color. Siblings were crossed and homozygous stocks were maintained.

Drosophila Stocks and Culture

w1118 (FBal0018186), Actin5C-Gal4 (y1 w*; P[Act5C-GAL4]17bFO1/TM6B, Tb1; FBst0003954), how-Gal4 (w*; P[GawB]how24B; FBst0001767), D42-Gal4 (w*; P[GawB]D42; FBst0008816), Mef2-Gal4 (y[1] w[*]; P[w[+mC]=GAL4-Mef2.R]3; 27390), Ap-Gal4/Cy (y[1] w[1118]; P[w[+mW.hs]=GawB]ap[md544]/CyO; FBst0003041), Actin88F-Gal4/Cy (w*; P[Act88F-GAL4.1.3]81B, P[Mhc-tauGFP]2/SM6b; FBst0038460) were all obtained from Bloomington Stock Center. Bur-α RNAi (w1118; P[GD3951]v13520; FBst0451049) was obtained from Vienna *Drosophila* RNAi Center. Rk RNAi (8930R-1; FBsf0000434575) was obtained from NIG-FLY (National Institute of Genetics, Mishima, Shizuoka Japan). Rk^{Pan}-Gal4 (Diao and White, 2012) was a gift from Dr. Benjamin White at the National Institute of Health. All fly stocks were maintained at 25 °C on a standard 12 hour light-dark cycle. *Drosophila* growth media contained cornmeal, agar, brewer's yeast, dextrose, sucrose, wheat germ. All crosses between UAS stocks and Gal4 stocks were performed at 29 °C using standard *Drosophila* growth media. For crosses, virgin females of indicated UAS constructs or W1118 flies were collected and mated with Gal4 stock males. Flies were then counted after eclosion and assessed for wing expansion defects. Each fly cross was performed on a minimum of 3 separate occasions.

Data Analysis and Statistics

All luciferase and fly cross data were graphed and analyzed using GraphPad Prism 5 (GraphPad Software Inc, La Jolla, CA).

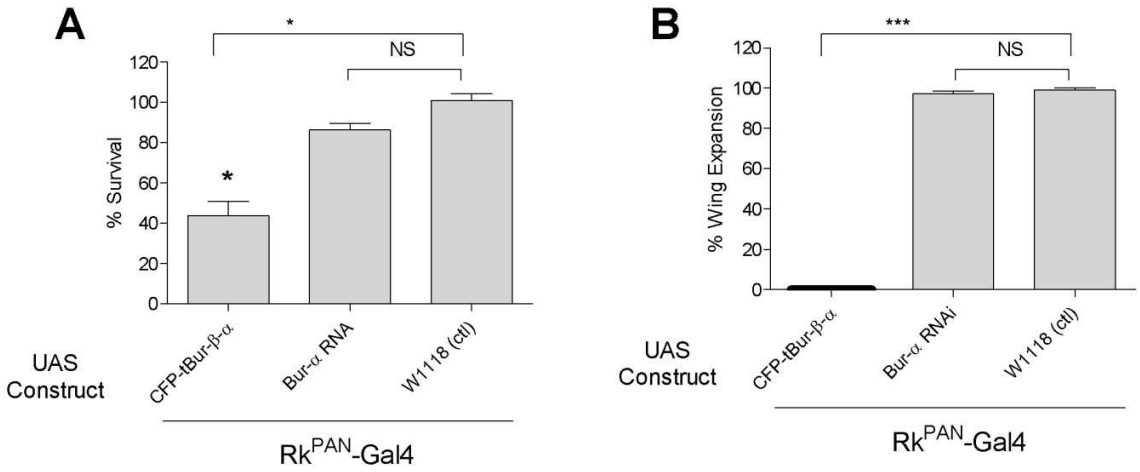


Figure 2.6. Expression of membrane tethered bursicon with Rk^{PAN}-Gal4 results in reduced survival and in wing expansion defects.
Expression of CFP-tBur-β-α, but not Bur-α RNAi in muscle results in a reduction in survival through eclosion (A) and wing expansion defects (B). Abbreviations: UAS= upstream activation sequence, CFP-tBur-β-α = cherry fluorescent protein membrane tethered bursicon, Rk^{PAN}= Gal4 expression under rk promoter sequence. Statistics: Data represent mean $\pm$ SEM of experiments done in triplicate. 134-143 progeny were counted for each independent cross. ***=P<.0001, *=P<.05, NS=P>0.05.

Chapter 3:

High throughput screen to identify novel modulators of the *Drosophila* GPCR rickets

Abstract

Malaria is a global health epidemic that results in widespread morbidity and mortality in regions where preventions and/or treatments are not widely available. One strategy to control the spread of malaria is to prevent the transmission of *Plasmodium* to humans by using insecticides to kill disease vector insects such as the mosquito *Anopheles gambiae*. A major issue with vector control strategies is the development of insecticide resistance due to the small number of compounds that are available. Any chemical that targets a novel pathway and compromises insect viability would be a valuable addition to the current armamentarium of insecticides. Rickets (rk) is a G protein-coupled receptor (GPCR) that controls wing expansion, melanization, and sclerotization in insects. Recent studies using RNA interference have also indicated that rk expression is required for developmental survival through eclosion of the model organism *Drosophila melanogaster*. Given the successful track record of GPCRs as drug targets and the role of rk in development and wing expansion; this receptor fulfills the criteria of a promising insecticide target. Currently the only known modulator of rk is the endogenous peptide agonist bursicon. No antagonists or small molecule ligands have been identified. To develop an antagonist as a putative insecticide, our lab has carried out screens of two different compound libraries. The first strategy utilized an *in vitro* screen of 3,709 plant extracts derived from Chinese medicinal herbs. The second *in vitro* assay was larger in scale and tested 378,311 synthetic small molecules. Both screens were cell based assays that yielded high throughput screening data of high quality. Initial screens were done using the *Drosophila melanogaster* rk ortholog (rk). Current efforts are now focused on validating putative hits as specific rk antagonists and demonstrating efficacy in an *in vivo*

model. We have also cloned the *Anopheles gambiae* rk ortholog (GPR-rk) and both its corresponding bursicon subunits. With this receptor we have adapted an *in vitro* signaling assay to determine the extent to which select compounds also target GPR-rk in this important disease vector.

Introduction

The development of insecticide resistance is a recurring problem in the control of both agricultural pests and disease transmitting insects. There is currently a need for novel insecticides with alternative molecular targets. There are 4 major classes of insecticides; pyrethroids, organochlorines, carbamates and organophosphates. These major insecticides work through two distinct modes of action in the nervous system. Pyrethroids and organochlorines open voltage gated sodium channels, while carbamates and organophosphates inactivate acetylcholinesterase (Enayati and Hemingway 2010). Environmental concerns have long been associated with a variety of insecticides. The disruption of natural ecosystems due to the inadvertent death of ecologically important non-pest insects was a consequence of heavy insecticide applications. Toxicity in humans from ingesting pesticides through contaminated water supplies or residue on agricultural products is also of great concern. The most commonly cited example of pesticide toxicity in humans is the organochlorine dichlorodiphenyltrichloroethane (DDT). While DDT acts on a specific voltage gated sodium channel in insects, it has shown genotoxicity and activity as an estrogen analog in humans. Because of this off target effect in humans, DDT was banned in the United States in 1972. Worldwide use is limited to the control of disease vectors such as the malaria transmitting mosquito, *Anopheles gambiae* (Turusov, Rakitsky et al. 2002). In addition to the environmental and health concerns, insecticide resistance to DDT has also emerged. Known point mutations have been identified in the *A. gambiae* sodium channel gene and have been shown to confer resistance to DDT (Pinto, Lynd et al. 2007). Unfortunately resistance is not just limited to DDT; examples of resistance have been identified for all 4 classes of

insecticides. Resistance conferring point mutations in acetylcholinesterase have also been indentified in *A. gambiae* populations, rendering organophosphates and carbamates less effective (Hemingway, Hawkes et al. 2004; Ramphul, Boase et al. 2009).

Cross reactivity of insecticides is also a concern. This is exemplified with a relatively new class of insecticides that act on nicotinic receptors, the neonicotinoids. Recent studies have suggested the neonicotinoids are having significant negative effects on the health and survival of global honey bee (Genus: *Apis)* populations. Given our reliance on honey bees to pollinate important food crops, this could have a significant impact on our global food supply (Blacquiere, Smagghe et al. 2012).

Given these concerns, efforts to identify new classes of insecticides are needed for three main reasons. First, there are human health and safety concerns associated with current insecticides. Second, there are a relatively small number of proteins that insecticides target resulting in the emergence of resistance. Finally, the lack of species specificity results in unintended effects on other species of insects resulting in potential harm to the greater ecosystem.

G protein-coupled receptors are attractive as potential insecticide targets as they have proven to be highly druggable. Select GPCRs found in insects have low sequence homology with corresponding vertebrate proteins and in some cases have no known vertebrate homolog making cross reactivity of compounds less likely (Van Loy, Vandersmissen et al. 2008). Recent work on the insect GPCR rk has suggested that it could be a promising molecular target for the development of a novel class of insecticides. Much of this research has come from the study of rk using the model system *Drosophila melanogaster*. Classical *Drosophila* mutant's rk^1 and rk^4 have point

mutations encoding premature stop codons in the rk coding sequence. Flies that are homozygous for these mutations show characteristic defects in wing expansion, cuticle hardening, and melanization. However, it has become increasingly clear that these fly lines are hypomorphic rather than receptor null mutants. Further studies using RNA interference (RNAi) for rk have shown that global and tissue specific knockdown of this receptor results in developmental lethality (Loveall and Deitcher 2010). In addition, our own studies have shown that expression of a bursicon membrane tethered ligand in a variety of tissues results in receptor desensitization and developmental lethality (Chapter 2). The fact that both inappropriate activation or knockdown of rk causes lethality, in conjunction with the observation that impaired wing expansion and cuticle hardening would occur if the insect did survive development, makes rk a plausible target for the development of a putative insecticide.

As a first step toward validating rk as a viable insecticide target, we sought to find novel compounds that block bursicon mediated rk signaling. Currently there are no known rk receptor antagonists and lead compounds need to be identified in order to assess their potential use as insecticides. One common methodology employed to discover new chemical modulators of a molecular target is a cell culture based high throughput screen (HTS). Major hurdles of conducting an HTS are the high cost, access to large compound libraries, automated screening equipment, and support from chemists for probe optimization after compounds have been identified.

We used two different approaches to conduct an HTS to identify a potential rk receptor antagonist. The first strategy we pursued was a screen of a small series of natural products. Purified plant extracts were derived through a rigorous separation

procedure. Libraries of natural product fractions are available from ICCB-Longwood Screening Facility at Harvard Medical School. One of these collections is the Authenticated Traditional Chinese Herbal Extract Collection (ATCHEC). This collection contains 3,709 plant fractions (Eisenberg, Harris et al. 2011). To generate this library, a series of plants were collected by botanists in China, separated into distinct parts, and chemically fractionated using high pressure liquid chromatography (HPLC). Each fraction is partially purified, but still maintains a mix of potentially active compounds. By screening plant fractions, many compounds can be tested from a relatively small number of samples. Once an active fraction is found, the active component must then be isolated. The identification of a natural product derivative that acts as an rk antagonist would be desirable given that a naturally derived product may have less potential for an adverse environmental impact than a synthetic chemical (Isman 2006).

The second approach involved screening a library of hundreds of thousands of synthetic small molecules. Unlike natural products, one advantage of small molecule screens is that once hits are found, the identity of the compound is immediately known. However, small molecule screens require substantially more resources that are often only available to large pharmaceutical companies. To address the accessibility issue, the National Institute of Health has made funding and resources available to the academic community through the Molecular Libraries Probe Production Centers Network (MLPCN). The MLPCN is comprised of multiple screening centers around the country that maintain an ever growing library of small molecules. Screening centers are able to implement HTS assays and can provide follow up chemistry support for probe optimization once lead compounds have been identified (Roy, McDonald et al. 2010).

Another advantage of synthetic small molecule screens is the feasibility of probe optimization. The use of small molecules enables the modification of a base structure to make a series of derivatives which can then be tested for structure activity relationships (SAR). For our purposes, this approach can be used to increase the affinity of antagonists, thereby making them more effective as insecticides. Also through this process, receptor selectivity of compounds on insect rk orthologs can be pursued. Specifically we are interested in the cross reactivity of compounds between *Anopheles gambiae* (the intended target) and *Apis mellifera* (an environmental and commercially important insect that needs to be preserved). The ultimate goal would be to identify a compound that acts specifically on rk in *A. Gambiae* but not on rk in honey bee species such as *A. mellifera.* In addition it would be important to test that other evolutionarily related vertebrate receptors are not affected by rk antagonists. Here we could use structure activity relationships to test and possibly modify any compounds that also antagonize mammalian glycohormone receptors, such as the FSH receptor.

To address these challenges we have begun the process of HTS in order to identify novel natural product and small molecule derived inhibitors of rk mediated signaling. To enable these studies we have developed several iterations of *in vitro* based screens. In addition we have also successfully cloned the *Anopheles gambiae* ortholog of rk as well as both *A. gambiae* bursicon subunits to initiate understanding the potential species specificity of any novel inhibitors identified in the primary screens.

Results

While RNAi studies suggest rk is a promising putative target, the identification of small molecules or natural products capable of blocking rk signal transduction is required to validate rk as a viable insecticide target. To identify novel compounds we set up two different high throughput screens utilizing a natural product library and a small molecule library.

Natural Product Screen

The first library of compounds tested was comprised of 3,709 plant fractions collected from Chinese medicinal herbs. The library was pre-plated in 384-well plates and the experimental protocol illustrated in Figure 3.1 was used. The primary screen assessed the ability of each compound to inhibit rk. For an initial test of rk specificity, all compounds were counterscreened on the *Drosophila* dopamine receptor 2 (DopR2) which also signals through $G\alpha_s$. Based on the ability to inhibit rk, but not DopR2 signal transduction, 27 fractions (of 3,709 tested) were selected for further analysis. The 27 isolates were assessed at 3 concentrations for their ability to inhibit rk signaling (Figure 3.2). The most potent four fractions (C3, C9, E7, E11) were then selected for additional analysis which included a retest with rk as well as a second $G\alpha_s$ coupled receptor counterscreen using the human Melanocortin 4 receptor (MC4R). The second counterscreen was used to further ensure receptor selectivity (Figure 3.3).

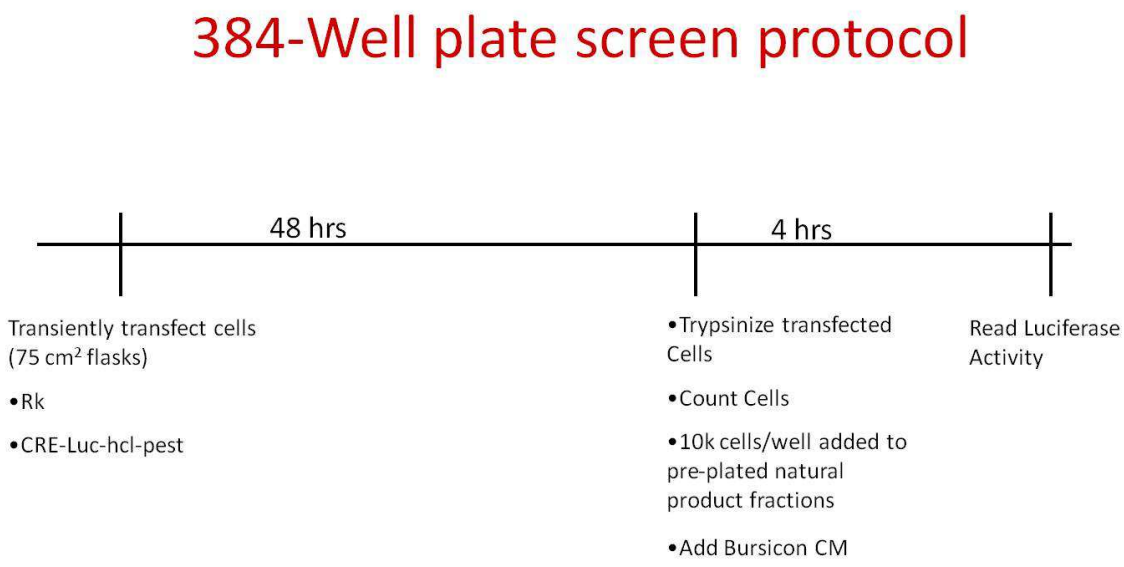


Figure 3.1 Experimental protocol for 384-well plate natural product screen.

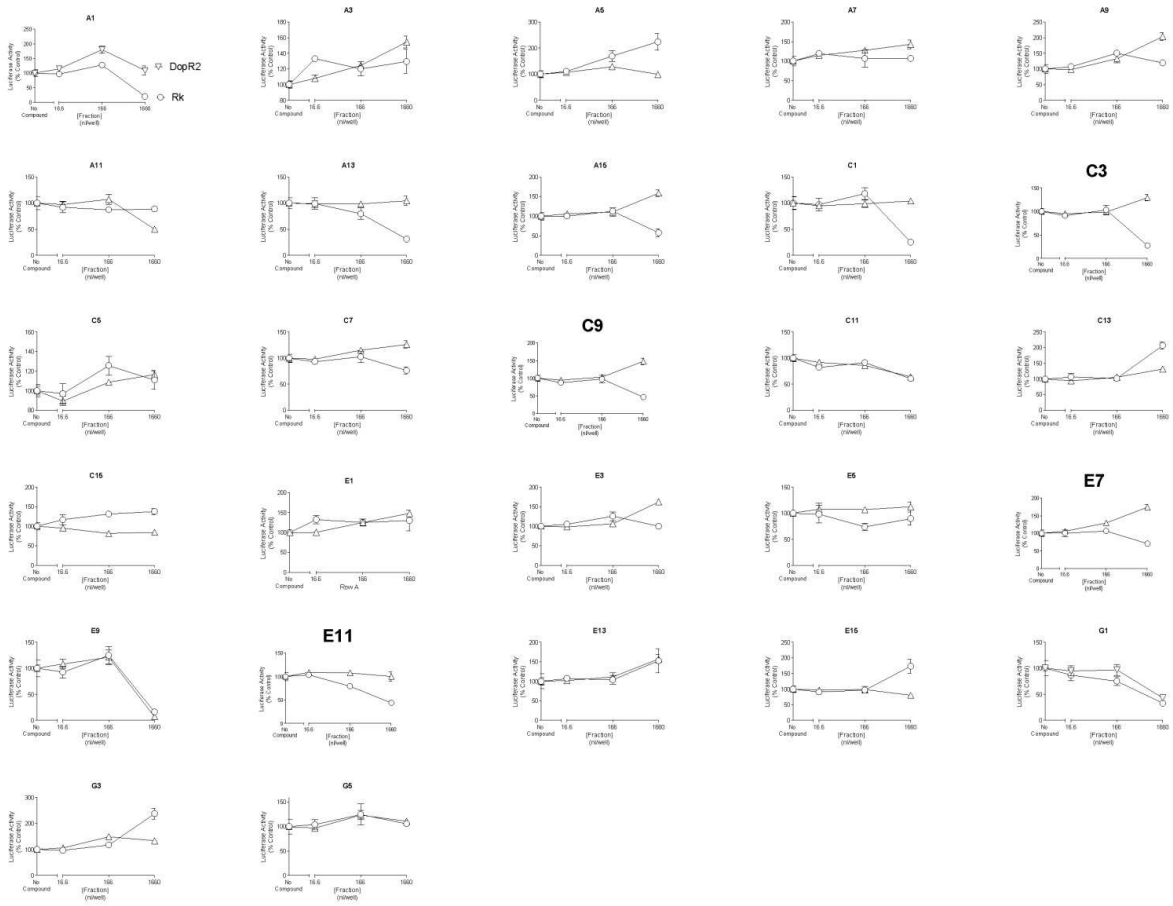


Figure 3.2. Follow up analysis for 27 fractions. Three point concentration response analysis of 27 compounds identified in primary natural product screen. Bolded fraction numbers C3,C9,E7, and E11 were selected for further analysis. Inhibition was assessed for both rk and DopR2 in a 96-well plate format. Plant extracts were added to cells for 15 minutes and then bursicon CM or dopamine was added for 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk or DopR2, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from a single experiment, performed in triplicate. Experiment performed by B. Harwood and J. Doyle.

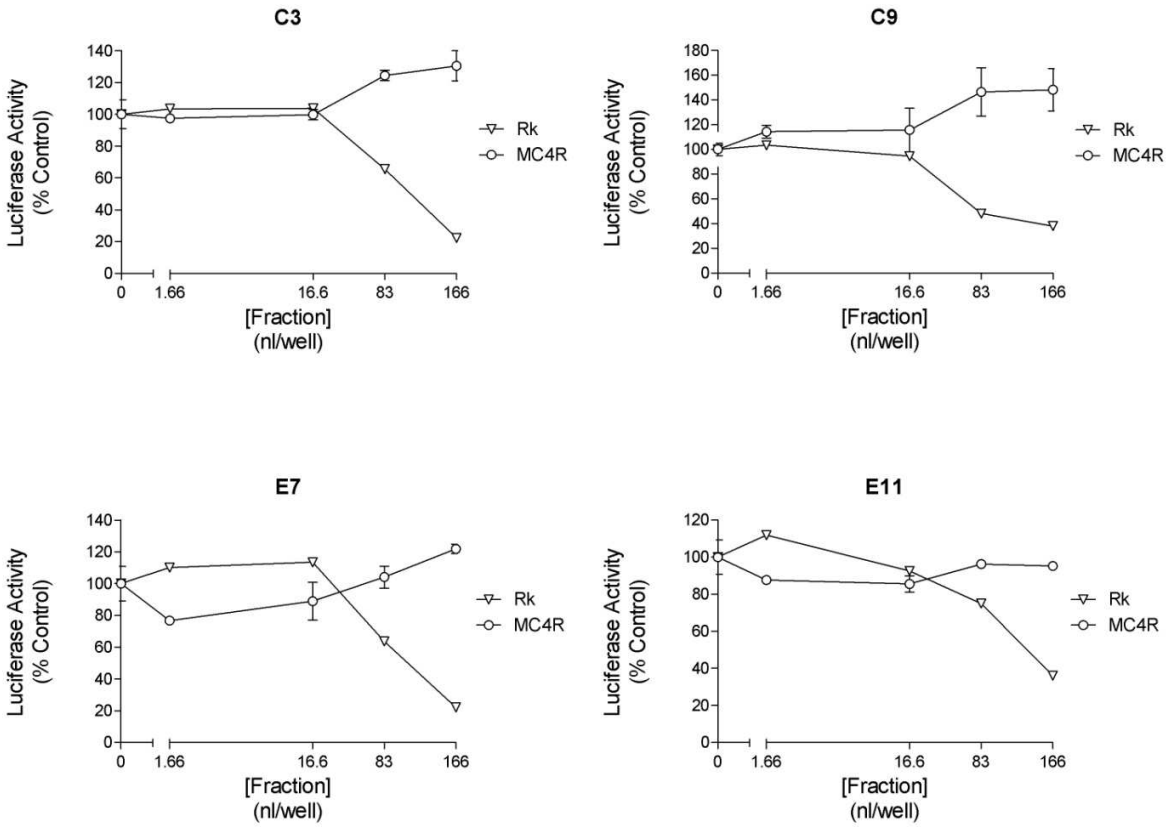


Figure 3.3. Final four fractions were retested on rk and a second counterscreen receptor MC4R. Fraction C3,C9,E7, and E11 were further analyzed. Inhibition was assessed for rk and MC4R in a 96-well plate format. Plant extracts were added to cells for 15 minutes and then bursicon CM or α-MSH was added for 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk or MC4R, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from a single experiment, performed in triplicate. Experiment performed by B. Harwood and J. Doyle.

From the final four, a single fraction (C3) was selected for subfractionation to identify potential active components in collaboration with the laboratory of Jon Clardy at Harvard Medical School. Plant extract C3 was derived from the dried aerial portion of the patouchli plant, *Pogostemon cablin*. Subfractionation was done by HPLC. Sixty subfractions from C3 were screened and counterscreened using rk and MC4R assays in 96-well plates. C3 subfractions 46 and 47 were shown to selectively inhibit rk but not MC4R (Figure 3.4). A second round of HPLC was then performed on an independent stock of C3 to determine whether the same subfractions could successfully be identified. In a second round of assays, fractions 46 and 47 were shown again to selectively inhibit rk mediated signaling, while not reducing MC4R signal transduction (Figure 3.5). Following confirmation of active subfractions, fractions 46 and 47 were analyzed using nuclear magnetic resonance (NMR) and mass spectrometry to identify any enriched molecules. From this analysis, the compound Verbascoside (IUPAC: [(2R,3R,4R,5R,6R)-6-[2-(3,4-dihydroxyphenyl)ethoxy]-5-hydroxy-2-(hydroxymethyl)-4-[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl]oxyoxan-3-yl](E)-3-(3,4-dihydroxyphenyl)prop-2-enoate) was identified. Verbascoside is a phenylpropanoid with a caffeine moiety that is found in many plants (Figure 3.6). Powders of Verbascoside and two structurally similar derivatives Forsythoside A and Isoacteoside were ordered from a commercial vendor for further confirmation. Upon further assessment with enriched purified powders, Verbascoside and its two derivatives, while showing inhibitory activity on rk also appeared to inhibit signaling of MC4R (the counterscreen receptor) (Figure 3.7). Since selectivity for rk was one of the main criteria for development of a potential

inhibitor, given the findings with verbascoside and its derivatives, the decision was made not to pursue this compound class further. However, 3 other fractions (C9, E7, E11) still have potentially active components and could be subfractionated and analyzed at a later date (Figure 3.3).

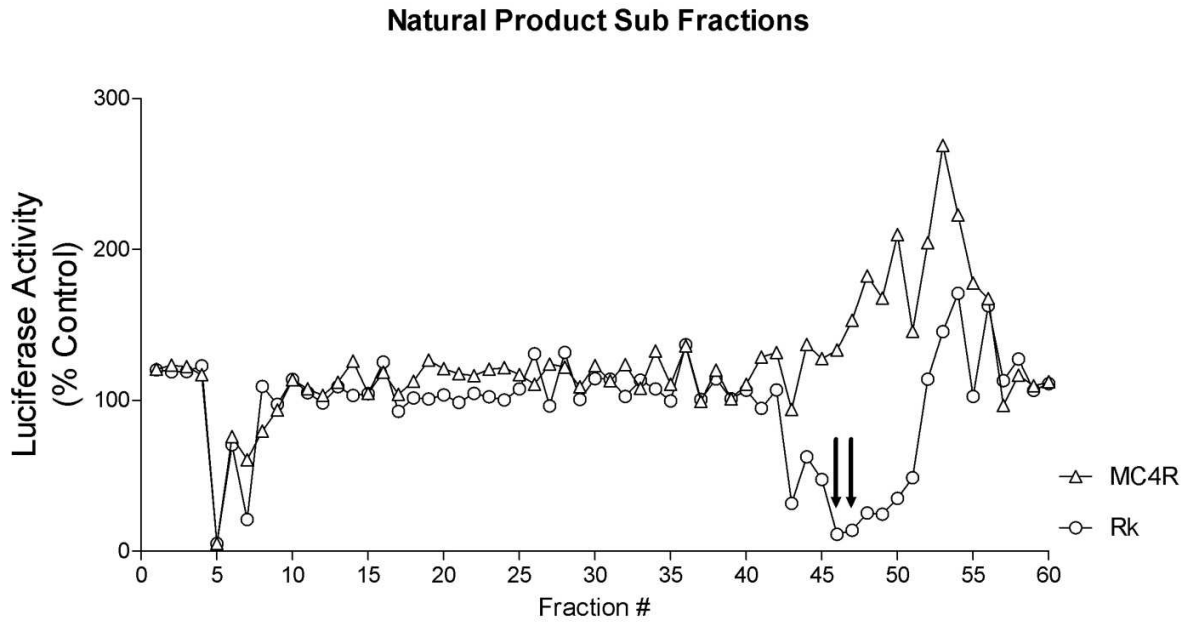


Figure 3.4. Subfractionation of fraction C3 derived from the dried aerial portion of *Pogostemon cablin*. Fractions 46 and 47 had the most differential activity. Inhibition was assessed in a 96-well plate format. Subfractions (15µg/well) were added to cells for 15 minutes and then bursicon or α-MSH was added for 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean from a single experiment, performed in duplicate. Experiment performed by B. Harwood and J. Doyle.

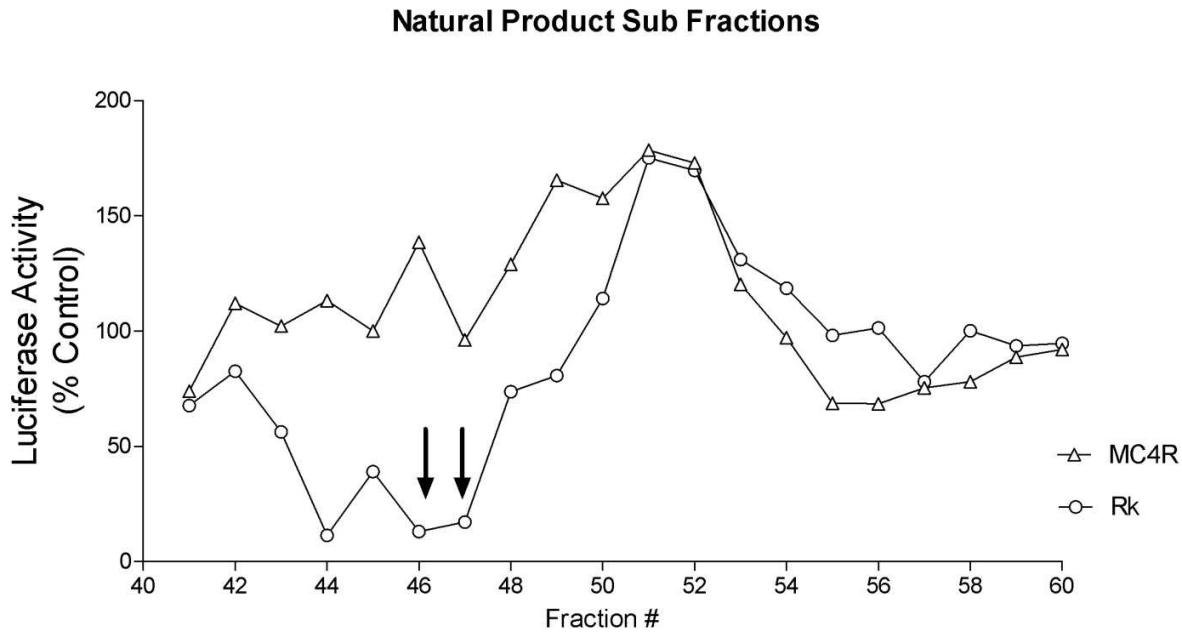


Figure 3.5. Second fractionation process also identified fractions 46 and 47 as differentially active between rk and MC4R. Repeated subfractionation of fraction. Fractions 46 and 47 again had the most differential activity. Inhibition was assessed in a 96-well plate format. Subfractions (15µg/well) were added to cells for 15 minutes and then bursicon or α-MSH was added for 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean from a single experiment, performed in duplicate. Experiment performed by B. Harwood and J. Doyle.

Verbascoside

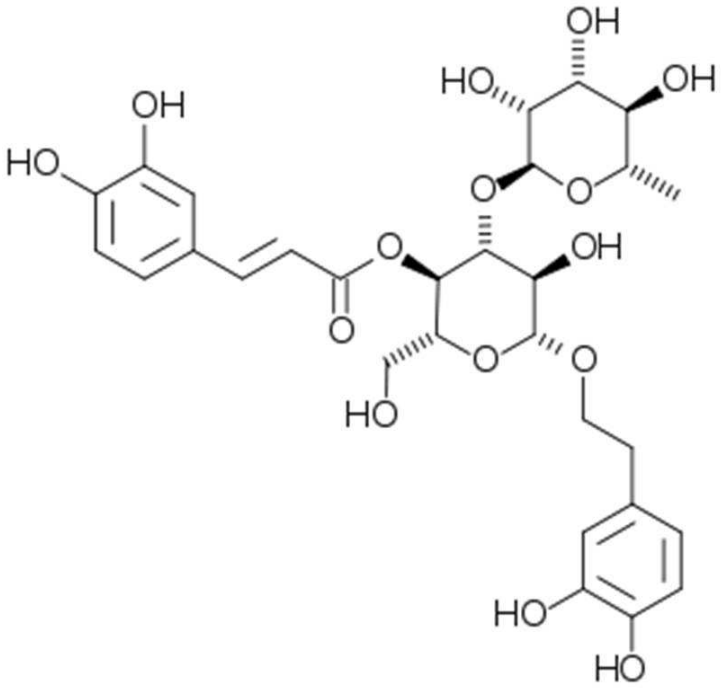


Figure 3.6. Structure of Verbascoside, the most prominent component identified in subfractions 46 and 47.

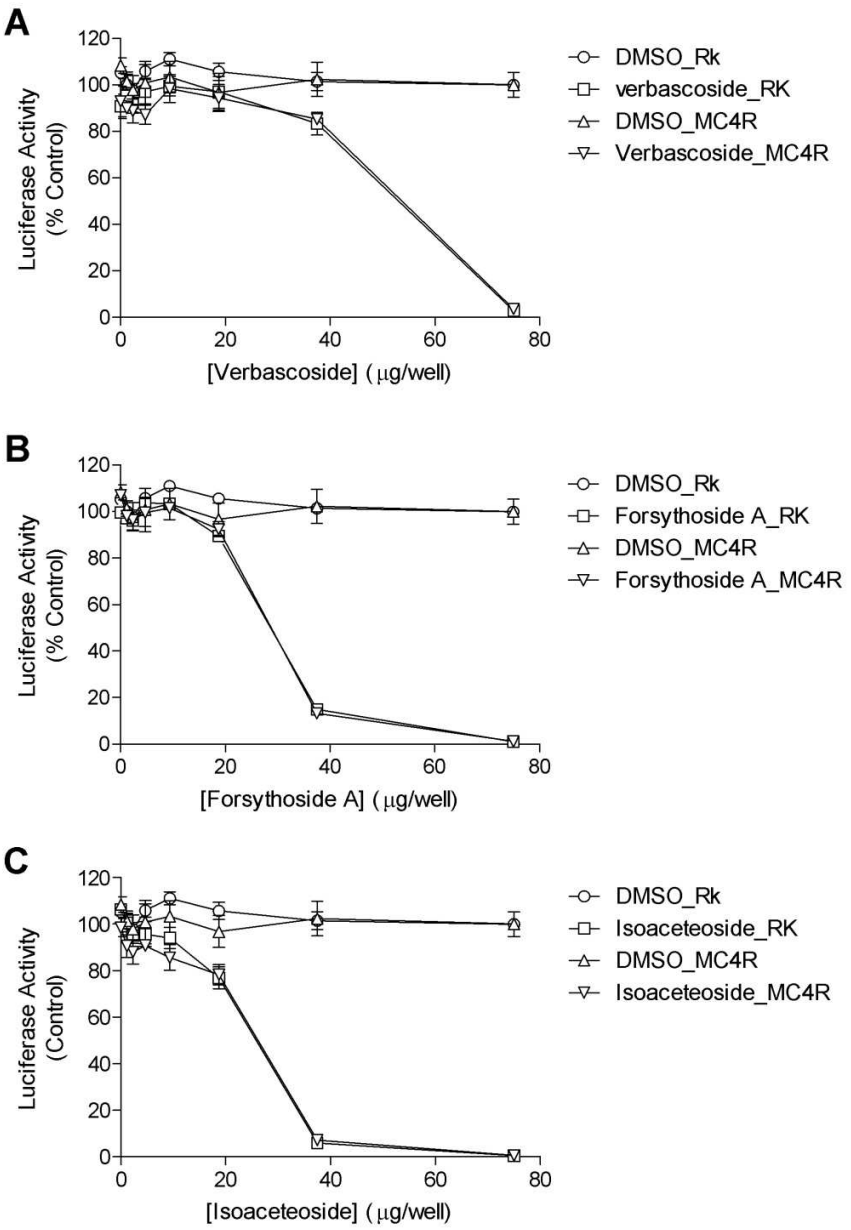


Figure 3.7. Powders of Verbascoside and two structurally similar compounds are not receptor specific. A) Inhibition curves of rk and MC4R with verbascoside. B) Inhibition curves of rk and Mc4R with Forsythoside A. C) Inhibition curves of rk and MC4R with Isoaceteoside. Indicated compounds or DMSO were added to cells for 15 minutes and then bursicon CM or α-MSH was added for 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk or MC4R, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from a single experiment, performed in triplicate.

MLPCN small molecule screen

The natural product screen used a relatively small library and to date has not yielded any selective rk inhibitors. As a second strategy, a screen of synthetic small molecules was also pursued. Through the NIH HTS program, a screen of the MLPCN library of small molecules was conducted using a strategy outlined in Figure 3.8. Prior to the large scale 1536-well screening protocol (outlined in Figure 3.9), a 2,000 compound pilot screen and counterscreen was performed to validate the screening system. The smaller library contains representative examples for all of structures in the MLPCN library. In 384-well plates, compounds were screened for inhibition of rk (Figure 3.10) or MC4R agonist induced signaling. Since no rk inhibitor currently exists, serum free media without bursicon was used as a positive control to mimic rk antagonism (no receptor activation). Screening was performed with bursicon and α-MSH at a their respective EC_{50} concentrations. Each plate of compounds was tested in duplicate using the format seen in Figure 3.12. The main performance indicator that had to be achieved during the pilot was that each plate must have must have a Z′ factor that exceeds 0.5. The Z′ factor is a statistical measure to assess the signal to noise ratio in relation to the variability for each plate. The formula to calculate Z′ is found in Figure 3.13.

As seen in Figure 3.8, the full screen is a multi-step process that began with a primary screen of 378,311 compounds from a common NIH library of small molecules, as well as a proprietary compound collection from the Broad Institute Diversity Oriented Synthesis (DOS) library. For the large scale screen, the 384-well plate assay was adapted to a 1536-well plate system in order to allow for higher throughput capabilities.

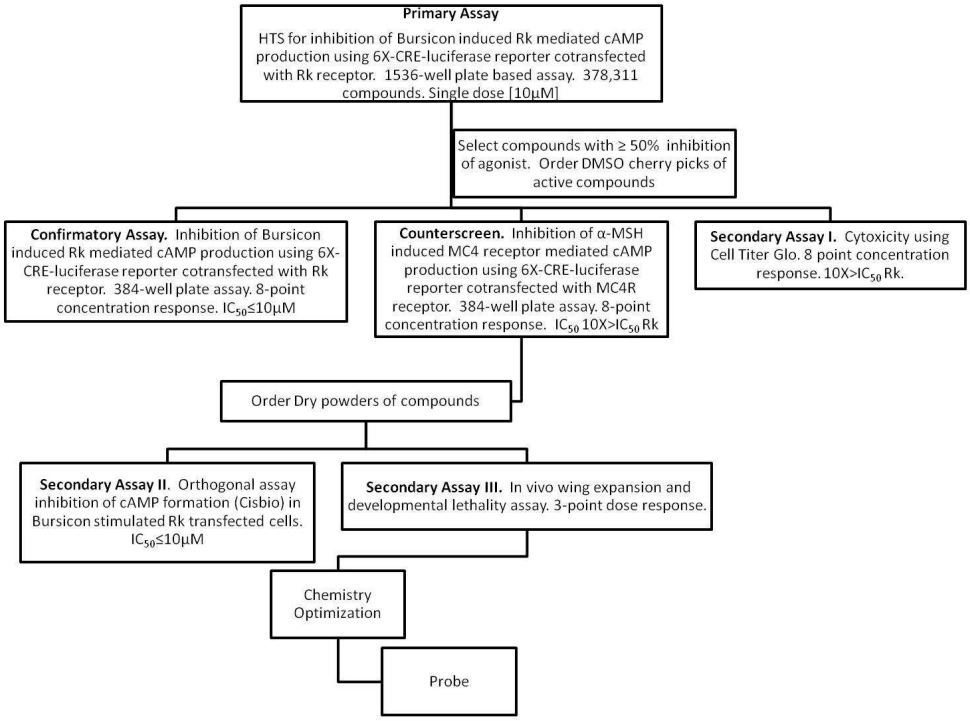


Figure 3.8. Flow chart of high throughput screen at the Broad Institute.
The chart outlines the stepwise approach to screening for rk antagonists in a high throughput system and subsequent validation of potential compounds.

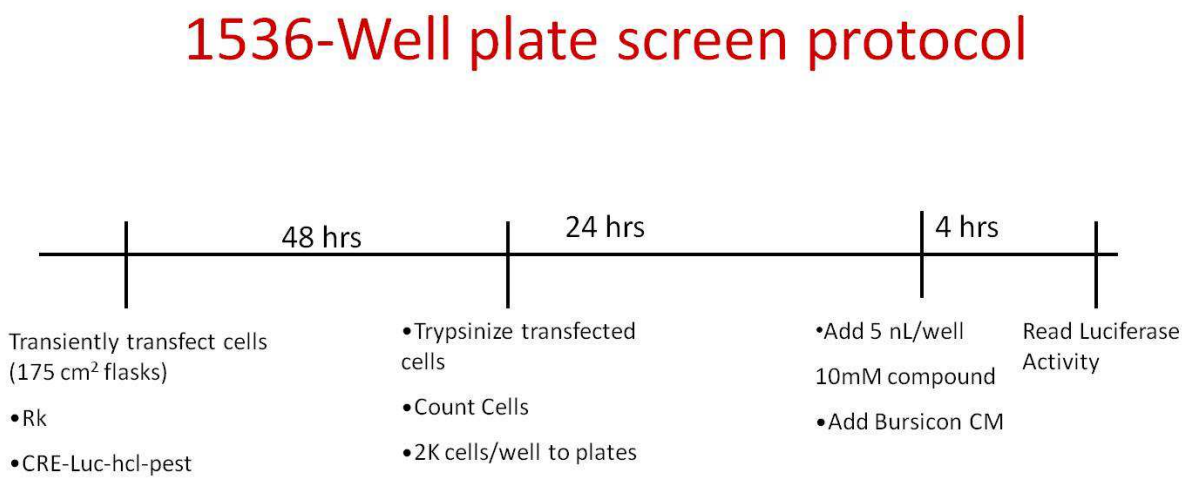


Figure 3.9. Experimental protocol for 1536-well plate natural product screen.

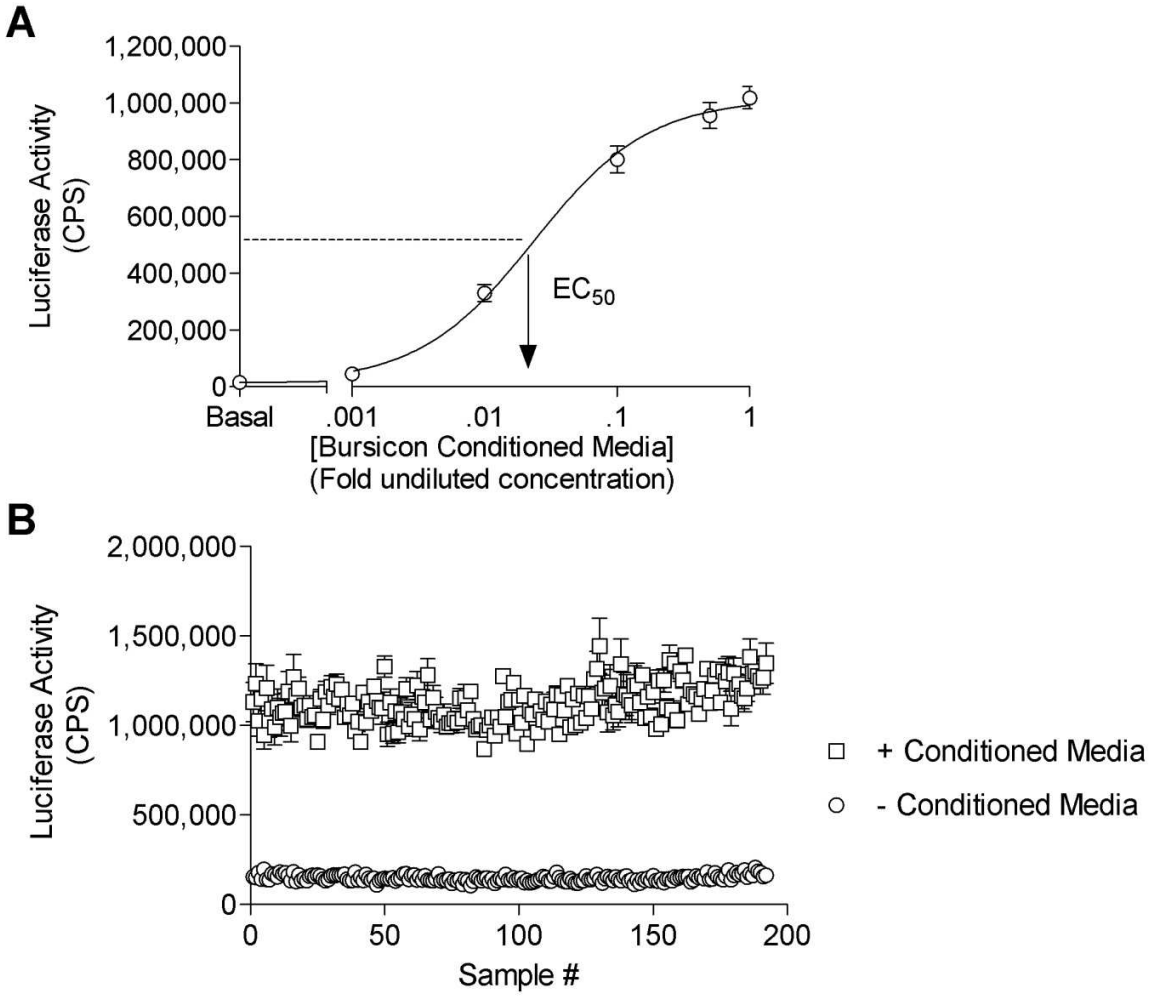


Figure 3.10. High throughput screening for an rk antagonist screen validation. A) Concentration dependent activation of rk with bursicon conditioned media made in triple 175 cm^2 flasks. For activity assay, a series of tenfold dilutions of conditioned media (1=undiluted conditioned media) was added to cells 20 hours after transfection in 96-well plates; the duration of ligand stimulation was 4 hours. Arrow indicates dilution of bursicon required for 50% of maximal rk stimulation. B) Validation of 384-well plate screening format with or without the addition of bursicon conditioned media. The duration of ligand stimulation was 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from a single experiment, performed in triplicate. Abbreviations: CPS=counts per second. EC_{50}= 50% of maximum receptor activation. Experiments performed by B. Harwood and J. Doyle.

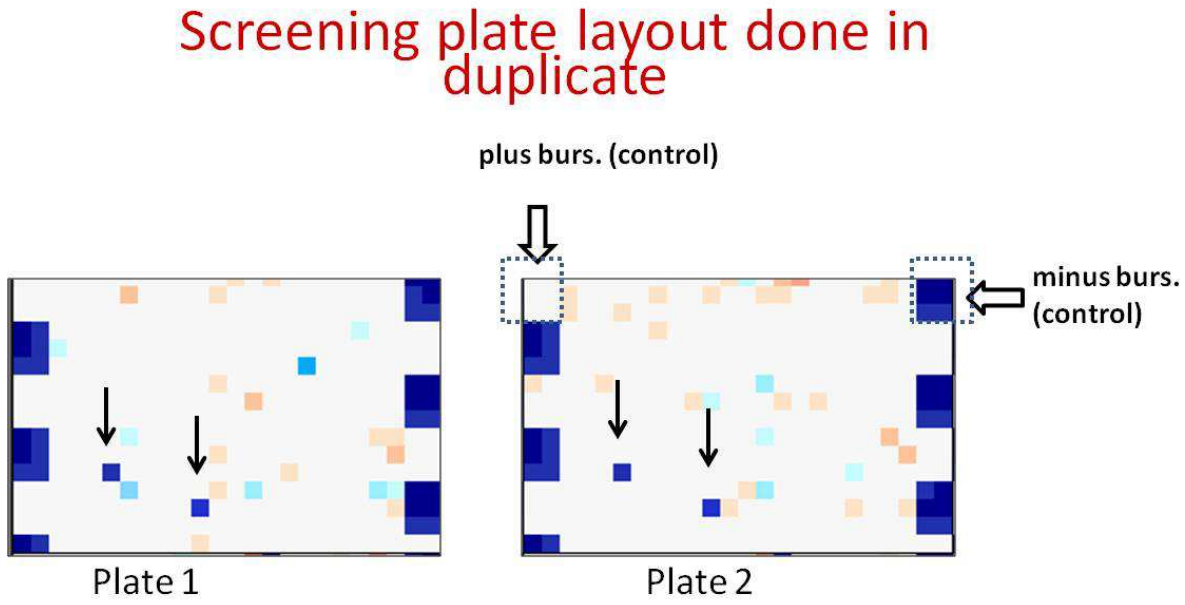


Figure 3.11. HTS screening plate layout and heatmap based read out.
HTS screen was done in 1536 well plates in duplicate. Thin arrows indicate 2 wells on independent plates with inhibition of signal (dark blue). Wide arrows point out positive and negative controls. Positive control that represents potential inhibition of rk signaling was wells with no compound, DMSO and media without bursicon (Dark blue). Negative control wells had no compound, DMSO, and bursicon conditioned media (White). All experimental wells were compared to negative control luciferase levels.

Rickets screen validation plate display

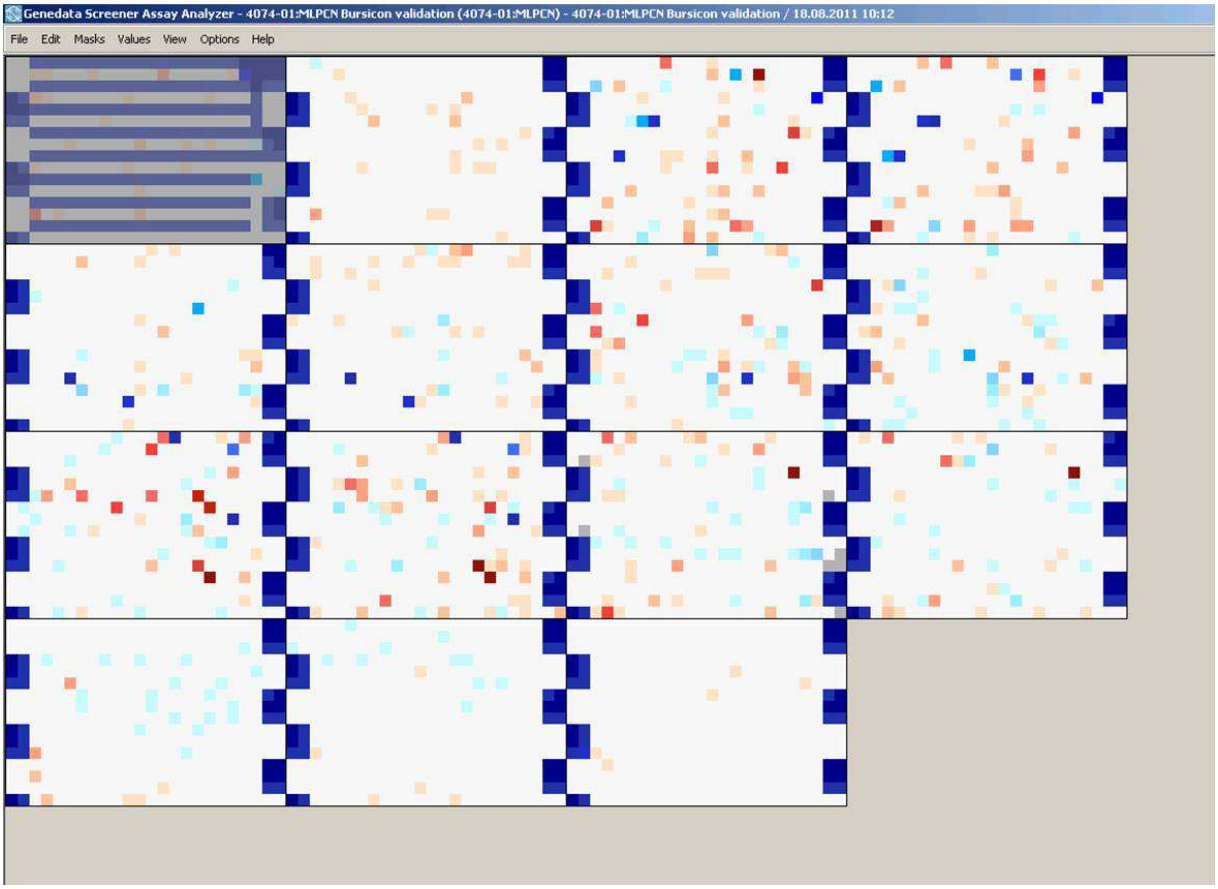


Figure 3.12. Screen shot of image representing luciferase from all plates for rk pilot screen done at the Broad in 384-well plates. Experiments performed by B. Harwood and J. Doyle.

Trend Display: Z'

Figure 3.13 Z factors for all plates in Broad screen.
Z factors for all plates in Broad screen exceeded cutoff threshold of 0.5. Equation to determine Z factor is shown in lower right hand corner. μ = mean, σ = standard deviation, $_p$ = positive control, $_n$ = negative control.

The primary screen was performed in duplicate. Following the screen, the data were analyzed to find compounds that resulted in ≥50% inhibition of rk mediated activity. Based on the primary screening results, 906 compounds were cherry picked for further assessment. The 906 compounds were then assessed in 384-well plates using the methods described for the pilot screen. Compounds were tested in duplicate using 8 point dose response curves to assess inhibitor activity. An MC4R counterscreen was also conducted. A cell titer glo assay (CisBio) was run in parallel to test for potential toxicity of compounds. The criteria for compounds to move beyond the 8 point concentration response curve step was that they must be potent (rk: IC_{50}<10μM), selective (MC4R IC_{50}: 10x> rk IC_{50}), and non-toxic (toxicity EC_{50}: 10x> rk IC_{50}). From the secondary screen of 906 compounds, 12 fulfilled these criteria. Structures and IC_{50} values for these 12 compounds are shown in Figure 3.14. Larger amounts of powder stocks will be ordered from commercial vendors. The powders of these compounds will be tested for purity and chemical composition to ensure that they match the compounds from the library stocks. In addition to retesting these compounds using the previously described assay, a secondary cAMP assay will be used to show inhibition of signal. These compounds will also be fed to *Drosophila* to see if any effects on wing expansion or development are observed *in vivo*. Once lead probes are identified through this experimental approach, compounds will be modified and optimized for enhanced affinity and species specificity. For analysis of species selectivity, we have cloned the *Anopheles gambiae* ortholog of rk (GPR-rk) as well as the corresponding α and β bursicon subunits. *Anopheles gambiae* bursicon conditioned media can be made as previously described for *Drosophila* and we have confirmed that GPR-rk also signals via a cAMP dependent pathway using our *in*

vitro luciferase based signaling assay (Figure 3.15). We have also used powders from the 4 most potent compounds and tested them on both rk and GPR-rk. As seen in Figure 3.16, there appears to be cross reactivity of identified compounds between rk orthologs. Importantly the compounds do not inhibit hMC4R or DopR2, suggesting rk selectivity (data not shown). Studies to assess whether these four compounds are effective rk antagonists *in vivo* are underway.

DOS Hit Structures

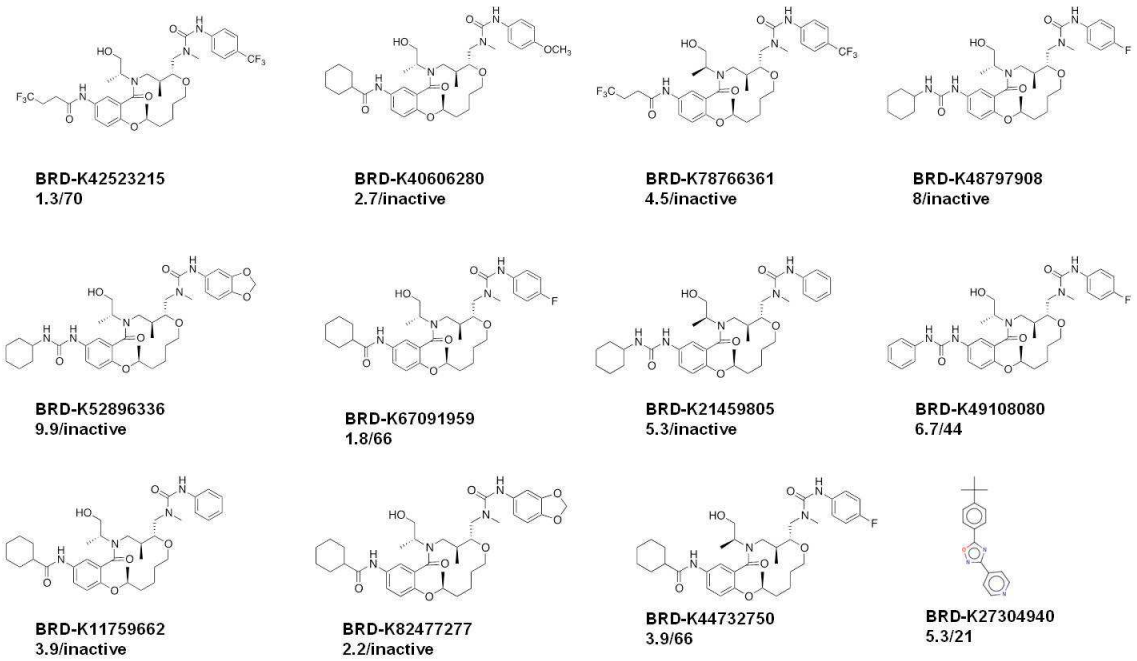


Figure 3.14. DOS hit structures. Chemical structures of 12 compounds ordered as dry powders that passed analysis during the primary screen, secondary retest, counterscreen, and toxicity analysis. BRD= Broad Institute identification number. Values below BRD Number=μM IC_{50} for rk/MC4R. DOS=diversity oriented synthesis.

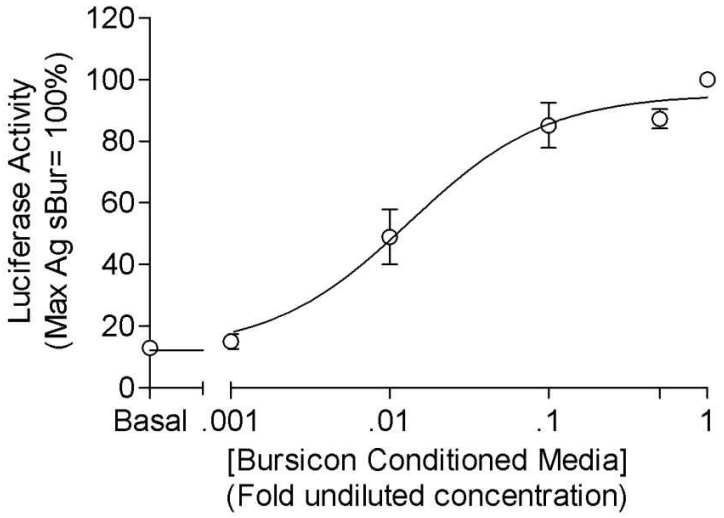


Figure 3.15. The *Anopheles gambiae* (Ag) rk homolog GPR-rk is activated by Ag bursicon conditioned media
A) Concentration dependent activation of GPR-rk with Ag bursicon conditioned media. A series of tenfold dilutions of conditioned media (1=undiluted conditioned media) was added to cells 20 hours after transfection; the duration of ligand stimulation was 4 hours. HEK293 cells were transiently co-transfected with cDNAs encoding: GPR-rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate.
Abbreviations: Ag sBur= *Anopheles gambiae* soluble bursicon.

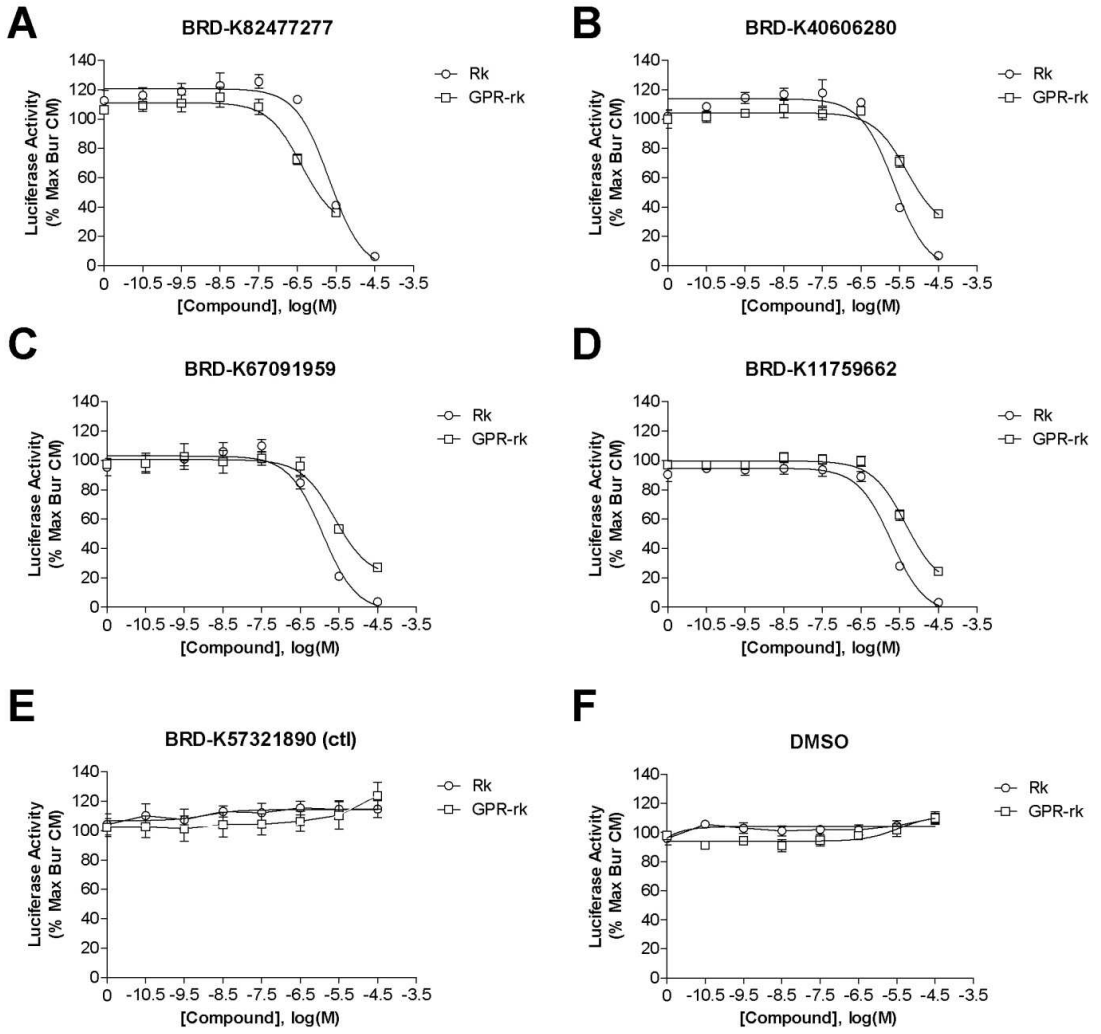


Figure 3.16. Inhibition by the four most potent small molecules identified during HTS. Inhibition of rk or GPR-rk signaling by A) BRD-K82477277, B) BRD-K40606280, C) BRD-K67091959, D) BRD-K11759662, E) BRD-K57321890 (inactive control), or F) DMSO in 96-well plate format. *Drosophila* or *Anopheles* bursicon CM was added 15 minutes after compound, followed by an additional 4 hour incubation. HEK293 cells were transiently co-transfected with cDNAs encoding: rk or GPR-rk, a 6X-CRE-Luc reporter gene, and a β-galactosidase gene for normalization. Data represent the mean ± SEM from a single experiment, performed in triplicate. Abbreviations: Bur CM= bursicon conditioned media, rk= *Drosophila* rk receptor, GPR-rk= *Anopheles* rk receptor.

Discussion

Natural Product Screen

As an initial approach we utilized a library of natural products to identify a *Drosophila* rk antagonist. Natural product extracts were appealing because a large number of compounds can be screened from a relatively small number of samples. Also, naturally derived insecticides have previously been successful. For example, Pyrethrin, structurally similar to pyrethroids (a major class of insecticide) is isolated from the flower of *Chrysanthemum cinerariifolium* and its use has gained popularity over the past several decades. It is also considered much less toxic than other pesticides (Hitmi, Coudret et al. 2000). In fact, Pyrethrin applications are frequently used to control the mosquito populations in Massachusetts to prevent the spread of Eastern Equine Encephalitis (Commonwealth of Massachusetts 2014).

For the rk assay, a 384-well plate format was used to screen 3,709 plant extracts. From the initial screen, twenty seven fractions were tested to determine whether they warranted further analysis (Figure 3.2). Based on preferential inhibition of rk compared to two other G_s coupled receptors *Drosophila* DopR2 and human MC4R, four extracts were selected and a single fraction (C3) of the patchouli plant was selected for subfractionation (Figures 3.3-3.5). This analysis suggested the active compound was in subfractions 46 and 47. Based on HLPC and NMR analysis the compound was determined to be Verbascoside. Unfortunately, when powdered stocks of Verbascoside were tested using inhibition assays, the results suggested that this compound was not selective due to the fact that it also inhibited MC4R (Figure 3.7). The discrepancy between the subfraction analysis and results with commercial powders could be the result

of one or more factors. One explanation is that Verbascoside was not the active component in subfractions 46 and 47. Alternatively, a mix of compounds may be required for rk specific inhibition and Verbascoside alone may be non-specific. Although disappointing, this information may still be useful if subfractionation of the other 3 primary plant extracts is pursued. Our natural product screening assay has been established. We have demonstrated that individual compounds can be isolated using our approach. As funding allows, future efforts will pursue the other plant extracts which had activity (C9, E7 and E11).

MLPCN Screen

While a natural products screen was one approach to identify an rk antagonist, we also developed a methodology for screening a large library of synthetic small molecules as potential inhibitors of rk. To date, we have successfully implemented a 1536-well screen of 378,311 small molecules from the MLPCN library. A full summary of the screen and all of the data for individual compounds has been deposited in the NCBI Pubchem bioassay database (AID 720647) for public access. From the 378,111 compounds assayed, 906 were selected and successfully analyzed using a 384-well plate assay. Based on toxicity, potency, and receptor specificity, 12 compounds were selected for further analysis (Figure 3.14).

Another long term goal as previously mentioned is to develop an rk antagonist that targets the *Anopheles gambiae* rk ortholog GPR-rk. To this end we have cloned GPR-rk receptor cDNA from *Anopheles gambiae* as well as both bursicon subunits. We have also confirmed the activity of the receptor and corresponding ligands using an *in*

vitro luciferase based assay. As with *Drosophila* rk, GPR-rk is also activated by heterodimeric bursicon in a concentration dependent manner (Figure 3.15). During probe optimization with *Drosophila* rk, potency on GPR-rk will also be measured to ensure the compound remains active on this receptor. The identification of a compound that is potent on GPR-rk would be ideal as it could be useful for disease vector control.

With this in mind, powder stocks of the four most potent compounds have been synthesized. All previous assays were performed using compounds that were previously diluted in DMSO for use within a small molecule library, so quantities are limited. Availability of independently synthesized compound provided an opportunity to confirm results obtained with library stocks and also to test species cross reactivity. Compounds have been tested on rk and GPR-rk (Figure 3.16). The four compounds selected appear to inhibit both rk and GPR-rk, suggesting that there is some species cross reactivity. Future experiments will test compounds on dLGR1, and a panel of human $G\alpha_s$ coupled receptors (including human glycohormone receptors) to better understand the extent of cross reactivity. These receptors were chosen for testing based on their amino acid sequence conservation with rk. Once all of this information is obtained, a lead compound will be selected for probe optimization. During the optimization process corresponding structural derivatives will be tested to develop a high affinity rk specific antagonist.

As a complementary approach, assessment of compounds in an *in vivo* assay is also currently underway. The four compounds are being administered to developing fruit flies to determine if any of the molecules affect survival through eclosion or wing expansion. Functional antagonists of rk would hypothetically have two effects. The first is developmental lethality if compounds are able to completely block rk stimulation. The

second is wing expansion and cuticle hardening defects if compounds only partially block rk stimulation. Wing expansion is also an excellent index to confirm that these compounds target rk *in vivo*. While not an obligate step, these results will also be used for lead probe selection if phenotypes are observed.

Using a GPR-rk antagonist to control populations of *Anopheles gambiae* is a strategy that could help prevent the transmission of malaria. In 2010 there were an estimated 216 million clinical cases of malaria diagnosed worldwide. Malaria caused 655,000 deaths, most of which occurred in young children. Africa in particular bears much of the malaria burden with approximately 90% of all cases taking place on this continent (World Health 2013). As exemplified by the controversial but successful use of insecticides such as DDT, vector control is a viable strategy to prevent the spread of malaria (Raghavendra, Barik et al. 2011). Therefore it would be highly desirable to have GPR-rk selective compounds as a new class of insecticide. While actual compound use in the environment is realistically a long way off, we can employ *in vivo* developmental assays in mosquitoes through a collaboration with the laboratory of Marc Muskavitch at Boston College. Therefore lead compounds can also be tested directly in *Anopheles gambiae* to determine if they cause defects in wing expansion, melanization, cuticle hardening, and/or developmental lethality. The observation of any of these predicted phenotypes *in A. Gambiae* would be a major step toward the promise of a new insecticide.

In conclusion, the development of an rk antagonist would provide a putative insecticide with a novel target and mode of action. A chemical inhibitor of the rk receptor could also be used to better understand the developmental significance of rk and

bursicon in insect species other than *Drosophila*. This would be especially important with insect species such as *A. gambiae*, where the use of genetic models and RNAi technologies are only in the beginning stages (Catteruccia and Levashina 2009). A novel receptor inhibitor would be useful to address questions regarding the role of rk at different points during the life cycle of *A. gambiae*. The question of temporal requirements is especially important as insects vary greatly in the progression of their life cycles and understanding when rk is expressed would be especially relevant for the application strategy of an insecticide. Current knowledge based on the *Drosophila* life cycle suggests that a GPR-rk antagonist would fall into the class of larvicides, and needs to be applied during insect development. Having an active inhibitor is important to confirm this as a viable strategy for other insect species.

We have utilized two distinct compound libraries, one comprised of natural products and another of synthetic small molecules. While we have yet to confirm our lead compound, the assays conducted so far have yielded promising candidates for further study. While the nature of molecular screening requires some luck to be successful, having sensitive, well performing assays is critical. Thus far both screens have these characteristics and at the very least can be used as templates to screen other insect GPCRs for antagonists. We remain confident that we will find an active and selective rk antagonist.

Materials and Methods

Rk Screen and MC4R counterscreen

HEK293 cells were seeded in T75 cm^2 flasks in DMEM (containing 10% Fetal Bovine Serum (FBS) and 100U/mL Penicillin/100µg/mL Streptomycin). Approximately 1.2 million cells/flask were added and grown overnight at 37°C in 5% CO_2.

Once cells reached 90 % confluence, they were then transfected with receptor and reporter. Transfections were performed under serum free conditions using PEI (23.4µL/ flask of a 1mg/mL stock solution). Flasks were transfected with the following amounts of cDNA: rk pcDNA1.1 (234ng/flask) or MC4R pcDNA1.1 (234ng/flask) and 1170ng/flask of CRE luciferase reporter gene (PGL 4.22, Promega). In brief, PEI and cDNA mixes diluted in serum free DMEM were combined and incubated at room temperature for 20 minutes. The transfection mix was then added to previously aspirated flasks. Final transfection volume was 11.7mL/T75 cm^2 flask. Transfected cells were then incubated at 37°C in 5% CO_2 for 48 hours.

Two days after transfection, cells were plated in 384-well plates for screening. In brief, the transfection mix was aspirated and cells were trypsinized using a .05% trypsin DMEM solution. Once cells were fully detached from the flask, the trypsin was neutralized using DMEM with 10% FBS. Cells were then centrifuged for 5 minutes at 1000rpm and the medium was aspirated. Cells were then resuspended in DMEM without phenol red (GIBCO 21063-029) containing 10% NuSerum (BD 355500) and counted with a hemocytometer. Cells were then plated at 8,000 cells/well (in 30µL per well) in 384-well plates (Cell Star T-3037-17) using a Multidrop Combi liquid dispenser (Thermo Scientific). Plated cells were then incubated overnight at 37°C in 5% CO_2.

The following day 100nL of individual compounds in DMSO were added to each well using a pin transfer apparatus. Bursicon conditioned media was diluted 1:33 with DMEM without phenol red. Following pin transfer of compounds, 30μL of bursicon conditioned media was added to each well resulting in a final 1/66 dilution of bursicon conditioned media. For MC4R stimulation, a stock of α-MSH peptide (H-1075, Bachem) in dH_2O (3mM) was diluted in serum free DMEM without phenol red to a working concentration of 20nM. Diluted peptide (30μL) was added to the cells using an 8 channel multidispense pipette (Biohit) to obtain a final concentration of 10nM. For both rk and MC4R, cells were then incubated for 4 hours with agonist. For a positive control (representing receptor inhibition), serum free DMEM without phenol red was added to alternating wells in columns 1, 2, 23, and 24 (no bursicon or α-MSH was added to these wells). The remaining wells in the columns were used as stimulated compound control wells (DMSO alone). The wells were stimulated with bursicon conditioned media or α-MSH without candidate compounds (representing no receptor inhibition). The layout of each plate is shown in Figure 3.11.

Four hours following the addition of bursicon conditioned media, 30 μL of steady light luciferase reagent (Perkin Elmer) diluted 1:2 in dH_20 was added to each well. Luminescence was then measured. Data were analyzed using Graphpad software. All data was expressed as a percentage of stimulated control wells (DMSO alone), defined as 100%.

Bursicon conditioned media

Bursicon conditioned media (CM) was made in the same way as described in Materials and Methods of Chapter 1. All batches of conditioned media were validated by standard concentration response curves to ensure maximum stimulation levels could be reached. For the Broad MLPCN screen enough bursicon conditioned media for the entire screen was produced in a single batch. In order to make the required 2 liters of conditioned media, Triple layer T175cm^2 flasks (353143, BD Biosciences) were transfected rather than the standard T75 cm^2 flasks. For Triple flasks, all transfection amounts were scaled up based on effective growth surface area. cDNA amounts per flask were: Bur-α pcDNA3.1 (27μg/flask); Bur-β pcDNA3.1 (27 μg/flask); PEI (160 μL per flask of a 1 mg/mL stock). Final volume was 84mL serum free DMEM per flask. Following a 24 hour transfection, media was aspirated and 72mL per flask of fresh serum free DMEM was added. Bursicon CM was then collected after 48 hours and centrifuged as previously described. Aliquots of media were made and frozen at -80°C. Validation of media made for screening can be seen in Figure 3.10.

1536-well plate adaptations

For the 1536-well plate screen of the MLPCN library conducted at The Broad Institute the following changes were made to the previously described 384-well plate assay:

Transfections

Transfections were done in T175 cm^2 flasks. Methodology for transfection was the same as for 384-well plate assay with the following changes: rk pcDNA1.1

(2.9µg/flask), destabilized CRE luciferase reporter (14.7 µg/flask), and PEI (61µL of a 1mg/mL stock) were used per flask.

Cell Plating

Transfected cells were plated in MaKO™ 1536-well Plates (Aurora Biotechnologies) two days after transfection. 2000 cells/well in 5µl of DMEM + 10% NuSerum were plated using a ViaFill system (Integra Biosciences, Hudson, NH).

Compound Transfer

The day following cell plating, experimental compounds (5nL/well, 10mM in DMSO) were added to plates using a Echo liquid handling system (Labcyte, Sunnyvale California). Final compound concentration was 10µM.

Bursicon Conditioned Media

Frozen aliquots of bursicon CM were thawed and filtered through a .22µM filter to remove any residual particulates remaining after media conditioning. Filtered bursicon conditioned media was then diluted 1/5 in serum free DMEM and 1µl/well of media (1:30 final dilution) was added to plates using a CombiNL liquid handling system (Thermo Scientific).

Luciferase assay

Four hours following stimulation with bursicon conditioned media, 1µl/well of steady glow was added and plates were incubated at RT for 10 minutes. Luminescence was then read using a ViewLux (Perkin Elmer) luciferase plate reader.

GPR-rk Activity Assay

Activity assessment of the *Anopheles gambiae* ortholog of rk was done as previously described in Materials and Methods of Chapter 1. GPR-rk (AGAP008347) was cloned into pcDNA1.1 and transfected into HEK293 cells. To make *A. gambiae* bursicon conditioned media, the same methodology was used as described in Chapter 1; however *A. gambiae* orthologs of Bur-α (AGAP002537) and Bur-β (AGAP004506) subunits were cloned into pcDNA1.1 and used in place of *Drosophila* bursicon subunits.

Chapter 4

A systematic strategy expedites the conversion of low affinity peptides to high potency anchored ligands

Abstract

Novel strategies are needed to expedite the generation and optimization of peptide probes targeting G protein-coupled receptors (GPCRs). As an initial step we previously showed that membrane tethered ligands (MTLs), recombinant proteins comprised of a membrane anchor, an extracellular linker, and a peptide ligand, could be used to identify targeted receptor modulators. Although MTLs provide a useful tool to identify and/or modify functionally active peptides, a major limitation of this strategy is its reliance on recombinant protein expression. We now report the generation and pharmacological characterization of prototype peptide-linker-lipid conjugates, soluble membrane anchored peptide ligands (SMALs), which are designed as mimics of corresponding MTLs. In this study, we systematically compare the activity of selected peptides as MTLs versus SMALs. As prototypes, we focused on the precursor proteins of mature substance P (SubP) and cholecystokinin 4 (CCK4). As low affinity soluble peptides these ligands each presented a challenging test case for assessment of MTL/SMAL technology. We report that both MTLs and corresponding SMALs show comparable receptor activity and subtype selectivity. In addition, our results illustrate that membrane anchoring increases ligand potency. Furthermore, both MTL and SMAL induced signaling can be blocked by specific non-peptide antagonists suggesting that the anchored constructs are orthosteric agonists. In conclusion, MTLs offer a streamlined approach for identifying peptides which are readily converted to SMALs. The ability to recapitulate MTL activity with SMALs provides an extended range of opportunities to utilize anchored peptides as functional probes and eventually as therapeutics.

Introduction

The development of peptide ligands has increased over the past two decades in parallel with an expansion in the diversity of corresponding therapeutic targets, e.g. ion channels, pumps/transporters, enzymes, G protein-coupled receptors (GPCRs) (Kaspar and Reichert 2013). Among these protein classes, GPCRs are the most widely explored targets with over 39% of peptides in the clinical pipeline interacting with these receptors (Kaspar and Reichert 2013). Although GPCRs remain one of the most important drug targets, only a small fraction of these receptors have been exploited by marketed therapeutics (Wise, Gearing et al. 2002; Schlyer and Horuk 2006). Therefore, novel strategies to activate or block GPCRs are needed as tools to probe the corresponding physiological functions as well as to validate potential drug targets.

We have previously reported that membrane tethered ligands (MTLs) offer a novel approach to modulate GPCR activity both *in vitro* and *in vivo* (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Fortin, Chinnapen et al. 2011; Harwood, Fortin et al. 2013). An MTL complementary DNA (cDNA) encodes a single protein which includes a peptide ligand (localized outside the cell), an intervening linker, and a transmembrane domain (TMD) anchor (Figure 4.1A). Although the MTL approach allows ligands and subsequent modifications to be studied without the need for traditional peptide synthesis, a major limitation of this strategy is that it relies on delivery and expression of cDNA. The ability to optimize a construct using the MTL system and then deliver the corresponding peptide ligand directly to a target tissue as a synthetic soluble membrane anchored ligand (SMAL) would open a range of new possibilities for both *in vitro* and *in vivo* use of such experimental probes. We postulated that a relatively simple synthetic

construct, a lipid-linker-peptide conjugate (Figure 4.1B and Figure 4.2), would anchor in the membrane and mimic the activity of the corresponding recombinant MTL. In addition to assessing this hypothesis, we have explored whether this strategy could provide an efficient platform for converting MTLs to deliverable synthetic SMALs.

As prototype ligands, we focused on two peptides, substance P (SubP) and cholecystokinin 4 (CCK4), both well characterized neuroendocrine hormones that activate selected cognate receptor subtypes. SubP is an eleven amino acid peptide that activates 3 neurokinin receptor subtypes: NK1, NK2, and NK3 (Leeman and Ferguson 2000; Bellucci, Carini et al. 2002). CCK4 is a tetrapeptide fragment of cholecystokinin that preferentially activates the cholecystokinin receptor subtype 2 (CCK2R) versus receptor subtype 1 (CCK1R) (Hughes, Boden et al. 1990; Ito, Matsui et al. 1993). The processing of CCK and SubP are similar with each existing as a C terminal glycine extended pro-hormone. Following cleavage of the glycine residue, peptidylglycine α-amidating monooxygenase catalyzes the addition of a C-terminal amide group, thought to be important as both an affinity and efficacy determinant (Eipper and Mains 1988; Eipper, Stoffers et al. 1992; Cuttitta 1993; Eipper, Milgram et al. 1993). During the course of our initial pilot studies with these two peptides, we observed that non-amidated SubP and glycine extended CCK4 both demonstrated significant agonist activity as MTLs (4.3 and 4.5). These reagents provided tools to systematically examine how the pharmacological features of low potency soluble peptides are altered with membrane anchoring. Therefore we compared recombinant MTLs with corresponding SMALs. Pharmacological features that were explored included receptor mediated activity, subtype

specificity, and the susceptibility to inhibition by known antagonists of corresponding free peptides.

Our results suggest that MTLs offer an expedited approach to screen for peptides that will be active as SMALs. Once identified as an active MTL, SMALs offer a more feasible mode of delivery *in vivo*. This combination of technologies may be utilized to enhance the identification and optimization of a novel class of GPCR probes (MTLs) that can be easily administered (as SMALs) to define the effects of tissue selective receptor modulation.

Materials and Methods

Cell Culture and Transfections

Human embryonic kidney cells (HEK293) were maintained at 37 °C in a humidified 5% CO_2 atmosphere and cultured with Dulbecco's modified Eagle's medium (Invitrogen, Chicago, IL) containing 10% fetal bovine serum, 100 U/mL penicillin, and 100μg/mL streptomycin. Cells were seeded into 96-well plates and grown to ~80% confluence. Cells were transfected for 24 hours using polyethylenimine (Sigma, Atlanta, GA) in serum-free medium (Doyle, Fortin et al. 2012) with cDNAs encoding, a) tethered ligand (where noted), b) 3 ng of indicated receptor, c) 25ng of pGL4.33 a luciferase reporter gene under the control of a serum response element (Promega, Madison, WI), and d) 5 ng of β-galactosidase to control for transfection efficiency.

Plasmids

Neurokinin receptors were purchased from the Missouri S+T cDNA Resource center (Rolla, MO). The CCK2 receptor was cloned as previously reported (Lee, Beinborn et al. 1993); the CCK-1R cDNA was PCR amplified based on a published sequence (Pisegna, de Weerth et al. 1994). Each receptor cDNA was subcloned into pcDNA1.1. Tethered SubP and CCK4 constructs were generated using a MTL with a type II transmembrane domain as a template cDNA which results in a free extracellular C-terminus of the corresponding peptide (Figure 4.1) (Harwood, Fortin et al. 2013). Oligonucleotide-directed, site-specific mutagenesis was used to introduce sequences encoding the following peptides with a free carboxy terminus: SubP, RPKPQQFFGLM and CCK4, WMDF. For the glycine extended construct, the corresponding oligonucleotide encoded

an additional glycine residue at the C-terminus of the CCK4 peptide (i.e. WMDFG). The nucleotide sequences of all receptor coding regions and tethered ligands were confirmed by automated DNA sequencing and analyzed using Vector NTI software (Invitrogen, Chicago, IL).

Peptides

All lipidated and non-amidated peptides were synthesized in the laboratory of Dr. Krishna Kumar at Tufts University, Medford Ma. Peptides illustrated in Figure 4.2 were synthesized using the *in-situ* neutralization protocol for *t*-Boc chemistry (Schnolzer, Alewood et al. 1992) on PAM resin on a 0.5 mmol scale. Amino acids were used with the following side chain protecting groups: Arg(Tos), Asp(OBzl), Gln(Xan), Lys(Fmoc), Lys(2-Cl-Z) and Trp(For). Peptide coupling reactions were carried out with a 4-fold excess (2.0 mmol) of activated amino acid for at least 15 min. The *t*-Boc protecting group on the *N*-terminus was removed using trifluoroacetic acid (TFA). The PAM resin from the CCK4 peptide synthesis was split into two equal portions. One portion of the resin was used for synthesizing non-lipidated peptides. The CCK4 (s-CCK-Gly-COOH) peptide was left unmodified at the *N*-terminus. This peptide served as positive control for the lipidated counterparts. The second portion of the CCK4 peptide and the SubP peptide were modified on resin as follows to yield test lipidated peptides (l-SubP-COOH and l-CCK4-Gly-COOH). Spacers (these are AA's used between the peg linker and the peptide of interest) were introduced on the peptides before pegylation (KGG for SubP and GG for CCK4). The free *N*-terminus of the peptide on resin was first pegylated with *N*-Fmoc-PEG8-propionic acid using standard HBTU coupling conditions. The *N*-Fmoc protecting

group was removed by treatment with 10% piperidine in DMF (N,N-Dimethylformamide) for 5 min. Palmitic acid was subsequently coupled with the *N*-terminal free amine of the pegylated peptide. Peptides were cleaved from the resin using high HF conditions (Pennington 1994) with minor modifications to the usual procedure. For the SubP peptide, longer times were used to ensure removal of Arg(Tos) protecting group (90% anhydrous HF/10% anisole at 0 °C for 2 h). For the CCK-4 peptides, 1,3-propanedithiol was used in the HF cleavage mixture to ensure deprotection of the formyl protecting group and prevent oxidation of methionine to its sulfoxide derivative: 85% anhydrous HF/10% anisole/5%PDT (1,3-propaneithiol) at 0 °C for 2 h) (Matsueda 1982). Following cleavage from resin, peptides were precipitated with cold Et_2O. Unmodified peptides were extracted using 10% AcOH in water and the lipidated peptides were extracted using 10% AcOH in H_2O followed by 10% AcOH in 50% EtOH/H_2O. Crude peptides were purified by RP-HPLC [Vydac C18, 10 μm, 22 mm x 250 mm]. The purities of the peptides were assessed by analytical RP-HPLC [Vydac C18, 5 μm, 4 mm x 250 mm]. The molar masses of peptides were determined by MALDI-TOF MS. The CCK4 peptide concentrations were determined using tryptophan absorbance ($\varepsilon = 5580\ M^{-1}\cdot cm^{-1}$ at 278 nm) (Gill and von Hippel 1989) and concentration of the lipidated SubP peptide was measured using amino acid analysis performed at the Molecular Biology Core Facility at Dana-Farber Cancer Institute, Boston, MA. The lipidated SubP peptide (l-SubP-COOH) included a KGG spacer coupled to the *N*-terminus to allow attachment of the corresponding PEG8/palmitic acid. In comparison, the lipidated CCK4 analog (l-CCK-Gly-COOH) contains only a GG spacer used for subsequent anchoring. A general

scheme of lipidated peptides is illustrated in Figure 4.1B. Detailed chemical structures, purities, and molecular weights of the synthetic peptides are shown in Figure 4.2.

Assessment of Ligand Activity

Tethered agonist induced signaling was assessed in HEK293 cells 24 hours after transfection. For soluble and lipidated peptides, 20 hours following transfection, cells were stimulated for an additional 4 hours. For antagonist assays, CP 99994 or YM022 (Tocris, Minneapolis, MN) were added concurrently with soluble agonist for 4 hours. With tethered ligands, antagonists were added 4 hours after transfection; activity was assessed following an additional 20 hour incubation. Quantification of luciferase and β-galactosidase activities were performed as previously described (Fortin, Chinnapen et al. 2011). Data were analyzed by nonlinear curve fitting using Graph Pad Prism 5.0 software.

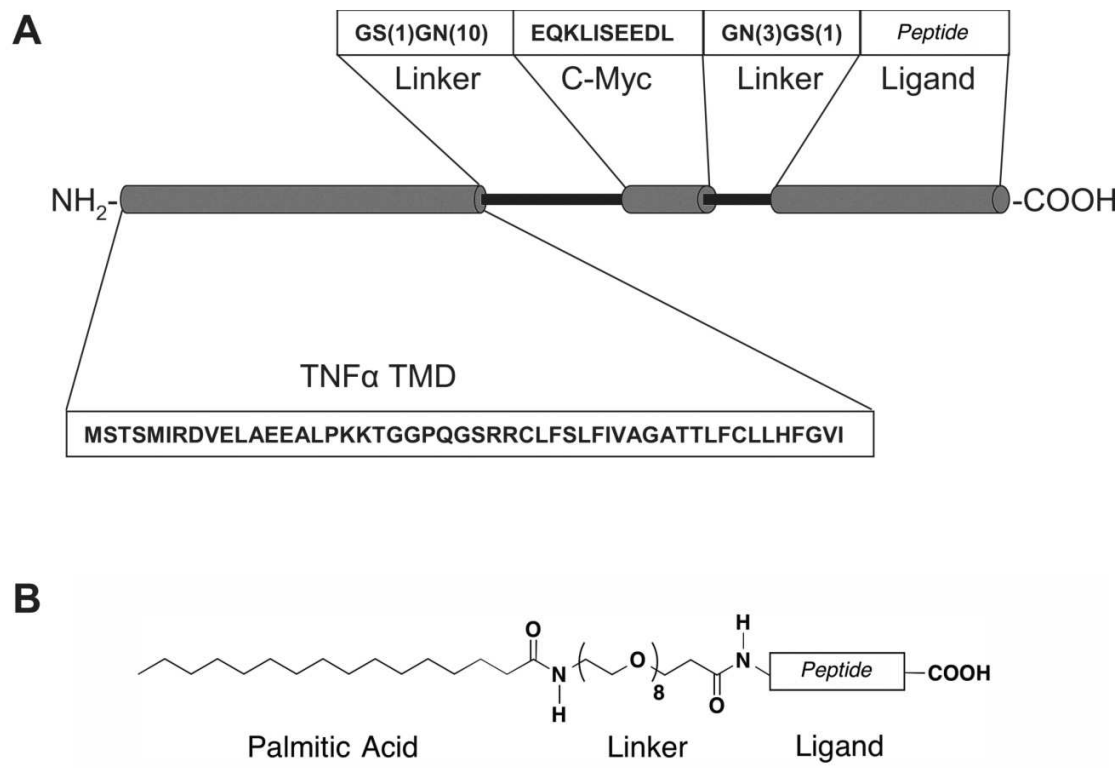
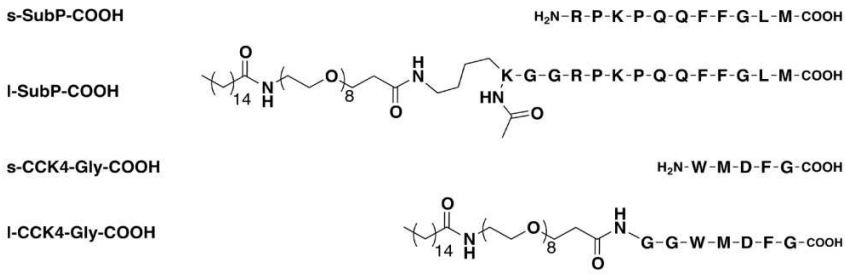


Figure 4.1. Cartoon models depicting A) recombinant membrane tethered ligand and B) corresponding soluble membrane anchored ligand.
Abbreviations: TNFα TMD =Tumor necrosis factor α transmembrane domain; PEG= polyethylene glycol; amino acids are represented in single letter code. Models were made by B. Harwood and S. Krishnaji.



Peptides	HPLC Purity[a] (%)	Molecular Weights (Da)	
		Calculated[b]	Observed[c]
l-SubP-COOH[d]	98	2294.8	2295.1 $[M+H]^+$
s-CCK-Gly-COOH	99	654.7	655.5 $[M+H]^+$
l-CCK-Gly-COOH[e]	96	1430.7	1453.3 $[M+Na]^+$

Figure 4.2. Chemical structure, purity, and molecular weight of synthesized peptides
A) Purity as determined by analytical RP-HPLC [Vydac C18, 5 μm, 4 mm × 250 mm] using a binary solvent system [A: $H_2O/CH_3CN/TFA$ (99/1/0.1); B: $CH_3CN/H_2O/TFA$ (90/10/0.07)] with a linear gradient of 65-80% solvent B over 20 min. The flow rate was set at 1 mL/min and elution was monitored by absorbance at 230 nm.
B) Expected molecular weights were calculated using Peptide mass calculator v3.2 and confirmed by the analysis tool in ChemBioDraw Ultra v12.0.3.
C) Observed molecular weights as determined using MALDI-TOF MS in reflectron positive mode using α-cyano-4-hydroxycinnamic acid as the matrix.
D) KGG and E) GG spacer coupled to the *N*-terminus of the peptide before pegylation. Analysis performed by S. Krishnaji.

Results

Although amidated SubP and CCK4 are well characterized peptides, fewer studies have examined the activity of precursor forms of these hormones. These pro-peptides as low potency ligands provide a useful framework to investigate how membrane anchoring can modulate peptide activity. On this basis, we have focused this study on elucidating the pharmacological properties of non-amidated SubP and glycine extended CCK4, as freely soluble peptides versus anchored counterparts.

We initiated our study with investigations focused on non-amidated SubP (SubP-COOH) as a recombinant MTL (tSubP). Activity of this construct was examined on each of the three human neurokinin receptor subtypes. When coexpressed with either NK1 or NK3 receptor, tSubP led to a concentration dependent increase in receptor mediated signaling (Figure 4.3A and C) whereas tSubP did not activate the NK2R (Figure 4.3B). In contrast, as a freely soluble ligand, s-SubP-COOH activated only the NK1R. Efficacy/potency comparisons were then carried out using a corresponding SMAL, a SubP peptide with the addition of a PEG linker and a palmitic acid at the amino terminus, i.e. lipidated SubP-COOH (l-SubP-COOH). This synthetic lipidated peptide mimicked the pharmacological activity of its genetically engineered tethered counterpart (tSubP). Both NK1 and NK3 receptors were activated by l-SubP-COOH (Figures 4.3A and C). When assessed at the NK2R, no signaling was observed (Figure 4.3B). Comparison of soluble and lipidated-SubP-COOH at the NK1R (Figure 4.3A) revealed that the lipidated analog had enhanced potency; corresponding EC_{50} values are as follows: l-SubP-COOH (EC_{50}=6.1nM) and s-SubP-COOH (EC_{50}=443nM). To further probe the pharmacological properties of anchored ligand induced receptor activation, we assessed the efficacy of a

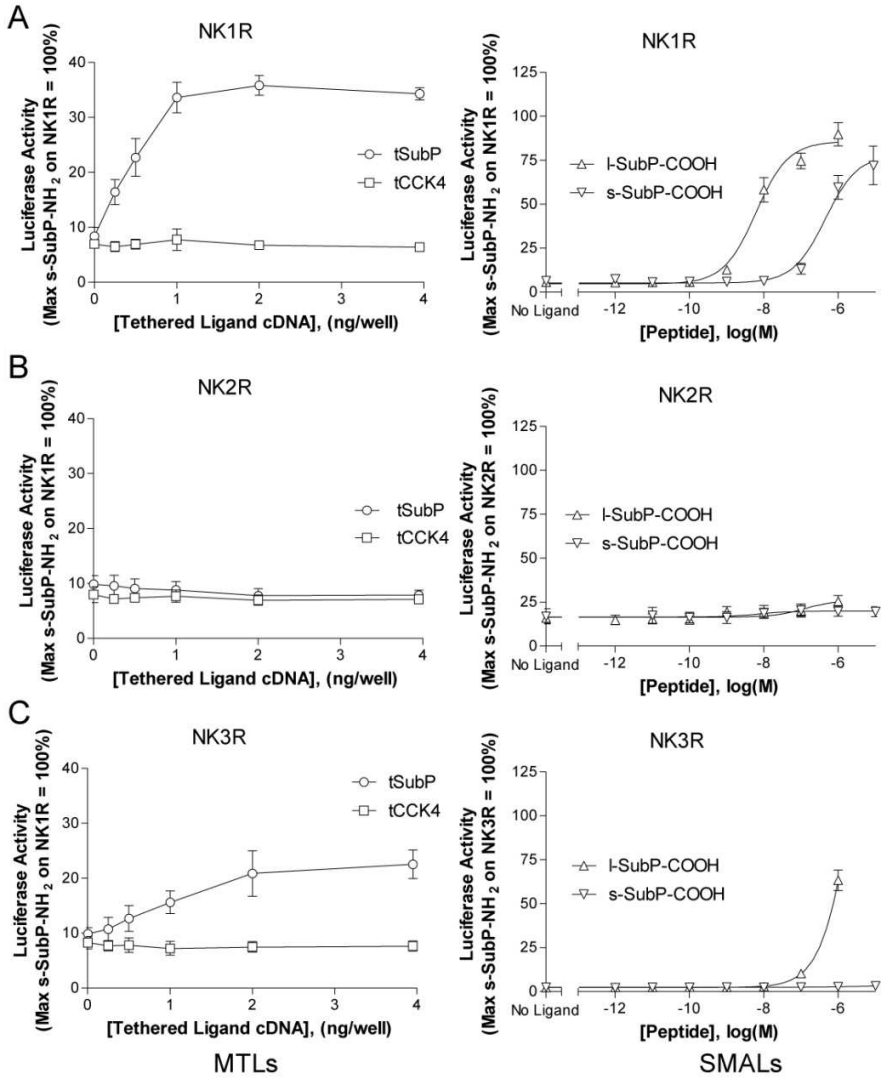


Figure 4.3. Recombinant SubP MTL mediated signaling predicts activity of a corresponding SMAL on neurokinin receptors.
Both tSubP and l-SubP-COOH activate NK1R (A) and NK3R (C) with no observed activity at NK2R (B). HEK293 cells were transiently cotransfected for 24 hours with cDNAs encoding: the designated NK receptor subtype, a 5X-SRE-Luc-pest reporter gene (pGL4.33), tethered ligand (for MTL assays, left panels), and a β-galactosidase gene to control for transfection variability. For assessment of SMAL induced signaling, cells were stimulated with ligand for 4 hours. Luciferase activity was quantified and normalized relative to a 4 hour stimulation with 1μM soluble amidated substance P (s-SubP-NH_2) on the corresponding NK receptor subtype. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tSubP, tethered substance P; tCCK4, tethered CCK4; s-SubP-COOH, soluble substance P with a C-terminal free acid; l-SubP-COOH, lipidated substance P with a C-terminal free acid; NK1R, neurokinin 1 receptor; NK2R, neurokinin 2 receptor; and NK3R, neurokinin 3 receptor. Experiments performed by B. Harwood and J. Doyle.

non-peptide inhibitor to block NK1R mediated signaling. CP 99994, a small molecule neurokinin receptor antagonist (McLean, Ganong et al. 1993; Lindstrom, von Mentzer et al. 2007), inhibited signaling by freely soluble as well as anchored SubP, either as a recombinant MTL or a SMAL. As illustrated in Figure 4.4A, tSubP activity was inhibited with an IC_{50} of 69.5nM. Agonist activity of l-SubP-COOH and s-SubP-COOH were also effectively blocked by CP 9994 (Figure 4.4B) with IC_{50} values of 18.0nM and 6.7nM, respectively.

In addition to studying how membrane anchoring influences SubP activity, we also focused on a second low affinity precursor peptide, glycine extended CCK4 (CCK4-Gly). Like SubP, the mature CCK4 peptide is also endogenously α amidated. CCK4-NH_2 is a well-established CCK2R ligand (Ito, Matsui et al. 1993; Lee, Beinborn et al. 1993). In preliminary experiments, membrane tethered CCK4 (tCCK4), minimally activated CCK2R. In contrast to tCCK4, we noted that with addition of a C-terminal glycine residue (tCCK4-Gly), activity of this construct at the CCK2R significantly increased. As illustrated in Figure 4.5A, tCCK4-Gly activated the CCK2R in a concentration dependent manner. In contrast, this construct showed no activity on the CCK1R (Figure 4.5B). To determine if the activity of the corresponding lipidated peptide would again (as with SubP) parallel the signaling observed with the tethered ligand, we next tested signaling induced by lipidated, glycine extended CCK4 (l-CCK4-Gly-COOH). As with tethered glycine extended CCK4, l-CCK4-Gly-COOH activated the CCK2R (Figure 4.5A) and lacked activity at the CCK1R (Figure 4.5B). Furthermore, lipidation of CCK4-Gly increased the potency of this ligand when compared with its soluble counterpart (s-CCK4-Gly-COOH) at the CCK2R. To further explore the

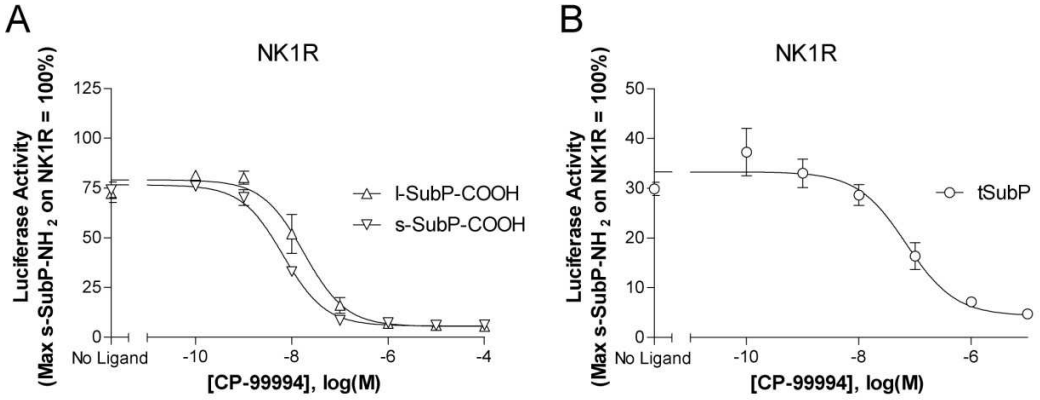


Figure 4.4. CP 99994 inhibits NK1R signaling induced by either a recombinant SubP MTL, soluble SubP with a C-terminal free acid (s-SubP-COOH), or the corresponding SMAL (l-SubP-COOH).
A small molecule, CP 99994, inhibits NK1R activation by tSubP (A), s-SubP-COOH and l-SubP-COOH (B). HEK293 cells were transiently cotransfected with cDNAs as outlined in Methods. For tSubP experiments (A), 4 hours following transfection, cells were treated with increasing concentrations of CP 99994 for 20 hours. For s-SubP-COOH and l-SubP-COOH experiments (B), 20 hours after transfection cells were treated with increasing concentrations of CP 99994 and 1µM of indicated soluble ligands for an additional 4 hours. Luciferase activity was quantified and normalized relative to a parallel preparation of NK1R expressing cells stimulated for 4 hours with s-SubP-NH_2 (1 µM). Abbreviations: tSubP, tethered substance P; s-SubP-COOH, soluble substance P with a C-terminal free acid; l-SubP-COOH, lipidated substance P with a C-terminal free acid; NK1R, neurokinin 1 receptor. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Experiments performed by J. Doyle.

mechanism underlying agonist mediated signaling, we evaluated the potential of a well-established CCK2R non-peptide antagonist, YM022 (Nishida, Miyata et al. 1994; Beinborn, Chen et al. 1998), to block receptor activation. As illustrated in Figure 4.6, YM022 inhibits CCK2 receptor signaling induced by tethered CCK4-Gly-COOH (Figure 4.6A) as well as soluble and lipidated CCK4-Gly-COOH (Figure 4.6B). IC_{50} values are as follows: tCCK4-Gly (IC_{50}=0.54 nM), l-CCK4-Gly-COOH (IC_{50}=10.2 nM), and s-CCK4-Gly-COOH (IC_{50}=0.84 nM).

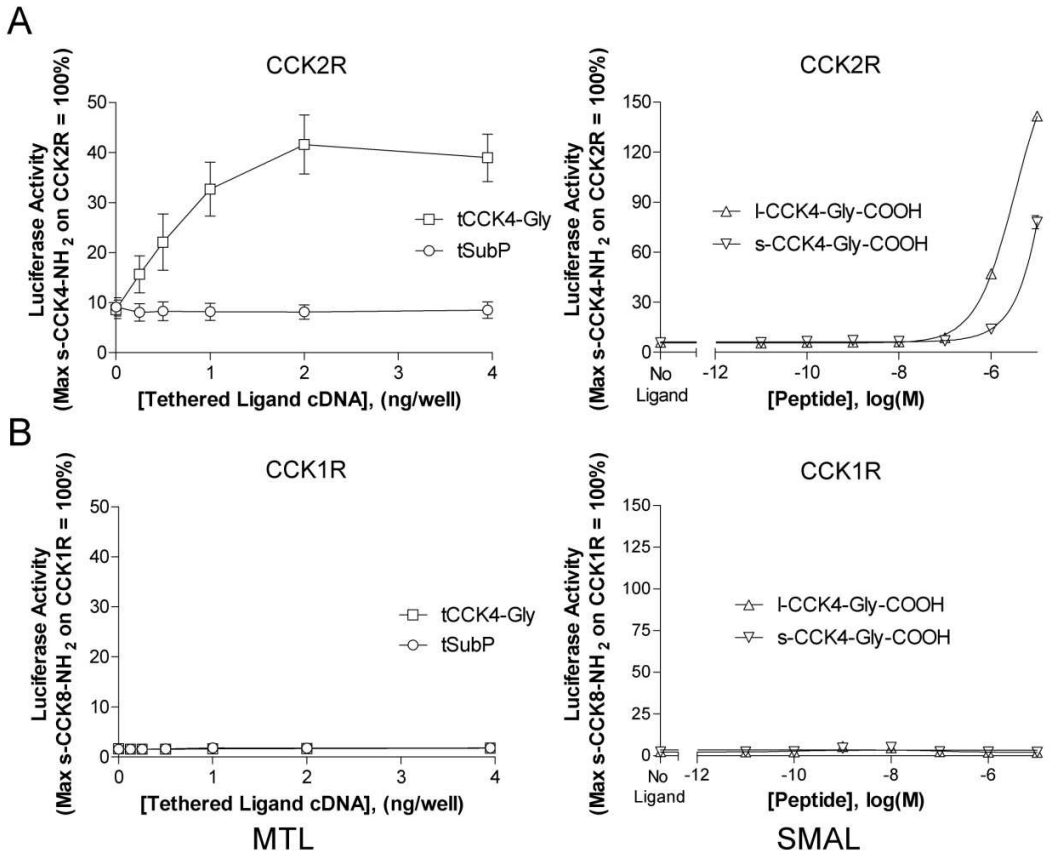


Figure 4.5. Tethered glycine extended CCK4 (tCCK4-Gly) anticipates activity of the corresponding soluble membrane anchored ligand.
A) Tethered CCK4-Gly activates the CCK2 receptor (left panel). Potency of the corresponding lipidated SMAL (l-CCK4-Gly) exceeds that of soluble CCK-4-Gly (right panel). B) CCK4-Gly as a tethered (left panel), soluble or lipidated (right panel) ligand fails to activate the CCK1 receptor. HEK293 cells were transiently cotransfected with cDNAs encoding: the designated CCK receptor subtype, a 5X-SRE-Luc-pest reporter gene (pGL4.33), tethered ligand (as indicated) and a β-galactosidase gene to control for transfection efficiency. Tethered ligand activity was measured 24 hours following transfection. To assess activity of soluble and lipidated CCK-4-Gly, cells were stimulated for 4 hours with ligand. Both soluble and tethered ligand activity was quantified relative to a parallel preparation of CCK receptor expressing cells stimulated for 4 hours with soluble amidated CCK-4 (s-CCK4-NH_2, 10μM) for CCK-2R and soluble amidated CCK-8 (s-CCK8-NH_2, 10μM) for CCK-1R. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tCCK4-Gly, tethered glycine extended CCK4; tSubP, tethered substance P; s-CCK4-Gly-COOH, soluble glycine extended CCK4 with a C-terminal free acid; l-CCK4-Gly-COOH, lipidated glycine extended CCK4 with a C-terminal free acid; CCK2R, cholecystokinin 2 receptor; CCK1R, cholecystokinin 1 receptor. Experiments performed by B. Harwood and K. Gao.

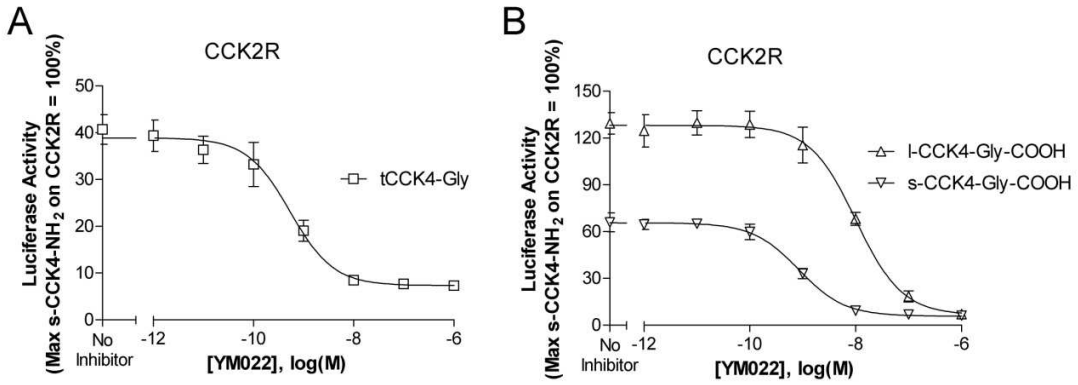


Figure 4.6. YM022 inhibits CCK2R signaling induced by tethered CCK4-Gly, soluble CCK4-Gly, or lipidated CCK4-Gly.
A) YM022 blocks tethered CCK4-Gly mediated CCK2R signaling. HEK293 cells were cotransfected with cDNAs encoding: CCK2R, a 5X-SRE-Luc-pest reporter gene, tCCK4-Gly (as indicated), and a β-galactosidase gene to control for transfection efficiency. Four hours following transfection, cells were with treated with increasing concentrations of YM022 for 20 hours. Luciferase activity was quantified and normalized relative to a parallel preparation of CCK2R expressing cells stimulated for 4 hours with soluble amidated CCK-4 (s-CCK4-NH_2, 10μM). B) YM022 blocks s-CCK-4-Gly-COOH and l-CCK4-Gly-COOH mediated activation of CCK2R. HEK293 cells were transfected as indicated above. Twenty hours after transfection, cells were with treated with increasing concentrations of YM022 together with either 10 μM of l-CCK4-Gly-COOH or s-CCK4-Gly-COOH. Following an additional four hour stimulation, luciferase activity was quantified and normalized as outlined for panel A. Data represent the mean ± SEM from 3 independent experiments, each performed in triplicate. Abbreviations: tCCK4-Gly, tethered glycine extended CCK4; s-CCK4-Gly-COOH, soluble glycine extended CCK4 with a C-terminal free acid; l-CCK4-Gly-COOH, lipidated glycine extended CCK4 with a C-terminal free acid; CCK2R, cholecystokinin 2 receptor.

Discussion

Our results demonstrate that relatively simple synthetic lipidated constructs mimic the activity of corresponding recombinant MTLs. This observation suggests a powerful two-step strategy that can be broadly applied to developing anchored peptide ligands for a wide range of targets. As a first step, an MTL with activity is identified. The recombinant nature of an MTL provides a highly efficient platform for generating and screening corresponding peptide variants should efficacy of the tethered ligand need to be fine-tuned. Once optimized using this strategy, the peptide is incorporated into a lipid-PEG linker backbone enabling expedited synthesis of soluble ligands that may be directly administered to a specific target tissue.

To test this two-step strategy, precursor forms of two well established peptide hormones, SubP and CCK4, were used. We noted that non-amidated SubP and glycine extended CCK4, respectively, showed agonist activity when assessed as MTLs (Figures 4.3 and 4.5). In light of the known importance of amidation for the function of numerous biologically active peptides (Eipper and Mains 1988; Kim and Seong 2001), including SubP and CCK4, the activity of the non-amidated precursor peptides was not anticipated (Eipper, Stoffers et al. 1992; Cuttitta 1993). This finding suggests that tethered precursor peptides may be active and that the requirement for post-translational modification does not necessarily preclude activity as an MTL. Thus, MTLs can provide a tool to facilitate the rapid identification of other active precursor peptides that can then be used as templates for further peptide optimization, the generation of recombinant transgenic activators, and/or as a template for SMALs.

The low potency of many precursor peptides, including CCK and SubP, is due in part to the absence of amidation as an affinity determinant (Eipper, Stoffers et al. 1992). We speculate that membrane anchoring, by virtue of holding the corresponding ligand in proximity to its cognate GPCR, bypasses the need for selected affinity determinants, in this case the C-terminal amide. For peptides where MTLs are active, anchoring appears to facilitate direct ligand-receptor interaction. The observed increase in potency of both SubP and CCK4 precursors with synthetic anchoring (i.e. lipidation) is consistent with this hypothesis.

Despite bypassing affinity determinants, anchored precursor proteins of SubP and CCK (either as MTLs or SMALs) show receptor subtype selectivity. Like tSubP, l-SubP-COOH activates NK1 and NK3R with no activity observed at the NK2R. This phenomenon is recapitulated with CCK4: tCCK4-Gly and l-CCK4-Gly-COOH both activate the CCK2R with no activity on CCK1R. In addition to illustrating receptor subtype selectivity, these data also highlight the fact that MTLs are good predictors of the activity of SMALs. This attribute of MTLs fits well with our assertion that MTLs provide an efficient system for identifying and optimizing peptides of interest and underscores the utility of MTL-SMAL technology.

The predictive nature of MTLs both with regard to activity and subtype selectivity make them powerful tools to detect low potency activators of GPCRs that may otherwise be missed using conventional screening techniques. As an example, whereas both anchored versions of the low potency SubP precursor (tSubP and l-SubP-COOH) activate the NK3 receptor, no signaling is observed with s-SubP-COOH. Generalizing from this illustration, if MTL technology were used to screen for low potency ligands, we can

anticipate the identification of additional agonists (in contrast to those identified when screening corresponding soluble ligands alone). These membrane tethered agonists could include precursor peptides which have not been post-translationally processed, an area of great interest when trying to match orphan receptors with putative ligands based on genome mining (Ozawa, Lindberg et al.). These examples highlight the potential utility of MTLs in identifying ligands for GPCRs of interest.

To better understand the mechanism underlying membrane anchored ligand activity, we completed a series of experiments using well established small molecule antagonists. Our studies utilized both genetically engineered and synthetic anchored ligands. Like their soluble counterparts, MTLs and SMALs act as orthosteric activators as opposed to allosteric modulators. With both SubP and CCK-4, all forms of ligand activity are inhibited by CP 99994 or YM022, respectively. The IC_{50} values for antagonism at both NK1R and CCK2R are in the nanomolar range, similar to those previously reported for inhibition of mature forms of SubP and CCK proteins, i.e. amidated peptides (McLean, Ganong et al. 1993; Nishida, Miyata et al. 1994). The ability to block SMAL activity with these highly selective antagonists further underscores the potential of anchored peptides as receptor specific functional probes.

Prior studies have examined the effects of N-terminal lipidation of the amidated cholecystokinin tetrapeptide, CCK4-NH_2, with a focus on enhancing membrane permeability. Both acetylation and/or caproylation of CCK4-NH_2 resulted in increased peptide stability, permeability and intestinal absorption (Tenma, Yodoya et al. 1993; Yodoya, Uemura et al. 1994; Setoh, Murakami et al. 1995; Fujita, Kawahara et al. 1998). In addition to CCK, lipidation has been utilized to modify a wide variety of other peptide

ligands (Zhang and Bulaj 2012). Such modifications have led to enhancing peptide stability (Myers, Yakubu-Madus et al. 1997; Cui, Webber et al. 2010; Bellmann-Sickert, Elling et al. 2011), prolonging half-life by facilitating binding to circulating albumen (Bhattacharya, Grune et al. 2000; Fujiwara and Amisaki 2008; Bellmann-Sickert, Elling et al. 2011), and/or targeting excretion to the liver rather than the kidney (Yuan, Wang et al. 2005; Wang, Hogenkamp et al. 2006; Bellmann-Sickert, Elling et al. 2011). Additional studies have shown that lipidation can improve intestinal absorption by increasing the lipophilic properties of a ligand (Yodoya, Uemura et al. 1994; Setoh, Murakami et al. 1995; Tanaka, Fujita et al. 1996; Yamamoto 1998; Yamamoto 2001). Based on existing literature, It appears that lipidataion of peptides may either increase or decrease affinity (Dasgupta and Mukherjee 2000; Zarandi, Varga et al. 2006; Bellmann-Sickert, Elling et al. 2011). MTL/SMAL technology complements and improves the above empiric approaches by providing a recombinant format in which to assess (and modify as needed) the pharmacological effects of anchoring. This may be done prior to SMAL synthesis to optimize ligand activity.

The combined MTL/SMAL approach provides a highly sensitive system for identifying putative receptor ligands. Without the use of an MTL approach, many low potency peptides could be missed during screening. MTLs also offer an index of how anchoring will effect peptide activity and provide a rational approach for defining the site of synthetic anchoring, e.g. lipidation at the N or C terminus. Active MTL peptides, once identified, can be converted to SMALs using standard synthetic chemical methods. These lipid-peptide constructs show enhanced potency versus the free ligand and have the added advantage that they can be directly administered using conventional delivery

methods. It can be anticipated that with libraries of cDNA encoded tethered peptides, it should be possible to identify novel peptides that can modulate receptors of interest. Given the power of MTL technology in peptide design as well as the potential for SMALs to enhance ligand potency and enable delivery, this combination of strategies is well-suited to expedite the development of peptide therapeutics.

Acknowledgements

We would like to thank the synapse neurobiology training program (BNH). We also thank Ci Chen for her help with tissue culture throughout the course of this research and Martin Beinborn for helpful discussions.

Footnotes:

a) This work was supported by the National Institutes of Health Heart, Lung, and Blood Institute [Grant T32-HL069770]; the National Institutes of Health Neurological Disorders and Stroke [Grant T32-NS061764]; and the National Institutes of Health General Medical Sciences [Grant R01-GM65000].

General Discussion

In my thesis we have utilized MTL technology as a tool to better understand the structure function relationships and biology of bursicon and rk. Extending from this foundation, we have screened for small molecule inhibitors of rk and explored the pharmacological properties of SMALs, synthetic constructs designed to mimic the advantages of MTLs. The discussion that follows highlights the advancements we have made in each of these research areas. We also frame each achievement in a historical context acknowledging the earlier research that enabled our efforts. Finally, we provide an overview of where each project may lead in the future.

Membrane tethered ligands in vitro

We have shown that MTL technology is a highly adaptable system enabling optimization of novel ligands to modulate GPCRs. Prior to initiating the work described in my thesis, all MTLs had been designed using a type I transmembrane domain (TMD). This configuration resulted in a tethered ligand with a free extracellular N terminus (Figure i.5: cartoon of tethered ligand). A Type I TMD derived from the human herpes virus 1 glycoprotein C. Type I transmembrane domains require a signal peptide for proper trafficking to the plasma membrane. The signal peptide of chymotrypsin was used in these constructs. Prior to bursicon, earlier versions of MTLs incorporated relatively small peptides; the longest being 39 amino acids (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Fortin, Chinnapen et al. 2011). In addition, all MTLs made prior to bursicon only included a single protein ligand (vs. a heterodimer). Building on the foundation of

existing MTLs, an initial goal of this thesis was to extend and explore the limits of this technology.

An initial question addressed in Chapter 1 was the feasibility of generating an active MTL using a Type II TMD. A Type II transmembrane domain has the opposite orientation as a Type I, resulting in a free extracellular C-terminus. Since many peptides require a free C-terminus for agonist activity, achieving success with this opposite orientation would greatly expand the range of peptides that could be converted into functionally active MTLs. For our prototype Type II MTL, the TMD sequence from Tumor necrosis factor α (TNFα) was used (Figure i.5; cartoon of tethered ligand). A convenient feature of Type II TMDs is that they do not require an additional signal peptide, as the transmembrane sequence itself serves this purpose (Marmenout, Fransen et al. 1985). In contrast, as previously mentioned, Type I TMDs require a signal peptide to be incorporated into the MTL sequence. Bursicon α and β subunits were the first ligands to show functional activity when incorporated in an MTL with a Type II TMD sequence.

The Type II TMD template developed for bursicon MTLs has since been successfully used to tether other ligands. For example, a series of membrane tethered chemerin constructs have been made using the Type II MTL template (data not yet published). Chemerin is an important peptide mediator of inflammation and is thought to be pro-resolving in asthma (Shimamura, Matsuda et al. 2009). We have also used the Type II MTL template to successfully develop membrane tethered chemokines. For example we have made a chemokine (C-C motif) ligand 20 (CCL20) MTL that activates the GPCR chemokine receptor 6 (CCR6). The success with CCL20 opens the possibility of

generating MTLs for a whole series of chemokines that are involved in various immunological responses. The Type II orientation was also used for substance P and CCK4 MTLs discussed in Chapter 4. For these two peptides, a free C-terminus is required for activity and therefore a Type II orientation was critical for MTL construction. The ability to flip the orientation of ligands by using either a Type I or a Type II TMD has greatly expanded the capabilities of MTL technology.

As previously mentioned all MTLs made prior to bursicon were with relatively small peptides (<39 amino acids). In comparison bursicon is both larger and more complex. Not including signal peptide sequences the bursicon α and β subunits are 141 and 121 amino acids in length, respectively. In addition to being much larger than anything previously constructed, bursicon was thought to require dimerization for activity. Given the ligand complexity, we wanted to determine if a heterodimeric ligand could be membrane tethered. Surprisingly, bursicon was amenable to tethering through several strategies. Co-expression of both bursicon subunits as independent MTLs led to rk mediated signaling. In contrast, neither single construct was active when expressed alone (Chapter 1). Even more remarkable, when bursicon was expressed as either an α–β fused heterodimer or a β–α fused heterodimer, both showed activity. These findings set the stage to generate heterodimer MTLs (Chapter 1). Since our success in generating an active bursicon MTL, other large ligands have been converted to MTLs by our group using parallel technology. Both the previously mentioned chemerin and CCL20 ligands are large (136 and 69 amino acids respectively); tethering each of these peptides led to a functionally active MTL agonist.

A second implication of bursicon MTLs retaining activity is that complex ligands such as other cystine-knot proteins may also be amenable to membrane tethering. The previously discussed series of glycohormones, LH, FSH, and TSH would all be interesting candidates. As discussed in Chapter 1, many of the advantages of membrane anchoring could be leveraged to study glycohormones. In addition to confirming that other cystine-knot proteins can be converted to MTLs, tethering glycohormones would also test whether heavily glycosylated proteins will retain activity when anchored. The cystine-knot heterodimer GPA2/GBP5 (the endogenous ligand for the LGR1 receptor) is another potential ligand that could be assessed using the bursicon MTL template (Sudo, Kuwabara et al. 2005). This is of particular interest because a recent study suggested the *Aedes egypti* ortholog of LGR1 regulates ion transport in the gut. Insect LGR1 function was previously unknown (Paluzzi, Vanderveken et al. 2014). We also recently cloned orthologs corresponding to the LGR receptor and GPA2/GBP5 ligands from *Anopheles gambiae*. The construction of a GPA2/GBP5 MTL might be useful to better understand the biology LGR1 receptors. In particular it could be used to determine if *Drosophila* LGR1 serves a similar function as is reported for mosquitoes.

The successful development of membrane tethered bursicon constructs has also led to other enhancements in MTL technology. Among these areas of improvement is the ability to monitor expression of MTLs. All tethered ligands initially contained an epitope tag, typically a C-myc tag. We have recently modified this sequence to an HA epitope with no loss of activity. This allows us to simultaneously monitor surface expression of two independent MTLs (Chapter 1). In Chapter 1 to enable live imaging, fluorescent proteins, either GFP or monomeric Cherry Fluorescent Proteins (CFP) have been used to

demonstrate localization of tethered construct at the cell membrane (Fortin, Chinnapen et al. 2011). This visualization tool has been used to detect MTL expression in selected CCAP/bursicon neurons (data not shown). In addition to optimizing detection, we have found that minor modification of MTLs can fine tune activity. Such modifications include altering the linker composition and/or length. The linker portions of membrane tethered ligands are constructed using repetitive amino acid sequences. Both glycine-asparagine (GN) and glycine-serine (GS) repeats have been used in constructs. We have shown that varying the number of linker repeats can be used as a tool to modulate activity of a given construct (data not shown). The ability to alter the linker lengths allows for a customizable system that can be tailored to optimize activity of each tethered ligand.

The ability to alter the sequence of MTLs using recombinant technology offers a major advantage in utilizing these constructs. Mutation-induced alterations in MTL amino acid sequence enable rapid identification of residues which define peptide activity. For example in Chapter 1, using this approach we converted a bursicon membrane tethered agonist to an antagonist. We were able to efficiently screen a series of MTLs which included varying bursicon domain deletions for loss of receptor activation. Once domains critical for receptor activation were identified, we were then able to test them for antagonist activity. Using this approach we identified the C-terminal tail of the bursicon α subunit as being important for receptor activation. The bursicon MTL deletion mutant was the first Type II tethered antagonist. A similar approach was also successful for Type I tethered exendin where two amino acids substitutions converted the corresponding MTL from an agonist to antagonist on the GLP1 receptor (Fortin, Zhu et al. 2009). Another advantage of the recombinant approach using MTLs is that they avoid the

bottleneck which would occur if individual proteins had to be synthesized and purified. For many of our studies, generating corresponding soluble peptides would be a highly labor intensive and cost prohibitive endeavor.

B***ursicon membrane tethered ligand in vivo***

Our *in vitro* based studies with MTLs have highlighted the versatility of these constructs. During the course of our structure-function studies with bursicon, we successfully generated a Type II MTL, tethered a complex heterodimeric ligand, improved the detection methods for MTLs, and generated a type II membrane tethered antagonist. As a next step we sought to utilize the bursicon MTL to explore the biology of the rk burs system using an in vivo model system.

Prior to the construction of bursicon MTLs, tethered pigment dispersion factor (PDF) had been generated and expressed in *Drosophila* and demonstrated to rescue arrhythmic PDF null flies (Choi et al. 2008). The tethered PDF construct included a Type I TMD, incorporated a relatively short (19 amino acid) peptide, and activated its cognate GPCR. As anticipated from *in vitro* studies, the construct showed agonist activity *in vivo*, resulting in partial rescue of the abnormal circadian rhythm observed in PDF null flies.

Bursicon, as a tethered construct to study *in vivo* function offered both challenges and advantages. Regarding challenges, the bursicon MTL was the first type II construct. In addition, the MTL incorporated two large heterodimeric cystine-knot proteins. At the same time the bursicon MTL offered clear phenotypic outcomes (viability, wing expansion, tanning) by which *in vivo* function could be assessed.

As described in chapters 1 and 2 we developed a bursicon membrane tethered ligand and then used it *in vivo* to better understand bursicon/rk signaling. Bursicon was especially appealing to help address the question of whether a membrane tethered ligand would work *in vivo* due to the fact that phenotypic outcomes of perturbation were hypothesized to be clear cut. Agonist activity was postulated to have no effect on wing expansion, tanning and survival. Antagonist activity was postulated to result in a wing phenotype, impaired tanning, and compromised survival. It was unanticipated that the bursicon tethered agonist when expressed *in vivo* gave phenotypes consistent with an antagonist. The bursicon MTL appeared to block rk signaling leading to wing expansion defects and lethality. Given this unanticipated result, we sought to define a potential mechanism that could explain the observed negative regulation of rk. Using an *in vitro* model, we found that rk appears to be readily desensitized upon chronic stimulation with either soluble bursicon or membrane tethered ligand based stimulation. This is the first example of an MTL that causes receptor desensitization *in vitro*. In contrast, tethered PDF does not result in receptor desensitization (Choi, Fortin et al. 2009). As a potential underlying mechanism, the desensitization of rk by the bursicon MTL is consistent with the phenotypes observed *in vivo*. This highlights the potential use of an *in vitro* cell based assay to model the *in vivo* function of MTLs.

Also as described in Chapter 2, we successfully used the bursicon membrane tethered ligand to better understand the tissue specific requirements of rk activation. In particular, our studies demonstrate a key role for rk in adult muscle during development and wing expansion. Using our bursicon MTL complemented by RNAi fly lines, we were able to efficiently survey various tissues using Gal4 drivers. Our results indicate

where rk is required for normal development. Multiple Gal4 drivers indicated that downregualtion of rk in muscle resulted in developmental arrest and wing expansion abnormalities. In addition, we have been able to demonstrate that our bursicon MTL can target muscle in trans across the neuromuscular junction (NMJ) when expressed in motor neurons. This model is consistent with previous studies that have shown bursicon is released from the NMJ (Loveall and Deitcher 2010). The activity of the bursicon MTL *in vivo* and the observations made using this construct during development, highlight the utility of this technology in advancing the understanding of bursicon/rk biology.

In addition to the insights gained, the consequences of bursicon MTL expression in *Drosophila* has also led to additional questions about rk biology. For example, now that muscle has been implicated as a key tissue, determining which muscle groups in the fly actually require rk activation to help promote eclosion and wing expansion is of interest. Careful histological studies are required to help determine which muscles are affected by downregualtion of rk. Such studies could explore the hypothesis posed in Chapter 2 that leg and abdominal muscle may require rk expression for proper development. Histological analysis of these muscles in conjunction with tethered ligand expression or rk RNAi would determine if there are any clear structural defects. Understanding the molecular consequence of rk knockdown in muscle would also be of interest. This would help clarify the role of rk signal transduction during normal adult *Drosophila* muscle development.

A second area of interest where bursicon MTLs could be used is to address the confounding observation that Bur-α is sometimes expressed without Bur-β being present. Multiple studies have confirmed and reinforced that bursicon only activates rk as a

heterodimeric ligand (Luo, Dewey et al. 2005; Mendive, Van Loy et al. 2005). Given the well studied role of the bursicon heterodimer, could Bur-α alone have an alternative function independent of the β subunit? This hypothetical function remains unknown. One theory is that bursicon homodimers are able to activate NF-κB signaling to promote an immune response (An, Dong et al. 2012). However it is still unclear whether this occurs through rk or another receptor. Using our *in vitro* assays, our lab has conducted multiple experiments using both with MTLs and soluble subunits in varying combinations. Applying these tools, we have not been able to observe any $G\alpha_s$ dependent rk signaling (data not shown). However with the appropriate reporters, the assays used throughout this thesis could potentially be modified to test rk based NF-κB signaling in the future.

Our *in vitro* and *in* vivo studies have helped to better characterize the role of bursicon and rk in insects and further establish rk as a crucial molecule in *Drosophila* development. Given the success of membrane tethering we feel that there are many future applications of this technology and we have only begun to leverage their utility both *in vitro* and *in vivo***.**

Rk as an insecticide target

Currently there is a need for novel insecticides. As discussed in Chapter 3, insecticide resistance is an ongoing problem. With only a limited number of targets for current insecticides, there is an urgency to identify novel targets. Notably, no GPCR has been successfully leveraged as an insecticide target despite these proteins having an excellent track record as being highly druggable.

In vitro and *in vivo* studies conducted in chapters 1 and 2 highlight the fact that rk is signaling may be modulated by a variety of approaches (e.g. administration of soluble ligand, expression of tethered ligand, or RNAi). *In vivo* work with rk also suggests that this receptor is crucial for insect development. Experiments with both MTLs and RNAi suggest that disruption of rk signaling results in developmental lethality. In light of these results, rk appears to be an attractive insecticide target. By modifying assays that we used to assess rk mediated signal transduction (Chapter 1), we were able to establish a high throughput screening protocol. Using this cell based assay we utilized two strategies to identify an rk antagonist. The first approach described in Chapter 3 was a natural product screen. Using this library, we screened 3,709 plant extracts and found 4 that potentially contain useful rk antagonists. Through subsequent analysis, the compound isolated from the pursued extract ultimately did not retain the specificity for rk that was desired. The 3 other extracts remain to be explored.

As a 2nd strategy we initiated a larger scale screen of synthetic small molecules to identify novel rk antagonists. As described in Chapter 3, a promising class of molecules has been identified. Work is in progress to validate these rk antagonists as potential insecticides. This validation can be broken down into a multi-step process. The first step is to test whether these template molecules have activity *in vivo*. To assess their activity these compounds are being fed to developing *Drosophila* larvae. Using the same phenotypes monitored in Chapter 2 (i.e. wing expansion defects and developmental arrest), we can determine if these compounds are active as *in vivo* probes. If no phenotypes are observed with these initial compounds, further optimization to enhance affinity and/or compound stability may be required.

In parallel with testing molecules *in vivo*, it is important to better understand how existing compounds and their structural derivatives block rk activation. Two potential mechanisms can potentially explain how these antagonists work. In the first scenario, the compounds bind to the rk receptor thereby disrupting the ability of rk to be activated by bursicon. For example a competitive antagonist works by binding to a receptor thereby preventing an agonist from activating the GPCR. In the second scenario the compound may bind or disrupt the bursicon, thus rendering it unable to activate rk. One example of how this could theoretically occur is if small molecule somehow disturbs the conformation of the heterodimeric cystine-knot thereby inactivating the ligand. Understanding the mechanism of antagonist activity will be useful for additional probe optimization.

Another important step in probe optimization is to determine the selectivity of this class of compounds. While it is very difficult to test compounds on all GPCRs for selectivity, a focused strategy can be applied. For example, compounds can be tested on other LGR receptors. Specifically as discussed in Chapter 3, it is important to test the human glycohormone receptors and other *Drosophila* LGRs for specificity. Since all of these GPCRs come from the same subfamily, there is a possibility that compounds may target a conserved domain within LGR receptors.

Testing rk orthologs from other insect species will also be informative to help understand compound selectivity. Having an understanding of species cross reactivity prior to compound optimization can be used as criteria for selectivity (e.g. an antagonist on the *Anopheles gambiae* ortholog GPR-rk would be most desirable if it lacked activity on the *Apis mellifera* ortholog).

Future directions for the identification of an rk antagonist fall into two categories, further assessment of natural products and synthetic small molecules. For the natural products, we have yet to find a promising compound. However we have established proof of principle that screening natural products can be a viable strategy to identify plant extracts that modulate insect GPCRs. This sets the stage to use natural products in two ways. The first is to follow up on the 3 other plant extracts that inhibit rk but have not been further characterized. Follow up efforts would aim to identify an rk specific active component from these fractions. The second possible path forward is to screen other natural product libraries on the rk receptor. The library of Chinese medicinal herb extracts that was tested in Chapter 3 is only one of many natural product libraries available. For example, the ICCB Longwood screening facility from where we received the extracts that were tested in our study, has a total of approximately 70,000 natural product extracts available. We only assessed a small portion of their collection. Screening more plant fractions both may still yield a naturally derived rk antagonist. This is a desirable approach because bringing a natural product to market as an insecticide has far fewer hurdles than a corresponding path for a synthetic chemical.

An alternative future direction is to continue with the assessment of hits found in the synthetic small molecule screen. Identification of lead compounds was a key first step in developing a novel insecticide. The next step is to optimize these hits. With this in mind, we have already tested 38 structural derivatives of our initial hits to help understand which domains of the chemical structure confer activity. Once we have maximally optimized compound affinity, other questions will be need to be addressed. Understanding the chemical stability and how the compound is degraded in the

environment will be important. Any toxic byproducts or potential off target effects will have to be explored. How easily the compound can be synthesized and how it is formulated for delivery must be established. All of these questions will require collaboration from experts in scientific disciplines including vector biology, chemistry, and environmental science.

Finally, the strategy used for screening rk is an approach that can be applied to other insect GPCRs. Given how amenable GPCRs are to modulation by small molecules, it is surprising that to date only one class of insecticides has targeted a GPCR. Compounds that modulate octopamine receptors are currently the only example (Verlinden, Vleugels et al. 2010). Based on RNAi studies in *Drosophila*, our lab has generated a list of putative GPCR targets where inhibition of signaling will compromise viability. These receptors include dopamine receptors, LGR1, the sex peptide receptor, and gustatory receptors. A similar strategy as pursued in the current study could be leveraged for each of these targets to develop molecules that block corresponding GPCRs. Identifying chemical modulators is again the first step toward validating these receptors as insecticide targets. Given major issues with the development of insecticide resistance in various disease vectors, having an array of insecticides to choose from is one of the keys to long term successful control (Kelly-Hope, Ranson et al. 2008).

Extensions of membrane tethered ligands: SMALs

As a complement to membrane tethered ligand technology we wanted to extend the strategy of ligand anchoring beyond a system that relies on recombinant DNA. With this goal mind, we focused our studies on peptides that are active as MTLs and could be efficiently synthesized as soluble peptides. Substance P and CCK4 both fit these criteria and were selected as candidate ligands for the development of corresponding soluble membrane anchored ligands.

In general, MTL technology has been a successful strategy to anchor a wide variety of GPCR ligands. Bursicon is one example of a series of MTLs that have been developed in the Kopin lab over the years (Choi, Fortin et al. 2009; Fortin, Zhu et al. 2009; Fortin, Chinnapen et al. 2011; Harwood, Fortin et al. 2013). The bursicon MTL project also exemplifies how membrane tethered constructs, even ones that incorporates complex heterodimeric ligands can be used *in vivo* as tissue specific probes. One advantage of MTLs for *in vivo* studies in model organisms is that this approach relies on recombinant constructs. This was exemplified by bursicon where the genetically tractable *Drosophila* system was an appropriate model for these studies. At the same time, a limitation of MTLs for delivery in higher mammals (mice, rats, humans) is that these constructs require recombinant expression. This is less than ideal if we want to extend membrane anchored ligand technology to other *in vivo* systems or adapt this technology as a future therapeutic. In light of these considerations, we pursued an alternative strategy to achieve anchoring without using DNA constructs. As described in Chapter 4 we have generated and characterized soluble membrane anchored ligands (SMALs), peptides conjugated to a lipid designed to mimic corresponding MTLs.

Lipidation as a strategy to generate an anchored ligand had previously been used to a limited extent in the Kopin lab. A previous study showed that like tethered Exendin, a GM-1 ganglioside (lipid moiety) conjugated to exendin-4 is able to anchor the ligand to the cell membrane and facilitate activation of GLP1R (Fortin, Chinnapen et al. 2011). In addition to SMALs which activate GPCRs from the outside of the cell and mimic a tethered ligand, lipidated portions of the receptor have been shown to act as agonists or antagonists. These molecules are usually derived from sequence in the 2nd or 3rd intracellular loop of the corresponding GPCR and are known as pepducins. Notably, the mode of action of pepducins is distinct from SMALs as they are postulated to work through an intracellular mechanism. Pepducins do not incorporate traditional peptide ligands, but instead use amino acid sequences corresponding to the GPCR itself (O'Callaghan, Kuliopulos et al. 2012).

While lipidation of exendin was feasible to generate a receptor modulator as outlined above, it remained unknown whether SMAL technology could generally recapitulate features of membrane tethered ligands. If SMALs do mimic MTLs, the combination of technologies provides a powerful system to develop and optimize anchored ligands. As shown in Chapter 4, both CCK4 and substance P SMALs mimic pharmacological features of their corresponding MTLs suggesting that this approach can be generalized. This has been further supported by additional follow up studies in the Kopin laboratory.

Anchoring appears to provide some unique and advantageous pharmacological properties to conventional soluble ligands. These may include increases in potency, duration of activity, and stability. In the case of substance P and CCK4, anchoring the

ligand using either a recombinant or lipid based strategy makes C-terminal amidation less important for activity. The ability to detect activity without the need for post-translational processing has important implications for the identification of novel peptide ligands. In the context of large scale screening of genetic data for novel biologically active peptides, assessment of corresponding MTLs and SMALs greatly increases the likelihood of being able to demonstrate peptide induced function. If a post-translational modification of a ligand were not present when a soluble peptide was screened, a ligand might be falsely considered inactive. However in the context of anchored ligands, unmodified peptide activity is more likely to be detected.

The ability to readily transform a membrane tethered ligand to a SMAL using well established chemical approaches opens an array of possibilities. Screening for optimized peptides can be done using the MTL system, the corresponding SMAL can then be generated for *in vivo* delivery. Illustrating this approach, SMALs have been made in the Kopin lab based on the previously mentioned membrane tethered chemerin. This SMAL also recapitulates MTL activity well and has promising therapeutic applications as an anti-inflammatory compound.

SMAL technology provides an exciting platform for drug discovery. Additional questions remain that will enable the technology to be further refined. We are currently exploring the extent to which different lipids alter the activity of SMALs. In the studies described in Chapter 4, all SMALs that were tested used a palmitic acid as the lipid. However, there are many other lipid groups that can act as anchors and will potentially alter SMAL activity. For example lipids that have either higher or lower affinity for a cell membrane could be used to alter binding affinity. These might in turn influence the

duration of signaling, the degree of receptor internalization and/or the efficacy of a given SMAL peptide.

A second important feature of SMALs that deserves in depth study is how they work *in vivo*. The Kopin lab has generated a series of SMALs that have therapeutic potential. The previously mentioned chemerin SMAL has been used both in an asthma model and a neuropathic pain model, both in mouse. In each case, the SMAL shows promise as an anti-inflammatory and/or analgesic (data not yet published).

It is possible and remains to be determined whether MTLs and SMALs result in biased signaling. For the studies of MTLs and SMALs which were highlighted in this thesis, we assessed canonical G protein-dependent signal transduction. It will be of interest in future studies to determine the extent to which MTLs and SMALs activate alternative pathways including beta-arrestin and/or ERK/MAP kinase. The outcome of these studies will provide a preview of whether anchored ligands preferentially activate specific pathways (i.e. are biased ligands); an area of research that is of particular interest with respect to drug development (Wootten, Christopoulos et al. 2013).

Summary and final concluding remarks

In Chapter 1, we generated and used membrane tethered bursicon to better understand rk mediated signaling *in vitro*. In so doing, we have extended the capabilities of MTLs. We were able to successfully convert bursicon into the first heterodimeric membrane tethered ligand. Bursicon was also the first MTL made using a type II transmembrane domain with a free extracellular C-terminus. We were also able to perform structure function studies using bursicon MTLs, converting an α-β bursicon fusion construct from an agonist to an antagonist. The template system developed for bursicon has been useful to generate other complex membrane tethered ligands. In addition, the bursicon constructs set the stage to generate MTLs in which the free C-terminus of a peptide is critical for activity.

Bursicon MTLs allowed us to better understand the *in vivo* consequences of bursicon/rk signaling (Chapter 2). Using the single subunit fusion construct (CFP-tBur-β-α) we identified an unexpected result of chronic rk activation; the phenotype resembled receptor inactivation. When over stimulated, rk appears to be downregulated most likely through receptor desensitization. Using this inhibitory, we were able to better understand the tissue specific requirements of rk. We identified adult muscle as an important bursicon target tissue, one which requires rk activation for flies to survive through eclosion and subsequent wing expansion. These studies have broad implications for the use of MTLs *in vivo*. A tethered agonist may possess either agonist or apparent antagonist activity when expressed *in vivo*. The level of activity appears to be a function

of the degree to which the MTL results in receptor desensitization and/or internalization. Our study highlights the importance of evaluating MTL induced desensitization *in vitro* as part of a standard set of assays. Our *in vivo* studies reinforced the critical role of bursicon and rk in insect development (Arakane, Li et al. 2008; Loveall and Deitcher 2010; Lahr, Dean et al. 2012) and encouraged us to pursue rk as an insecticide target.

Leveraging the assays used during previous studies, we developed a high throughput screen to identify chemical antagonists of rk signal transduction (Chapter 3). Two separate libraries were screened, a natural product derived library and a synthetic small molecule library. Through these studies we found that synthetic small molecule screens had a distinct advantage. In contrast to natural products, with small molecules once hits were found, the identity of the compound was immediately known. A promising set of compounds derived from the small molecule screen. Plans are in place to characterize this series of molecules on a variety of insect disease vectors. In addition, work is in progress to generate additional derivatives of promising leads. Success in our screen sets an important precedent that will hopefully lead to the screening of other insect GPCRs to identify novel insecticides.

Finally, as a logical extension of membrane tethered ligands, we sought to develop another form of anchored ligands that maintain the pharmacological advantages of anchoring but bypass the need for recombinant DNA expression (Chapter 4). For this purpose, our lab uses lipidated peptides which also anchor to the cell membrane. This SMAL technology is complementary to MTLs; together the approaches provide a system to study and deliver anchored peptide ligands. We have shown that SMALs mimic the pharmacological properties of membrane tethered ligands. This means that SMALs can

be used as powerful probes to modulate GPCRs both *in vitro* and *in vivo*. The ability to apply SMALs directly to a target tissue using traditional methods of drug application (e.g. inhalation: airways, enema: colon, cream: skin) extends their utility as potential therapeutics.

In conclusion we have developed a novel complex membrane tethered bursicon construct. It has been successfully used both *in vivo* and *in vitro*. Information garnered from this MTL led to the development of high throughput screens for chemical antagonists of rk as well as the extension to SMAL technology. We feel that we have made significant contributions to the field bursicon/rk biology. Our studies have broader implications as a model for understanding GPCRs and corresponding peptides. Strategies developed throughout this thesis can be generalized and have already proven applicable to other receptor ligand pairs. The positive results obtained in each Chapter of this thesis and the broad applicability of the approaches underlines the potential impact of this work.

References

Al-Fulaij, M. A., Y. Ren, et al. (2007). "Identification of amino acid determinants of dopamine 2 receptor synthetic agonist function." J Pharmacol Exp Ther **321**(1): 298-307.

An, S., S. Dong, et al. (2012). "Insect Neuropeptide Bursicon Homodimers Induce Innate Immune and Stress Genes during Molting by Activating the NF-kappaB Transcription Factor Relish." PLoS One **7**(3): e34510.

An, S., S. Wang, et al. (2009). "Identification of a novel bursicon-regulated transcriptional regulator, md13379, in the house fly Musca domestica." Arch Insect Biochem Physiol **70**(2): 106-21.

Arakane, Y., B. Li, et al. (2008). "Functional analysis of four neuropeptides, EH, ETH, CCAP and bursicon, and their receptors in adult ecdysis behavior of the red flour beetle, Tribolium castaneum." Mech Dev **125**(11-12): 984-95.

Auer, S. and I. Ibanez-Tallon (2010). ""The King is dead": Checkmating ion channels with tethered toxins." Toxicon **56**(8): 1293-8.

Avsian-Kretchmer, O. and A. J. Hsueh (2004). "Comparative genomic analysis of the eight-membered ring cystine knot-containing bone morphogenetic protein antagonists." Mol Endocrinol **18**(1): 1-12.

Bai, H. and S. R. Palli (2010). "Functional characterization of bursicon receptor and genome-wide analysis for identification of genes affected by bursicon receptor RNAi." Dev Biol **344**(1): 248-58.

Baker, J. D. and J. W. Truman (2002). "Mutations in the *Drosophila* glycoprotein hormone receptor, rickets, eliminate neuropeptide-induced tanning and selectively block a stereotyped behavioral program." J Exp Biol **205**(Pt 17): 2555-65.

Baker, P. W., K. K. Tanaka, et al. (2005). "Adult myogenesis in *Drosophila* melanogaster can proceed independently of myocyte enhancer factor-2." Genetics **170**(4): 1747-59.

Bate, M., E. Rushton, et al. (1991). "Cells with persistent twist expression are the embryonic precursors of adult muscles in *Drosophila*." Development **113**(1): 79-89.

Beinborn, M., C. Chen, et al. (1998). "Small synthetic ligands of the cholecystokinin-B/gastrin receptor can mimic the function of endogenous peptide hormones." Yale J Biol Med **71**(3-4): 337-46.

Bellmann-Sickert, K., C. E. Elling, et al. (2011). "Long-acting lipidated analogue of human pancreatic polypeptide is slowly released into circulation." J Med Chem **54**(8): 2658-67.

Bellucci, F., F. Carini, et al. (2002). "Pharmacological profile of the novel mammalian tachykinin, hemokinin 1." Br J Pharmacol **135**(1): 266-74.

Bhattacharya, A. A., T. Grune, et al. (2000). "Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin." J Mol Biol **303**(5): 721-32.

Bjarnadottir, T. K., D. E. Gloriam, et al. (2006). "Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse." Genomics **88**(3): 263-73.

Blacquiere, T., G. Smagghe, et al. (2012). "Neonicotinoids in bees: a review on concentrations, side-effects and risk assessment." Ecotoxicology **21**(4): 973-92.

Butnev, V. Y., V. Singh, et al. (2002). "Truncated equine LH beta and asparagine(56)-deglycosylated equine LH alpha combine to produce a potent FSH antagonist." J Endocrinol **172**(3): 545-55.

Carmon, K. S., X. Gong, et al. (2011). "R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling." Proc Natl Acad Sci U S A **108**(28): 11452-7.

Catteruccia, F. and E. A. Levashina (2009). "RNAi in the malaria vector, Anopheles gambiae." Methods Mol Biol **555**: 63-75.

Chan, W. Y. (1998). "Molecular genetic, biochemical, and clinical implications of gonadotropin receptor mutations." Mol Genet Metab **63**(2): 75-84.

Chintapalli, V. R., J. Wang, et al. (2007). "Using FlyAtlas to identify better *Drosophila* melanogaster models of human disease." Nat Genet **39**(6): 715-20.

Choi, C., J. P. Fortin, et al. (2009). "Cellular dissection of circadian peptide signals with genetically encoded membrane-tethered ligands." Curr Biol **19**(14): 1167-75.

Christie, A. E., V. Roncalli, et al. (2013). "Peptidergic signaling in Calanus finmarchicus (Crustacea, Copepoda): in silico identification of putative peptide hormones and their receptors using a de novo assembled transcriptome." Gen Comp Endocrinol **187**: 117-35.

Commonwealth of Massachusetts, D. o. E. P. (2014). "Aerial Application of Pesticide to Control Eastern Equine Encephalitis (EEE) Carrying Mosquitoes."

Cripps, R. M., B. L. Black, et al. (1998). "The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during *Drosophila* myogenesis." Genes Dev **12**(3): 422-34.

Cui, H., M. J. Webber, et al. (2010). "Self-assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials." Biopolymers **94**(1): 1-18.

Currie, D. A. and M. Bate (1991). "The development of adult abdominal muscles in *Drosophila*: myoblasts express twist and are associated with nerves." Development **113**(1): 91-102.

Cuttitta, F. (1993). "Peptide amidation: signature of bioactivity." Anat Rec **236**(1): 87-93, 172-3; discussion 93-5.

Dai, L., E. M. Dewey, et al. (2008). "Identification, developmental expression, and functions of bursicon in the tobacco hawkmoth, Manduca sexta." J Comp Neurol **506**(5): 759-74.

Darling, R. J., R. W. Ruddon, et al. (2000). "Cystine knot mutations affect the folding of the glycoprotein hormone alpha-subunit. Differential secretion and assembly of partially folded intermediates." J Biol Chem **275**(20): 15413-21.

Dasgupta, P. and R. Mukherjee (2000). "Lipophilization of somatostatin analog RC-160 with long chain fatty acid improves its antiproliferative and antiangiogenic activity in vitro." Br J Pharmacol **129**(1): 101-9.

Delporte, C. (2013). "Structure and Physiological Actions of Ghrelin." Scientifica (Cairo) **2013**: 518909.

Diao, F. and B. H. White (2012). "A novel approach for directing transgene expression in *Drosophila*: T2A-Gal4 in-frame fusion." Genetics **190**(3): 1139-44.

Dietzl, G., D. Chen, et al. (2007). "A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*." Nature **448**(7150): 151-6.

Doyle, J. R., J. P. Fortin, et al. (2012). "Selected melanocortin 1 receptor single-nucleotide polymorphisms differentially alter multiple signaling pathways." J Pharmacol Exp Ther **342**(2): 318-26.

Draper, I., P. T. Kurshan, et al. (2007). "Locomotor activity is regulated by D2-like receptors in *Drosophila*: an anatomic and functional analysis." Dev Neurobiol **67**(3): 378-93.

Dulcis, D., R. B. Levine, et al. (2005). "Role of the neuropeptide CCAP in *Drosophila* cardiac function." J Neurobiol **64**(3): 259-74.

Eipper, B. A. and R. E. Mains (1988). "Peptide alpha-amidation." Annu Rev Physiol **50**: 333-44.

Eipper, B. A., S. L. Milgram, et al. (1993). "Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains." Protein Sci **2**(4): 489-97.

Eipper, B. A., D. A. Stoffers, et al. (1992). "The biosynthesis of neuropeptides: peptide alpha-amidation." Annu Rev Neurosci **15**: 57-85.

Eisenberg, D. M., E. S. Harris, et al. (2011). "Developing a library of authenticated Traditional Chinese Medicinal (TCM) plants for systematic biological evaluation--rationale, methods and preliminary results from a Sino-American collaboration." Fitoterapia **82**(1): 17-33.

Enayati, A. and J. Hemingway (2010). "Malaria management: past, present, and future." Annu Rev Entomol **55**: 569-91.

Eriksen, K. K., F. Hauser, et al. (2000). "Molecular cloning, genomic organization, developmental regulation, and a knock-out mutant of a novel leu-rich repeats-containing G protein-coupled receptor (DLGR-2) from *Drosophila* melanogaster." Genome Res **10**(7): 924-38.

Fares, F. A., S. Yamabe, et al. (1998). "Conversion of thyrotropin heterodimer to a biologically active single-chain." Endocrinology **139**(5): 2459-64.

Farooqui, T. (2007). "Octopamine-mediated neuromodulation of insect senses." Neurochem Res **32**(9): 1511-29.

Fernandes, J., M. Bate, et al. (1991). "Development of the indirect flight muscles of *Drosophila*." Development **113**(1): 67-77.

Fortin, J. P., D. Chinnapen, et al. (2011). "Discovery of dual-action membrane-anchored modulators of incretin receptors." PLoS One **6**(9): e24693.

Fortin, J. P., Y. Zhu, et al. (2009). "Membrane-tethered ligands are effective probes for exploring class B1 G protein-coupled receptor function." Proc Natl Acad Sci U S A **106**(19): 8049-54.

Fox, K. M., J. A. Dias, et al. (2001). "Three-dimensional structure of human follicle-stimulating hormone." Mol Endocrinol **15**(3): 378-89.

Fredriksson, R. and H. B. Schioth (2005). "The repertoire of G-protein-coupled receptors in fully sequenced genomes." Mol Pharmacol **67**(5): 1414-25.

Fujita, T., I. Kawahara, et al. (1998). "Permeability characteristics of tetragastrins across intestinal membranes using the Caco-2 monolayer system: comparison between acylation and application of protease inhibitors." Pharm Res **15**(9): 1387-92.

Fujiwara, S. and T. Amisaki (2008). "Identification of high affinity fatty acid binding sites on human serum albumin by MM-PBSA method." Biophys J **94**(1): 95-103.

Galet, C., F. Guillou, et al. (2009). "The beta104-109 sequence is essential for the secretion of correctly folded single-chain beta alpha horse LH/CG and for its FSH activity." J Endocrinol **203**(1): 167-74.

Garcia, M. I., M. Ghiani, et al. (2009). "LGR5 deficiency deregulates Wnt signaling and leads to precocious Paneth cell differentiation in the fetal intestine." Dev Biol **331**(1): 58-67.

Ghazi, A., S. Anant, et al. (2000). "Apterous mediates development of direct flight muscles autonomously and indirect flight muscles through epidermal cues." Development **127**(24): 5309-18.

Gill, S. C. and P. H. von Hippel (1989). "Calculation of protein extinction coefficients from amino acid sequence data." Anal Biochem **182**(2): 319-26.

Harwood, B. N., J. P. Fortin, et al. (2013). "Membrane tethered bursicon constructs as heterodimeric modulators of the *Drosophila* G protein-coupled receptor rickets." Mol Pharmacol **83**(4): 814-21.

Hearn, M. G., Y. Ren, et al. (2002). "A *Drosophila* dopamine 2-like receptor: Molecular characterization and identification of multiple alternatively spliced variants." Proc Natl Acad Sci U S A **99**(22): 14554-9.

Hearn, M. T. and P. T. Gomme (2000). "Molecular architecture and biorecognition processes of the cystine knot protein superfamily: part I. The glycoprotein hormones." J Mol Recognit **13**(5): 223-78.

Hemingway, J., N. J. Hawkes, et al. (2004). "The molecular basis of insecticide resistance in mosquitoes." Insect Biochem Mol Biol **34**(7): 653-65.

Hitmi, A., A. Coudret, et al. (2000). "The production of pyrethrins by plant cell and tissue cultures of Chrysanthemum cinerariaefolium and Tagetes species." Crit Rev Biochem Mol Biol **35**(5): 317-37.

Hoare, S. R. (2005). "Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors." Drug Discov Today **10**(6): 417-27.

Honegger, H. W., E. M. Dewey, et al. (2008). "Bursicon, the tanning hormone of insects: recent advances following the discovery of its molecular identity." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **194**(12): 989-1005.

Honegger, H. W., T. Y. Estevez-Lao, et al. (2011). "Bursicon-expressing neurons undergo apoptosis after adult ecdysis in the mosquito Anopheles gambiae." J Insect Physiol **57**(7): 1017-22.

Hughes, A. L. (2012). "Amino acid sequence coevolution in the insect bursicon ligand-receptor system." Mol Phylogenet Evol.

Hughes, J., P. Boden, et al. (1990). "Development of a class of selective cholecystokinin type B receptor antagonists having potent anxiolytic activity." Proc Natl Acad Sci U S A **87**(17): 6728-32.

Ibanez-Tallon, I. and M. N. Nitabach (2012). "Tethering toxins and peptide ligands for modulation of neuronal function." Curr Opin Neurobiol **22**(1): 72-8.

Ibanez-Tallon, I., H. Wen, et al. (2004). "Tethering naturally occurring peptide toxins for cell-autonomous modulation of ion channels and receptors in vivo." Neuron **43**(3): 305-11.

Isman, M. B. (2006). "Botanical insecticides, deterrents, and repellents in modern agriculture and an increasingly regulated world." Annu Rev Entomol **51**: 45-66.

Ito, M., T. Matsui, et al. (1993). "Functional characterization of a human brain cholecystokinin-B receptor. A trophic effect of cholecystokinin and gastrin." J Biol Chem **268**(24): 18300-5.

Jacoby, E., R. Bouhelal, et al. (2006). "The 7 TM G-protein-coupled receptor target family." ChemMedChem **1**(8): 761-82.

Kang, D. S., X. Tian, et al. (2013). "beta-Arrestins and G protein-coupled receptor trafficking." Methods Enzymol **521**: 91-108.

Kaspar, A. A. and J. M. Reichert (2013). "Future directions for peptide therapeutics development." Drug Discov Today **18**(17-18): 807-17.

Kelly-Hope, L., H. Ranson, et al. (2008). "Lessons from the past: managing insecticide resistance in malaria control and eradication programmes." Lancet Infect Dis **8**(6): 387-9.

Khan, S. M., R. Sleno, et al. (2013). "The expanding roles of Gbetagamma subunits in G protein-coupled receptor signaling and drug action." Pharmacol Rev **65**(2): 545-77.

Kim, K. H. and B. L. Seong (2001). "Peptide Amidation: Production of Peptide Hormones in vivo and in vitro." Biotechnology and Bioprocess Engineering **6**: 244-251.

Lahr, E. C., D. Dean, et al. (2012). "Genetic analysis of ecdysis behavior in *Drosophila* reveals partially overlapping functions of two unrelated neuropeptides." J Neurosci **32**(20): 6819-29.

Lee, Y. M., M. Beinborn, et al. (1993). "The human brain cholecystokinin-B/gastrin receptor. Cloning and characterization." J Biol Chem **268**(11): 8164-9.

Leeman, S. E. and S. L. Ferguson (2000). "Substance P: an historical perspective." Neuropeptides **34**(5): 249-54.

Lin, H., M. F. Sassano, et al. "A pharmacological organization of G protein-coupled receptors." Nat Methods **10**(2): 140-6.

Lindstrom, E., B. von Mentzer, et al. (2007). "Neurokinin 1 receptor antagonists: correlation between in vitro receptor interaction and in vivo efficacy." J Pharmacol Exp Ther **322**(3): 1286-93.

Loveall, B. J. and D. L. Deitcher (2010). "The essential role of bursicon during *Drosophila* development." BMC Dev Biol **10**: 92.

Luan, H., W. C. Lemon, et al. (2006). "Functional dissection of a neuronal network required for cuticle tanning and wing expansion in *Drosophila*." J Neurosci **26**(2): 573-84.

Lundstrom, K. (2009). "An overview on GPCRs and drug discovery: structure-based drug design and structural biology on GPCRs." Methods Mol Biol **552**: 51-66.

Luo, C. W., E. M. Dewey, et al. (2005). "Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2." Proc Natl Acad Sci U S A **102**(8): 2820-5.

Marmenout, A., L. Fransen, et al. (1985). "Molecular cloning and expression of human tumor necrosis factor and comparison with mouse tumor necrosis factor." Eur J Biochem **152**(3): 515-22.

Matsueda, G. R. (1982). "Deprotection of Nin-formyl tryptophan using 1,2-ethanedithiol in liquid hydrogen fluoride. Deformylation upon HF treatment of Merrifield peptidyl-resins." Int J Pept Protein Res **20**(1): 26-34.

Mayer, M. R. and M. J. Stone (2001). "Identification of receptor binding and activation determinants in the N-terminal and N-loop regions of the CC chemokine eotaxin." J Biol Chem **276**(17): 13911-6.

McLean, S., A. Ganong, et al. (1993). "Pharmacology of CP-99,994; a nonpeptide antagonist of the tachykinin neurokinin-1 receptor." J Pharmacol Exp Ther **267**(1): 472-9.

Mendive, F. M., T. Van Loy, et al. (2005). "*Drosophila* molting neurohormone bursicon is a heterodimer and the natural agonist of the orphan receptor DLGR2." FEBS Lett **579**(10): 2171-6.

Myers, S. R., F. E. Yakubu-Madus, et al. (1997). "Acylation of human insulin with palmitic acid extends the time action of human insulin in diabetic dogs." Diabetes **46**(4): 637-42.

Nathans, J. and D. S. Hogness (1983). "Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin." Cell **34**(3): 807-14.

New, D. C. and Y. H. Wong (2007). "Molecular mechanisms mediating the G protein-coupled receptor regulation of cell cycle progression." J Mol Signal **2**: 2.

Nishi, S., S. Y. Hsu, et al. (2000). "Characterization of two fly LGR (leucine-rich repeat-containing, G protein-coupled receptor) proteins homologous to vertebrate glycoprotein hormone receptors: constitutive activation of wild-type fly LGR1 but not LGR2 in transfected mammalian cells." Endocrinology **141**(11): 4081-90.

Nishida, A., K. Miyata, et al. (1994). "Pharmacological profile of (R)-1-[2,3-dihydro-1-(2'-methylphenacyl)-2-oxo- 5-phenyl-1H-1,4-benzodiazepin-3-yl]-3-(3-methylphenyl)urea (YM022), a new potent and selective gastrin/cholecystokinin-B receptor antagonist, in vitro and in vivo." J Pharmacol Exp Ther **269**(2): 725-31.

Nongthomba, U., S. Pasalodos-Sanchez, et al. (2001). "Expression and function of the *Drosophila* ACT88F actin isoform is not restricted to the indirect flight muscles." J Muscle Res Cell Motil **22**(2): 111-9.

O'Callaghan, K., A. Kuliopulos, et al. (2012). "Turning receptors on and off with intracellular pepducins: new insights into G-protein-coupled receptor drug development." J Biol Chem **287**(16): 12787-96.

Olson, E. N., M. Perry, et al. (1995). "Regulation of muscle differentiation by the MEF2 family of MADS box transcription factors." Dev Biol **172**(1): 2-14.

Overington, J. P., B. Al-Lazikani, et al. (2006). "How many drug targets are there?" Nat Rev Drug Discov **5**(12): 993-6.

Ozawa, A., I. Lindberg, et al. "Deorphanization of novel peptides and their receptors." Aaps J **12**(3): 378-84.

Paluzzi, J. P., M. Vanderveken, et al. (2014). "The Heterodimeric Glycoprotein Hormone, GPA2/GPB5, Regulates Ion Transport across the Hindgut of the Adult Mosquito, Aedes aegypti." PLoS One **9**(1): e86386.

Park, J. H., A. J. Schroeder, et al. (2003). "Targeted ablation of CCAP neuropeptide-containing neurons of *Drosophila* causes specific defects in execution and circadian timing of ecdysis behavior." Development **130**(12): 2645-56.

Park, J. I., J. Semyonov, et al. (2005). "Conservation of the heterodimeric glycoprotein hormone subunit family proteins and the LGR signaling system from nematodes to humans." Endocrine **26**(3): 267-76.

Peabody, N. C., F. Diao, et al. (2008). "Bursicon functions within the *Drosophila* CNS to modulate wing expansion behavior, hormone secretion, and cell death." J Neurosci **28**(53): 14379-91.

Peabody, N. C., J. B. Pohl, et al. (2009). "Characterization of the decision network for wing expansion in *Drosophila* using targeted expression of the TRPM8 channel." J Neurosci **29**(11): 3343-53.

Pennington, M. W. (1994). "HF cleavage and deprotection procedures for peptides synthesized using a Boc/Bzl strategy." Methods Mol Biol **35**: 41-62.

Pierce, K. L., R. T. Premont, et al. (2002). "Seven-transmembrane receptors." Nat Rev Mol Cell Biol **3**(9): 639-50.

Pinto, J., A. Lynd, et al. (2007). "Multiple origins of knockdown resistance mutations in the Afrotropical mosquito vector Anopheles gambiae." PLoS One **2**(11): e1243.

Pisegna, J. R., A. de Weerth, et al. (1994). "Molecular cloning, functional expression, and chromosomal localization of the human cholecystokinin type A receptor." Ann N Y Acad Sci **713**: 338-42.

Raghavan, K. V., N. Gendre, et al. (1996). "Transplanted wing and leg imaginal discs in *Drosophila* melanogaster demonstrate interactions between epidermis and myoblasts in muscle formation." Dev Genes Evol **206**(1): 46-53.

Raghavendra, K., T. K. Barik, et al. (2011). "Malaria vector control: from past to future." Parasitol Res **108**(4): 757-79.

Ramphul, U., T. Boase, et al. (2009). "Insecticide resistance and its association with target-site mutations in natural populations of Anopheles gambiae from eastern Uganda." Trans R Soc Trop Med Hyg **103**(11): 1121-6.

Ranganayakulu, G., B. Zhao, et al. (1995). "A series of mutations in the D-MEF2 transcription factor reveal multiple functions in larval and adult myogenesis in *Drosophila*." Dev Biol **171**(1): 169-81.

Robishaw, J. D. and C. H. Berlot (2004). "Translating G protein subunit diversity into functional specificity." Curr Opin Cell Biol **16**(2): 206-9.

Roy, A., P. R. McDonald, et al. (2010). "Open access high throughput drug discovery in the public domain: a Mount Everest in the making." Curr Pharm Biotechnol **11**(7): 764-78.

Sato, A., E. Perlas, et al. (1997). "Cystine knot of the gonadotropin alpha subunit is critical for intracellular behavior but not for in vitro biological activity." J Biol Chem **272**(29): 18098-103.

Sato, T., R. G. Vries, et al. (2009). "Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche." Nature **459**(7244): 262-5.

Schlyer, S. and R. Horuk (2006). "I want a new drug: G-protein-coupled receptors in drug development." Drug Discov Today **11**(11-12): 481-93.

Schnolzer, M., P. Alewood, et al. (1992). "In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences." Int J Pept Protein Res **40**(3-4): 180-93.

Seifert, R. and K. Wenzel-Seifert (2002). "Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors." Naunyn Schmiedebergs Arch Pharmacol **366**(5): 381-416.

Sen Gupta, C. and R. R. Dighe (2000). "Biological activity of single chain chorionic gonadotropin, hCGalphabeta, is decreased upon deletion of five carboxyl terminal amino acids of the alpha subunit without affecting its receptor binding." J Mol Endocrinol **24**(2): 157-64.

Setlur, S. R. and R. R. Dighe (2007). "Single chain human chorionic gonadotropin, hCGalphabeta: effects of mutations in the alpha subunit on structure and bioactivity." Glycoconj J **24**(1): 97-106.

Setoh, K., M. Murakami, et al. (1995). "Improvement of transdermal delivery of tetragastrin by lipophilic modification with fatty acids." J Pharm Pharmacol **47**(10): 808-11.

Shimamura, K., M. Matsuda, et al. (2009). "Identification of a stable chemerin analog with potent activity toward ChemR23." Peptides **30**(8): 1529-38.

Snyder, J. C., L. K. Rochelle, et al. (2013). "The stem cell-expressed receptor Lgr5 possesses canonical and functionally active molecular determinants critical to beta-arrestin-2 recruitment." PLoS One **8**(12): e84476.

Snyder, J. C., L. K. Rochelle, et al. (2013). "Constitutive internalization of the leucine-rich G protein-coupled receptor-5 (LGR5) to the trans-Golgi network." J Biol Chem **288**(15): 10286-97.

Sudo, S., Y. Kuwabara, et al. (2005). "Heterodimeric fly glycoprotein hormone-alpha2 (GPA2) and glycoprotein hormone-beta5 (GPB5) activate fly leucine-rich repeat-containing G protein-coupled receptor-1 (DLGR1) and stimulation of human thyrotropin receptors by chimeric fly GPA2 and human GPB5." Endocrinology **146**(8): 3596-604.

Sugahara, T., P. D. Grootenhuis, et al. (1996). "Expression of biologically active fusion genes encoding the common alpha subunit and either the CG beta or FSH beta subunits: role of a linker sequence." Mol Cell Endocrinol **125**(1-2): 71-7.

Tanaka, K., T. Fujita, et al. (1996). "Enhancement of intestinal transport of thyrotropin-releasing hormone via a carrier-mediated transport system by chemical modification with lauric acid." Biochim Biophys Acta **1283**(1): 119-26.

Tenma, T., E. Yodoya, et al. (1993). "Development of new lipophilic derivatives of tetragastrin: physicochemical characteristics and intestinal absorption of acyl-tetragastrin derivatives in rats." Pharm Res **10**(10): 1488-92.

Toullec, J. Y., E. Corre, et al. (2013). "Transcriptome and peptidome characterisation of the main neuropeptides and peptidic hormones of a euphausiid: the Ice Krill, Euphausia crystallorophias." PLoS One **8**(8): e71609.

Turusov, V., V. Rakitsky, et al. (2002). "Dichlorodiphenyltrichloroethane (DDT): ubiquity, persistence, and risks." Environ Health Perspect **110**(2): 125-8.

Van Hiel, M. B., H. P. Vandersmissen, et al. (2011). "An evolutionary comparison of leucine-rich repeat containing G protein-coupled receptors reveals a novel LGR subtype." Peptides **34**(1): 193-200.

Van Loy, T., M. B. Van Hiel, et al. (2007). "Evolutionary conservation of bursicon in the animal kingdom." Gen Comp Endocrinol **153**(1-3): 59-63.

Van Loy, T., H. P. Vandersmissen, et al. (2008). "Comparative genomics of leucine-rich repeats containing G protein-coupled receptors and their ligands." Gen Comp Endocrinol **155**(1): 14-21.

Verlinden, H., R. Vleugels, et al. (2010). "The role of octopamine in locusts and other arthropods." J Insect Physiol **56**(8): 854-67.

Wang, J., D. J. Hogenkamp, et al. (2006). "Reversible lipidization for the oral delivery of leu-enkephalin." J Drug Target **14**(3): 127-36.

Wheatley, M. and S. R. Hawtin (1999). "Glycosylation of G-protein-coupled receptors for hormones central to normal reproductive functioning: its occurrence and role." Hum Reprod Update **5**(4): 356-64.

Wise, A., K. Gearing, et al. (2002). "Target validation of G-protein coupled receptors." Drug Discov Today **7**(4): 235-46.

Woodruff, E. A., 3rd, K. Broadie, et al. (2008). "Two peptide transmitters co-packaged in a single neurosecretory vesicle." Peptides **29**(12): 2276-80.

Wootten, D., A. Christopoulos, et al. (2013). "Emerging paradigms in GPCR allostery: implications for drug discovery." Nat Rev Drug Discov **12**(8): 630-44.

World Health, O. (2013). World malaria report 2013, World Health Organization.

Xing, Y., R. V. Myers, et al. (2004). "Glycoprotein hormone assembly in the endoplasmic reticulum: IV. Probable mechanism of subunit docking and completion of assembly." J Biol Chem **279**(34): 35458-68.

Yamamoto, A. (1998). "[Improvement of intestinal absorption of peptide and protein drugs by chemical modification with fatty acids]." Nihon Rinsho **56**(3): 601-7.

Yamamoto, A. (2001). "[Improvement of transmucosal absorption of biologically active peptide drugs]." Yakugaku Zasshi **121**(12): 929-48.

Yin, Z., X. L. Xu, et al. (1997). "Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development." Development **124**(24): 4971-82.

Yodoya, E., K. Uemura, et al. (1994). "Enhanced permeability of tetragastrin across the rat intestinal membrane and its reduced degradation by acylation with various fatty acids." J Pharmacol Exp Ther **271**(3): 1509-13.

Yuan, L., J. Wang, et al. (2005). "Reversible lipidization prolongs the pharmacological effect, plasma duration, and liver retention of octreotide." Pharm Res **22**(2): 220-7.

Zaffran, S., M. Astier, et al. (1997). "The held out wings (how) *Drosophila* gene encodes a putative RNA-binding protein involved in the control of muscular and cardiac activity." Development **124**(10): 2087-98.

Zarandi, M., J. L. Varga, et al. (2006). "Lipopeptide antagonists of growth hormone-releasing hormone with improved antitumor activities." Proc Natl Acad Sci U S A **103**(12): 4610-5.

Zaric, V., D. Weltin, et al. (2004). "Effective polyethylenimine-mediated gene transfer into human endothelial cells." J Gene Med **6**(2): 176-84.

Zhang, L. and G. Bulaj (2012). "Converting peptides into drug leads by lipidation." Curr Med Chem **19**(11): 1602-18.

Zhang, M., Y. X. Tao, et al. (2007). "Intrinsic differences in the response of the human lutropin receptor versus the human follitropin receptor to activating mutations." J Biol Chem **282**(35): 25527-39.