2016

Homology Modeling and Molecular Docking of Antagonists to Class B G-Protein Coupled Receptor Pituitary Adenylate Cyclase Type 1 (PAC1R)

Suzanne Louise Stanton
University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/graddis

Part of the Chemistry Commons

HOMOLOGY MODELING AND MOLECULAR DOCKING OF ANTAGONISTS TO
CLASS B G-PROTEIN COUPLED RECEPTOR PITUITARY ADENYLATE
CYLCASE TYPE 1 (PAC1R)

A Thesis Presented

by

Suzanne Louise Stanton

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Organic Chemistry/Chemistry

October, 2016

Defense Date: May 9, 2016
Thesis Examination Committee:

Matthias Brewer, Ph.D., Advisor
Jose Madalengoitia, Ph.D.
Victor May, Ph.D.
Rory Waterman, Ph.D.
Cynthia J. Forehand, PhD., Dean of the Graduate College

ABSTRACT

Recent studies have identified the Class B g-protein coupled receptor (GPCR) pituitary adenylate cyclase activating polypeptide type 1 (PAC1R) as a key component in physiological stress management. Over-activity of neurological stress response systems due to prolonged or extreme exposure to traumatic events has led researchers to investigate PAC1R inhibition as a possible treatment for anxiety disorders such as post-traumatic stress disorder (PTSD). In 2008, Beebe and coworkers identified two such small molecule hydrazide antagonists and a general pharmacaphore for PAC1R inhibition. However, a relative dearth of information about Class B GPCRs in general, and PAC1R in specific, has significantly hindered progress toward the development of small molecule antagonists of PAC1R. The recent crystallization of the homologically similar glucagon receptor (GCGR) by Siu and coworkers in 2013, also a Class B receptor, has provided an experimentally resolved template from which to base computationally derived models of PAC1R.

Initially, this research was focused towards synthesizing small molecule antagonists for PAC1R which were to be biologically screened via a qualitative western blot assay followed by a radioisotope binding assay for those hydrazides exhibiting down-stream signaling inhibitory capabilities. However, the resolution of the GCGR crystal structure shifted research objectives towards developing a homology model of PAC1R and evaluating that computationally created model with Beebe's known small molecule antagonists. Created using academic versions of on-line resources including UniProtKB, Swiss-Model and Maestro, a homology model for PAC1R is presented here. The model is validated and evaluated for the presence of conserved Class B GPCR residues and motifs, including expected disulfide bridges, a conserved tyrosine residue, a GWGxP motif, a conserved glutamic acid residue and the extension of the transmembrane helix 1 (TM1) into the extra-cellular domain.

Having determined this virtual PAC1R an acceptable model, ligand docking studies of known antagonists to the receptor were undertaken using AutoDock Vina in conjunction with AutoDock Tools and PyMol. Computational docking results were evaluated via comparison of theoretical binding affinity results to Beebe's experimental data. Based on hydrogen bonding capabilities, several residues possibly key to the ligand-receptor binding complex are identified and include ASN 240, TYR 241 and HIST 365. Although the docking software does not identify non-bonding interactions other than hydrogen-bonding, the roles of additional proposed binding pocket residues are discussed in terms of hydrophobic interactions, π-π interactions and halogen bonding. These residues include TYR 161, PHE 196, VAL 203, PHE 204, ILE 209, LEU 210, VAL 237, TRP 297, PHE 362 and LEU 386. Although theoretical in nature, this reported homology modeling and docking exercise details a proposed binding site that may potentially further the development of drugs designed for the treatment of PTSD.

DEDICATION

To my father, James A. Stanton, Sr., for inspiring me to reach for the stars. To my mother, Diane S. Stanton, for teaching me persistence and perseverance. To my sister, Nancy Rae Stamm, for always being my hero. And to my son, Calvin James Clark-Stanton, for making everything in this life worth doing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1. Physiology of Post-Traumatic Stress Disorder

Anxiety disorders, such as Post-Traumatic Stress Disorder (PTSD), can be debilitating hindrances to a person's ability to function in everyday life. Recent studies have indicated that PTSD may result when a traumatic event, or a prolonged series of traumatic events, causes an over-reactive adrenaline response in the hypothalamus-pituitary-adrenal (HPA) axis.   This over-reactivity in the HPA axis can cause significant biochemical changes in the brain that are uncharacteristic of other psychiatric disorders (Yehuda et al., 2004).

Researchers have demonstrated that the HPA axis controls stress reactions by secreting corticotrophin-releasing hormone (CRH) from the hypothalamus. CRH in turn stimulates the pituitary gland to secrete adrenocorticotropic hormone (ACTH), which then stimulates the production of glucocorticoids, such as cortisol, in the adrenal cortex (Engelmenn et al., 2004). This glucocorticoid signaling cascade ultimately redistributes the body's available energy to the brain and major muscle groups for "fight or flight" responses. The presence of cortisol in the blood eventually signals the hypothalamus to stop secreting CRH and the hypothalamus-pituitary-adrenal axis slows. Dysfunction of the HPA axis, characterized by abnormally high levels of CRH and low levels of cortisol, is one of the distinct neuroendocrine profiles that differentiates PTSD from other mental disorders (Yehuda et al., 2002). Glucocorticoids have many important functions,

including modulation of stress reactions, but in excess they can be physiologically damaging. In humans and animals exposed to severe stress, atrophy of the hippocampus in the medial temporal lobe of the brain is believed to be caused by prolonged exposure to high concentrations of glucocorticoids (Finkelstein et al., 1985). Deficiencies of the hippocampus caused by this atrophy may reduce the memory resources available to help a body formulate appropriate reactions to external stress.

In addition to the hippocampus, the bed nucleus of the stria terminalis (BNST) is thought to suffer damaging effects from the neurotransmitter cascade induced by chronic and/or excessive stress. The BNST also plays an important role in the regulation of the HPA axis during stress. Located in the brain, the BNST is thought to act as a relay site within the HPA axis and regulate axis activity in response to acute stressors through the expression of CRH (Choi et al., 2007). BNST plays a central role in stress reactivity, particularly to the integration of fear and stress related circuits. Chronic exposure to stressors has been thought to not only overly activate the BNST, but also to alter its plasticity with significant increases measured in BNST volume, dendritic length and number of branch points in BNST neurons following exposure to long term stress (Hammack et al., 2010). Additionally, tissue studies conducted following subject exposure to extensive periods of stress have indicated that regions of the BNST with high levels of CRH also contain high levels of the neuropeptide pituitary adenylate cylclase-activating polypeptide (PACAP) and its class B G-protein coupled pituitary adenylate cyclase 1 receptor (PAC1R) (Choi et al., 2007, Hammack et al., 2010).

1.2. G-Protein Coupled Receptors (GPCRs)

Pituitary Adenylate Cyclase Activating Polypetide 1 receptor (PAC1R) is classified as a G-Protein coupled receptor (GPCR). GPCRs are the largest and most diverse group of membrane receptors in eukaryotes. They can be classified into four major families or classes based on functional similarities and sequence homology: A (rhodopsin), B (secretin), C (metabotropic glutamate) and F (frizzled). In order to be classified as a GPCR, the receptor must conform to two requirements: it must have a hepta-helical membrane spanning domain and it must be able to couple to a heterotrimeric guanosine nucleotide-binding protein or G-protein (Fredriksson et al., 2003). These cell surface receptors act as an inbox for messages received in the form of light energy, peptides, lipids, sugars, and proteins. Such messages inform cells about the presence or absence of life-sustaining light or nutrients in their environment, or they convey information sent by other cells. GPCRs play a role in an incredible array of functions in the human body, and increased understanding of these receptors has greatly impacted modern medicine. Researchers estimate that between one-third and one-half of all marketed drugs act by binding to GPCRs (Choi et al., 2007).

GPCRs consist of a single polypeptide that is folded into a globular shape and embedded in a cell's plasma membrane. Seven segments of this molecule span the entire width of the membrane, thus GPCRs are sometimes called seven-transmembrane receptors (7TM). Intervening portions of the protein loop both inside and outside the cell. The extracellular loops form part of the pockets at which signaling molecules bind to the GPCR. Figure 1-1 (adapted from Kuntal et al, 2012) illustrates the general

3
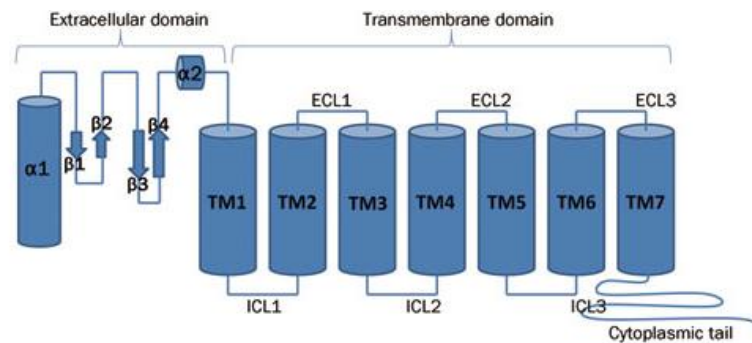
Figure 1-1: General Architecture for GPCRs

architecture of GPCRs. Of the four major classes of GPCRs, the vast majority (approximately 85%) belong to the Class A family of rhodopsin-like receptors.

1.2.1. Class B GPCRs

Of particular interest to this research are family B GPCRs, commonly known as secretin family GPCRs. The secretin family of GPCRs is so-named due to the discovery of the first receptor to belong to this class, isolated by Ishibara et al. in 1991 from rat tissue (Ishihara et al., 1991). There are 15 members of this class, including the secretin receptor (SCTR); glucagon receptor (GCGR); gastric inhibitory peptide receptor (GIPR); glucagon-like peptide-1 and peptide-2 receptors (GLP-1R/ GLP-2R); growth hormone-releasing hormone receptor (GHRHR); calcitonin and calcitonin-like receptor (CALCR/ CALCLR); corticotropin-releasing hormone receptors (CRHR1/ CRHR2); adenylate cyclase activating polypeptide receptor (PAC1/ ADCYAP1R1); parathyroid hormone receptors (PTHR1/ PTHR2) and the vasoactive intestinal peptide receptors (VIPR1/

4

VIPR2). Generally speaking, these receptors lack the signature sequences found in the majority of the much larger group of Class A GPCRs. The most typical features of Class B family GPCRs include a long extracellular amino-terminal tail region with six highly conserved cysteine residues that form 3 intra-domain disulfide bonds. Class B receptors also show a number of conserved proline residues within the TM segments which are thought to be essential for the conformational dynamics of the receptors (Worth et al., 2009).

Family B GPCRs have a similar morphology to several group A GPCRs, but they do not share any sequence homology. Their ligands include high molecular weight hormones such as glucagon, secretine, calcitonin, growth hormone-releasing hormone, corticotropin-releasing factor, VIP-PACAP and the Black widow spider toxin, α-latrotoxin (Jacoby, 2006).

1.3. G-Proteins

Understanding the function of G-proteins is critical to designing potential drugs aimed at their control. As the name implies, GPCRs interact with guanine nucleotide-binding, or G, proteins in the plasma membrane. When an external signaling molecule binds to a GPCR, it causes a conformational change in the GPCR. This change then triggers the interaction between the GPCR and a nearby G-protein. G-proteins are specialized proteins with the ability to bind the nucleotides guanosine triphosphate (GTP) and guanosine diphosphate (GDP). Some G-proteins, such as the signaling protein Ras,

are small proteins with a single subunit. However, the G-proteins that associate with

GPCRs are heterotrimeric, meaning they have three different subunits: an alpha subunit, a

beta subunit, and a gamma subunit. Two of these subunits — alpha and gamma — are

attached to the plasma membrane by lipid anchors. As Figure 1-2 illustrates (adapted

from Alberts et al., 2008), a G-protein alpha subunit binds either GTP or GDP depending

on whether the protein is active (GTP) or inactive (GDP). In the absence of a signal, GDP

attaches to the alpha subunit, and the entire G-protein-GDP complex binds to a nearby

GPCR (Figure 1-2 A). This arrangement persists until a signaling molecule joins with the

GPCR. At this point, a change in the conformation of the GPCR activates the G-protein,

and GTP physically replaces the GDP bound to the alpha subunit (Figure 1-2 B). As a

result, the G-protein subunits dissociate into two parts: the GTP-bound alpha subunit and

a beta-gamma dimer. Both parts remain anchored to the plasma membrane, but they
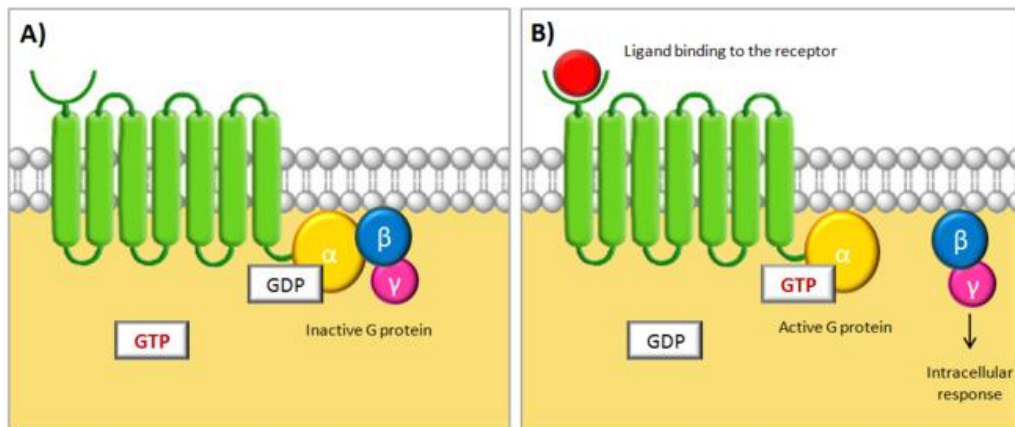
Figure 1-2: G Protein-Coupled Receptor and its Interacting G Protein.

are no longer bound to the GPCR, so they can now diffuse laterally to interact with other

membrane proteins. G-proteins remain active as long as their alpha subunits are joined with GTP. However, when this GTP is hydrolyzed back to GDP, the subunits once again assume the form of an inactive heterotrimer, and the entire G-protein reassociates with the now-inactive GPCR. In this way, G-proteins work like a switch, turned on or off by signal-receptor interactions on the cell surface (Alberts et al., 2008).

Whenever a G-protein is active, both its GTP-bound alpha subunit and its beta-gamma dimer can relay messages in the cell by interacting with other membrane proteins involved in signal transduction. Specific targets for activated G-proteins include various enzymes that produce second messengers, as well as certain ion channels that allow ions to act as second messengers. Some G-proteins stimulate the activity of these targets, whereas others are inhibitory. Vertebrate genomes contain multiple genes that encode the alpha, beta, and gamma subunits of G-proteins. The many different subunits encoded by these genes combine in multiple ways to produce a diverse family of G-proteins (Alberts et al., 2008).

Activation of a single G-protein can affect the production of hundreds or even possibly thousands of second messenger molecules. Agonist binding at the GPCR promotes the active state conformation, which vastly enhances the affinity of the GPCR for the G protein. The affinity of the individual subunits for their cognate secondary effectors is markedly increased, which the subunits then bind to and exert either inhibitory or stimulatory effects upon. The secondary effector then catalyzes the formation of secondary effector molecules, which in turn cause further signaling cascades throughout the cell. Activated G-proteins are capable of stimulating or inhibiting many

effector proteins. Indeed, one active GPCR is capable of activating more than one G-protein upon adopting active conformation, and one activated G-protein is capable of activating more than one secondary effector, and one secondary effector catalyzes the formation of more than one secondary signaling molecule. Therefore, the signal from the GPCR activating ligand is greatly amplified to potentially initiate a prodigious intracellular response from a relatively low concentration of ligand (Alberts et al., 2008).[16] One especially common target of activated G-proteins is adenylyl cyclase, a membrane-associated enzyme that, when activated by the GTP-bound alpha subunit, catalyzes synthesis of the second messenger cAMP from molecules of ATP. In humans, cAMP is involved in responses to sensory input, hormones, and nerve transmission, among other functions.

1.4 Class B GPCR Peptide and Non-Peptide Binding

Activation (and/or deactivation) of the G-protein and its cascade of intracellular signaling begins with the binding of a peptide (or small molecule) to the GPCR. A general mechanism of peptide binding has been proposed for class B GPCRs. In this mechanism, it is suggested that the C-terminal region of the peptide ligand binds to the extracellular N-terminal domain of the receptor (Olejniczak, 2007). This binding acts as an "affinity trap," allowing interaction of the N-terminus of the ligand with the receptor juxtamembrane, or J-domain (Warne et al., 2011). Peptide interaction with the J-domain then activates the receptor, promoting intracellular signaling (Hoare, 2005).

As illustrated in Figure 1-3 (adapted from Pal et al., 2012), the peptide and receptor are orientated for initial receptor-ligand binding. The initial complex forms between the C-terminus of the peptide and the N-terminus of the receptor extracellular domain (ECD). This interaction facilitates the binding of the free N-terminus of the peptide to the juxtamembrane region of the 7TM domain of the receptor. Such binding of the peptide induces a conformational change in the 7TM and cytoplasmic domains of the receptor, which mediates its interaction with a heterotrimeric G-protein. This peptide-protein binding paradigm is often referred to as the "two-domain" model (Hoare, 2005).
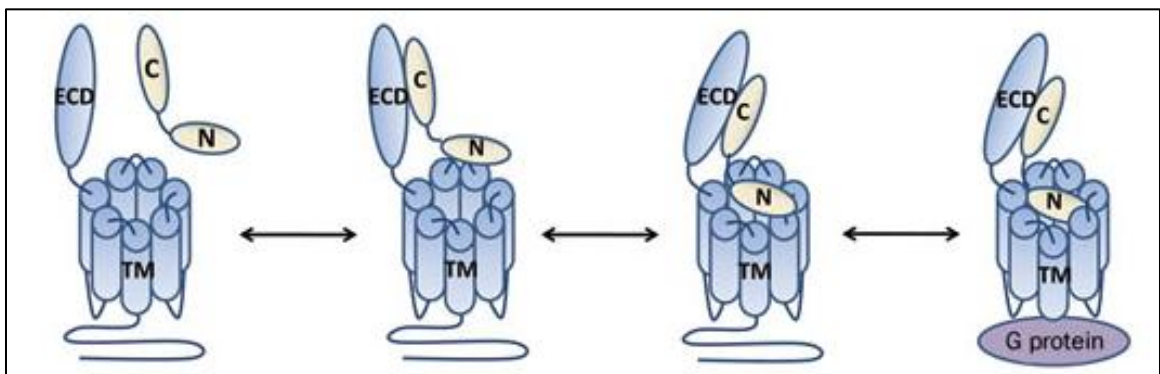
Figure 1-3: Peptide Binding to Class B GPCRs

Non-peptide ligands are thought to bind either the juxtamembrane or N-terminal domain in the activation of intracellular signaling. In Figure 1-4 (adapted from Pal et al., 2012), the non-peptide antagonist ligand (depicted as a solid blue circle), binds with the J-domain in the 7TM portion of the receptor. Binding of the non-peptide antagonist is then thought to induce a change in the J-domain which prevents peptide bonding to that domain. In blocking peptide bonding to the juxtamembrane, the

antagonist prevents peptide-stimulated receptor signaling as peptide-to-J-domain interaction is required for G-protein activation. It has been proposed that nonpeptide antagonists of GPCRs act allosterically, binding sites that are distinct from the peptide-ligand binding regions (Pal et al., 2012). Although non-peptide binding to the J-domain prohibits peptide binding to that region of the receptor, small molecule antagonists are not thought to prevent peptides from binding to the N-domain, as reflected in Figure 4 (adapted from Hoare, 2005). For disorders such as post-traumatic stress, which are
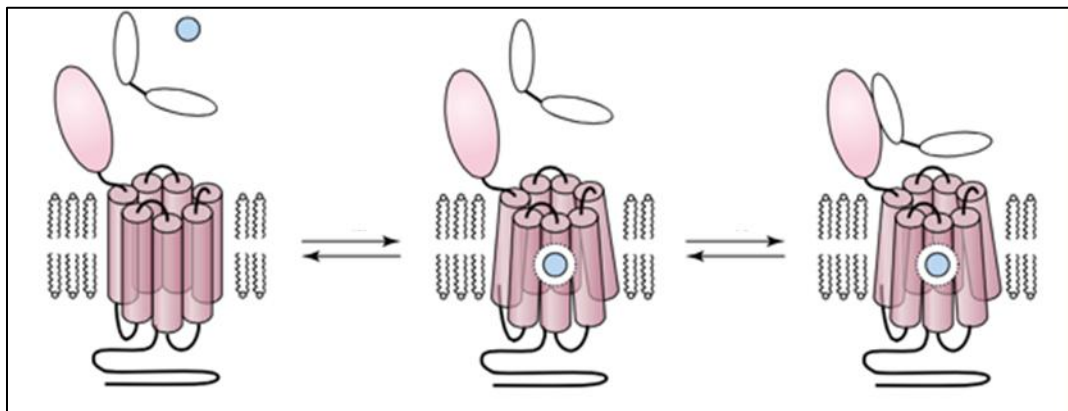
Figure 1-4: Non-Peptide Ligand Binding to Class B GPCRs.

thought to be perpetuated by overactive signaling cascades, small-molecule antagonists could ideally prevent the activation of the receptor and thereby circumvent peptide-stimulated signaling cascades.

1.5. Pituitary Adenylate Cyclase Activating Polypeptide (PACAP) and Corresponding Receptor (PAC1R)

In an effort to isolate novel hypophysiotropic neuropeptides, Arimura and coworkers screened fractions from the extracts of 4300 ovine hypothalamus by monitoring their stimulatory effect on adenylyl cyclase activity in cultured rat anterior pituitary cells. As a result of these screening efforts, Arimura and coworkers isolated a peptide, designated as pituitary adenylate cyclase-activating polypeptide (PACAP) (Arimura, 1991). It was discovered that this peptide strikingly increased the formation of cAMP. Characterization of PACAP indicated that it is comprised of 38 amino acid residues and is C-terminally α-amidated (Miyata et al., 1989). Two years following characterization, Chartrel et al. determined the primary structure of this 38-amino acid peptide (PACAP38) in the European green frog Rana ridibunda (Chartrel et al., 1991). Within the sequence of PACAP38, there exists an internal cleavage-amidation site at Gly28-Lys29-Arg30 (Arimura et al., 1991). This cleavage-amidation site suggests that the PACAP precursor can also generate a 27-residue α-amidated polypeptide (PACAP27). The primary structures for PACAP38 and PACAP27 are depicted in Figure 1-5.

PACAP38 HSDGIFYDSYSRYRKQMAVKKYLAAVLGKRYKQRVKNK-NH$_2$

PACAP27 HSDGIFYDSYSRYRKQMAVKKYLAAVL-NH$_2$

Figure 1-5: PACAP38 and PACAP27 Primary Structures

Following characterization and primary structure resolution of PACAP, the distribution of the peptide was determined in the brains of mammals and amphibians(Olejniczak, 2007, Koves et al., 1991, Yon et al., 1992). Using a radioimmunoassay (RIA) and rat brain tissue, Arimura et al. found the highest concentrations of PACAP to occur in the hypothalamic area (Arimura, 1991). Reversed-phase HPLC analysis indicated that PACAP38 accounted for more than 90% of the identified PACAP, while PACAP27 represented less than 10% of the total peptide content in brain tissue (Arimura, 1991).

Although the majority of PACAP was found in hypothalamic areas, the actual distribution of the peptide is much wider and includes such complex biological systems as the stomach, the pancreas, respiratory system and heart. This wide distribution of PACAP and its receptors throughout these complex biological systems suggests that the peptide may exert pleiotropic physiological functions. Indeed, it has been demonstrated that PACAP acts as a hormone, a neurohormone, a neurotransmitter, and a trophic factor in a number of tissues. This protein has been implicated in a broad range of biological processes including reproduction, development, growth, cardiovascular, respiratory, and digestive functions, immune responses, and circadian rhythms (Nussdorfer et al., 1998).

Important to this research is the central role the PACAP system plays in response to stress in activating the sympathoadrenomedullary and hypothalamic-pituitary- adrenal systems. As previously discussed, tissue studies conducted following subject exposure to extensive periods of stress have indicated that regions of the BNST with high levels of CRH also contain high levels of the neuropeptide PACAP and its class B G-protein

coupled pituitary adenylate cyclase 1 receptor (PAC1R) (Hammack et al., 2009, Finkelstein et al., 1985). Consequently, given the role of active PAC1 receptors in perpetuating symptoms of anxiety disorders such as PTSD, attention has recently turned towards studying the physiological effects of deactivating PAC1 receptors as a means of anxiety symptom management and control.

.

1.6. Non-Peptide Lead Compounds for PAC1R

A number of non-peptide antagonists have been identified for at least one third of class B GPCRs, however molecular details of their receptor interactions are not fully understood at this time, thus making the synthesis of non-protein PAC1R inhibitors difficult. Recently, Beebe et al. discovered small molecule hydrazide inhibitors of PAC1R (Beebe, 2008). Working from two lead hydrazides (Figure 1-6, Beebe et al., 2008) identified from the Abbot Laboratories compound library via a radioligand binding assay, a structure-activity relationship (SAR) was developed for small molecule compounds with high binding affinities for the PAC1 receptor. This hydrazide pharmacaphore (Figure 1-7, Beebe et al., 2008) consisted of an acyl phenol with a variable electron-withdrawing group in the meta position, a hydrazide linker, a middle
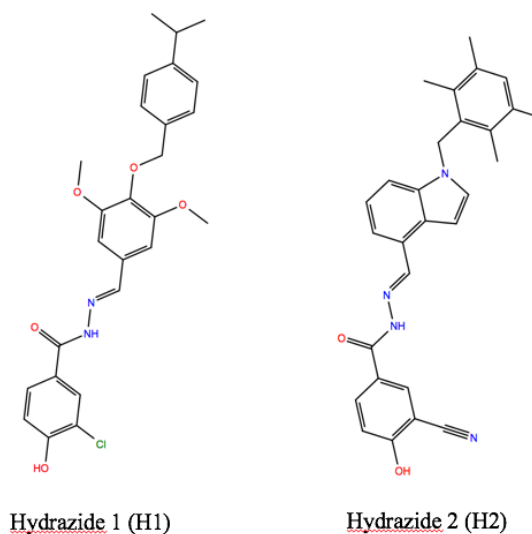
Figure 1-6: Small-Molecule Hydrazide Lead Compounds

arene ring with possible meta or para branching, and a distal ring with varying R-groups in the ortho, meta and/or para positions (Beebe et al., 2008). Prior to the work of Beebe and coworkers, drug-like small molecule inhibitors of the PAC_1 receptor were not known.
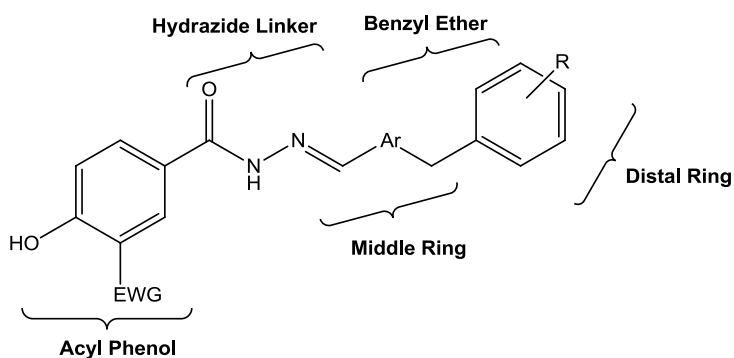
Figure 1-7: Hydrazide Pharmacaphore SAR

While Beebe and colleagues have successfully synthesized hydrazides that show significant binding affinity to the PAC1R, additional work is required to further elucidate the function and biological significance of this class B GPCR.

1.7. GPCR Crystallization

The crystallization of GPCRs has largely been hindered by their inherent flexibility, instability in detergents, and low levels of expression (Siu at al, 2013). Crystallization of Family B/ Secretin class receptors has been particularly tricky due to the presence of the large, globular extracellular domain coupled to the conformationally flexible, aliphatic 7TM bundle via a stalk structure, all of which are mobile as they have evolved to be highly dynamic in nature. Until 2013 there were experimentally solved structures available for the globular extracellular domains of some of these receptors, but no crystal structures for the 7TM bundles. In 2013, two structures of the TM bundle domains were released simultaneously, one documenting the glucagon receptor 7TM bundle (Siu et al., 2013) and the other showing the corticotropin-releasing factor receptor type 1 (CRFR-1) 7TM bundle (Hollenstein et al., 2013).

The glucagon receptor (GCGR) was crystallized in the inactive state using a small molecule antagonist, NNC0640 (Sui et al., 2013). Sui and co-workers synthesized an N-terminally truncated GCGR. Although the crystal structure contained truncated N- and C-terminal regions, Sui and co-workers built a model whereby they coupled the known crystal structure of GLP-1 bound GLP-1R N-terminal domain with the solved glucagon

15

TM domain structure in order to analyze potential positioning of the ligand-bound N-terminal domain with the TM region. From this model they proposed that the stalk region forms a complete α-helical structure leading to the N terminal domain, reinforced by intra-helical interactions, the extended extracellular loop 1 (ECL1) region and interactions with the middle α-helical region of glucagon. This suggests that the stalk region is critical to proper ligand orientation for optimal interaction with the TM domain (Siu et al., 2013).

This groundbreaking crystallization/modeling exercise has facilitated further strides in the attempt to elucidate possible small-molecule candidates for the treatment of PTSD based on their ability to bind to PAC1R. Using the newly crystallized glucagon receptor (GCGR) as a template, we can now use computer software to fold PAC1R via the model of another class B GPCR.

1.8. Computational Chemistry

Computational chemistry, or molecular modeling, encompasses all theoretical methods and computational techniques employed to mimic or simulate the behavior of molecules. This process is defined by a collection of computer based techniques for deriving, representing and manipulating the structures and reactions of molecules, and those properties that are dependent on these three dimensional structures. Molecular modeling allows analysts to use computers to visualize molecules, representing molecular structures numerically and simulating their behaviors with the principles and equations of

quantum and classical physics. In this process, the free energy of the system can be used to assess many interesting and informative aspects of the receptor-ligand complex (Krieger et al, 2003).

The energy of the system is a function of the type and number of atoms and their positions. Molecular modeling software is designed to calculate this energy efficiently and simulate molecular activities based upon this energy. In the absence of experimentally derived empirical data, such as crystal structures, information about chemical systems derived from these computational methods can be invaluable in informing research directions. It has been demonstrated that protein structures between homologues are conserved to a greater extent than sequences (Chothia et al., 1986). However, for proteins with less than 20% sequence identity, the three dimensional structures of related proteins can differ tremendously (Biasani et al., 2014). PAC1R shares 30% primary structure identity with its Class B homolog GCCR, rendering GCCR a reasonable template for this PAC1R-based computational research.

1.9. Homology Modeling

One division of computational chemistry, homology modeling, describes a software-based process by which to create theoretical prototypes for unsolved protein structures. This process involves the construction of a three dimensional atomic-resolution model of a protein (the target) based on its amino acid sequence and an experimentally resolved three-dimensional structure of a homologous protein (the

17

template).   Homology modeling software first arranges the target protein backbone identically to that of the template. In this process not only the positions of alpha carbons, but also the phi and psi angles and secondary structure, are made identical to the template. Following backbone alignment, sophisticated homology modeling packages then adjust sidechain positions to minimize collisions, and offer further energy minimization or molecular dynamics in an attempt to improve the model.

Swiss-Model is a highly refined bioinformatics web-based server dedicated to the homology modeling of protein structures (Biasini et al., 2014). The accuracy, stability and reliability of the Swiss-Model server had been previously validated by the EVA benchmark project. Created by Volker Eyrich at Columbia University, EVA was a large-scale and continuously running web server that automatically assessed protein secondary and tertiary structure prediction methods via weekly comparison to newly solved protein structures deposited in the Protein Data Bank (Eyrich et al., 2001). Currently, Swiss-Model is participating in the CAMEO3D (Continuous Automated Model EvaluatiOn) project which functions much like the now defunct EVA project in evaluating the accuracy and reliability of protein structure prediction services in a fully automated manner (Haas et al., 2013). As a well-established and reputable source for building theoretical protein structures, Swiss-Model was an integral part of this research.

1.10. Molecular Docking

Molecular docking describes a method which predicts the preferred orientation of

one molecule to a second when bound to each other to form a stable complex. This computational approach can be used to model the interaction between a small molecule and a protein at the atomic level. The docking process involves two basic steps: prediction of the ligand conformation as well as its position and orientation within docking sites (usually referred to as a pose) and assessment of the binding affinity. Additionally, when a binding site is unknown, the search for the binding site as well as for the ligand's position in that binding site, are performed simultaneously (Krieger et al., 2003). Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to predict the affinity and activity of the small molecule. As such, by enhancing understanding of the binding complex geometry, molecular docking studies have the potential, for example, to inform modifications of lead compounds in optimization of drug potency. It is with this goal in mind, to computationally evaluate potential small molecule antagonists for PAC1R, that molecular docking research has herein been undertaken.

Designed and implemented by Oleg Trott in the Molecular Graphics Lab at The Scripps Research Institute, AutoDock is one of many reputable and tested molecular docking software tools available to computational chemists. AutoDock is by far the most cited docking software in academic publications, accounting for more than 27% of citations in peer-reviewed journals with the remaining 72% of citations divided fairly evenly among 24 other docking software programs (Sousa et al., 2014). AutoDock Vina, an improved version of AutoDock, when used in conjunction with a molecular graphics visualization system, Schrodinger LLC's PyMol, allows for a thorough docking study

that includes binding site definition and adjustment, automatic file preparations for receptor definition, straightforward selection of flexible residues, ligand file preparation, viewing of docking poses, and analysis and export of virtual screening results.

CHAPTER 2: SYNTHESIS OF SMALL MOLECULE HYDRAZIDE ANTAGONISTS OF THE PITUITARY ANDENYLATE CYCLASE-ACTIVATING POLYPEPTIDE RECEPTOR TYPE 1

2.1. Synthetic Research Objectives

It was the initial goal of this research to synthesize and evaluate novel hydrazide compounds as potential antagonists for PAC1 receptors. To build and hone laboratory skills in the general process of producing these hydrazides, three previously synthesized compounds (Figure 2-1) were the focus of the first phase of this research.
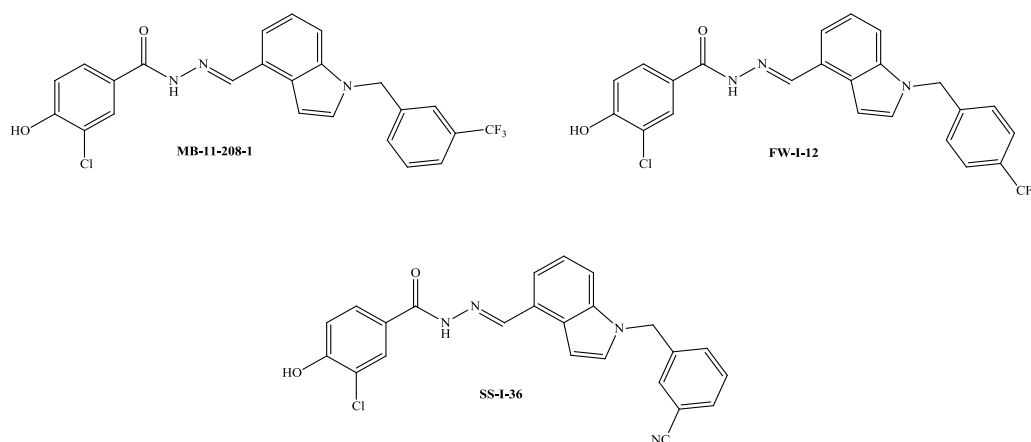
Figure 2-1: Initial Target Compounds

2.2. Retrosynthetic Plan

As shown in retrosynthetic Figure 2-2, these hydrazide compounds (**1**) were made from the condensation of 3-chloro-4-hydroxybenzohydrazide (**2**) and an N-alkylated indole species (**3**) with varying R- groups on the distal ring.
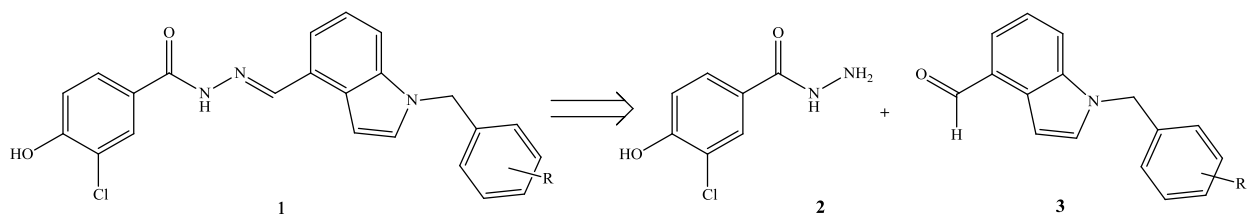
Figure 2-2: Retrosynthetic Plan

2.2.1. Synthesis of 3-Chloro-4-Hydrobenzohydrazide

The synthesis of 3-chloro-4-hydronezohydrazide, compound (**2**), was achieved via the addition of hydrazine to methyl-3-chloro-4-hydroxybenzoate (**4**) in ethanol under reflux conditions overnight (Figure 2-3). The product was isolated by suction filtration, purified by hot filtration recrystallization and characterized as sufficiently pure by
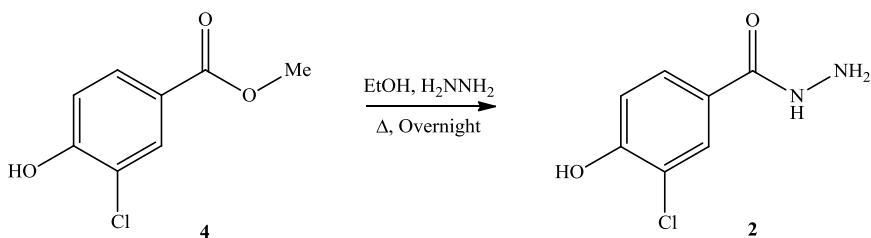
Figure 2-3: Synthesis of 3-Chloro-4-Hydroxybenzohydrazide (2)

comparison of ^{1}H NMR spectroscopy results to known spectra. This reaction proceeded at 82% yield and provided 4-hydroxy-3-chlorobenzohydrazide (**2**) as a fluffy white solid.

2.2.2. Synthesis of N-Alkylated Aldehyde Indole

The synthetic route shown in Figure 2-4 illustrated the production of the N-alkylated aldehyde indole. The N-alkylated aldehyde indole species (**3**) was made by treating 1*H*-indole-4-carbaldehyde (**5**) in dry dimethylformamide with sodium hydride at 0 °C to deprotenate the indole nitrogen. N-alkylation was then accomplished by allowing the deprotenated indole to stir at room temperature for 2 hours with variously substituted aryl bromide species. The reaction mixture was extracted with ethyl acetate, washed with water and brine, dried over magnesium sulfate and concentrated in vacuo. Product
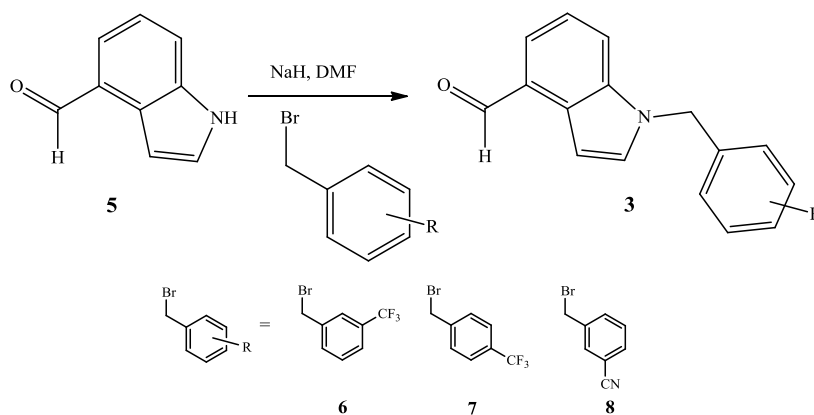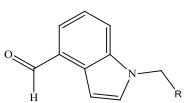
Figure 2-4: Synthesis of N-Alkylated Aldehyde Indole Species (3)

purification was achieved by flash column chromatography in dichloromethane and verified by ^{1}H NMR spectroscopy with a comparison to known spectra. Product yields are summarized in Table 1.

Table 1: N-Alkylated Indole Species Product Yields

Entry		Yield
6	(3-CF$_3$-benzyl)	29%
7	(4-CF$_3$-benzyl)	23%
8	(3-CN-benzyl)	65%

2.2.3. Synthesis of Completed PAC1 Receptor Antagonists

The final target compounds were synthesized via condensation reactions of the hydrazide moiety with the N-alkylated indole species (Figure 2-5). To the indole species (**3**) in dimethylsulfoxide was added 4-hydroxy-3-chlorobenzohydrazide (**1**) and a catalytic amount of acetic acid. The reaction was allowed to stir overnight at 50 °C. The resulting mixture was cooled and added to ethyl acetate, washed with water and brine, dried over magnesium sulfate and concentrated in vacuo. The product was purified by hot filtration recrystallization in 4:1 hexanes to ethyl acetate and verified sufficiently pure
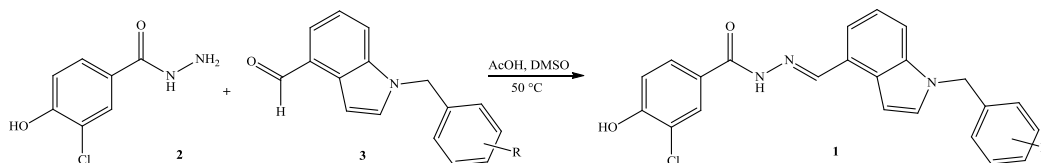
Figure 2-5: Synthesis of Completed PAC1R Antagonists (1)

by ^{1}H NMR spectroscopy comparison to known spectra. Product yields are summarized in Table 2.

Table 2: Completed Hydrazide Product Yields
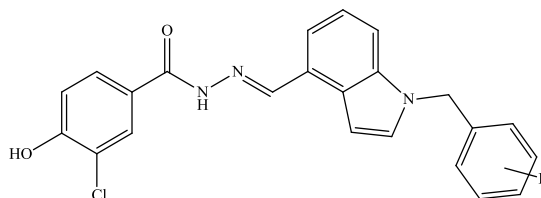
Entry		Yield
9		49%
10		18%
11		64%

2.3. Additional Synthetic Research Goals

Following the successful synthesis of the initial target compounds, research goals turned toward preparing novel hydrazide analogs. Initially, three compounds (Figure 2-6) were requested by Dr. Victor May's laboratory for screening purposes as potential PAC$_1$ receptor antagonists. Unfortunately, the halogen-benzyl starting materials for two of the compounds (3,4,5-triflouromethyl and 3,5-nitrile substituted moieties) were not commercially available and lengthy/complicated synthetic steps rendered in-house

preparation of these starting materials unrealistic. Furthermore, synthesis of the third novel hydrazide compound was temporarily delayed by Dr. May's discovery that his cell cultures used for determining ligand-receptor activity were potentially contaminated.
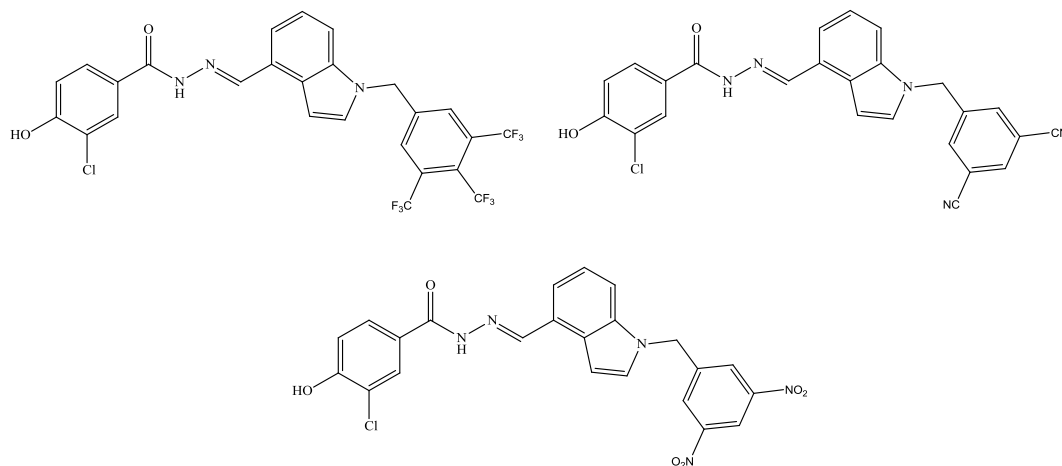
Figure 2-6: Novel Hydrazide Analog Targets

2.4 Shift in Research Focus

In addition to synthetic and cell-culture issues, the rigorous demands of full-time work outside of the academic setting and parenting a young child forced a dramatic change in research goals. Focus shifted to a non lab-based approach to developing small molecule antagonists for the PAC1R. The recent crystallization of GCCR discussed in the introduction above provided many opportunities for computer-based analysis of the PAC1R and possible ligands. This research thus shifted from a synthetic focus to a computational analysis.

2.5. Experimental

2.5.1. Synthesis of 3-Chloro-4-Hydrobenzohydrazide (2)

To methyl-3-chloro-4-hydroxybenzoate (2.001 g, 11 mmol) dissolved in ETOH (50 mL) was added Hydrazine (1.8 mL). The reaction mixture was refluxed overnight under nitrogen. When the mixture was cooled, the desired product crystallized. The white solid was isolated by filtration recrystallization from hot ETOH affording 1.64 g (82%) of 3-chloro-4-hydroxy benzoic acid hydrazide.

2.5.2. Synthesis of 3-((4-formyl-1H-indol-1-yl)methyl) benzotrifluoromethylene (3)

To a solution of 4-formylindole (0.303 g, 2.09 mmol) in dry DMF (9 mL) under an atmosphere of nitrogen was added slowly NaH (0.094 g, 60% wt/wt, 2.35 mmol) and allowed to stir at 0°C for 30 min. To the mixture was added 3 trifluoromethyl 1 benzyl bromide (0.39 mL, 2.56 mmol) and allowed to stir for 2 hours at room temperature. Reaction was monitored via TLC. The reaction mixture was extracted with EtOAc, washed with H2O (3x) and brine, dried over MgSO4 and concentrated in vacuo to form a crude product (0.575 g). The Crude product was purified by flash column chromatography in DCM to yield 0.029 g (29%).

2.5.3. Synthesis of 3-((4-formyl-1H-indol-1-yl)methyl) benzotrifluoromethylene (3)

To a solution of 4-formylindole (0.337 g, 2.322 mmol) in dry DMF (9 mL) under an atmosphere of nitrogen was added slowly NaH (0.126 g, 60% wt/wt, 3.150 mmol) and allowed to stir at 0°C for 30 min. To the mixture was added 4-trifluoromethyl-1-benzyl bromide (0.560 mL, 2.342 mmol) and allowed to stir for 2 hours at room temperature. Reaction was monitored via TLC. The reaction mixture was extracted with EtOAc, washed with H_2O (3x) and brine, dried over $MgSO_4$ and concentrated in vacuo to yield 0.378 g crude product as a dark brown oil. The crude product was purified by flash column chromatography in DCM to yield 0.016 g. (23%).

2.5.4. Synthesis of 3-((4-formyl-1H-indol-1-yl)methyl) benzonitrite (3)

To a solution of 4-formylindole (0.300 g, 2.07 mmol) in dry DMF (9 mL) under an atmosphere of nitrogen was added slowly NaH (0.095 g, 60% wt/wt, 2.38 mmol) and allowed to stir at 0°C for 30 min. To the mixture was added 3-(bromomethyl) benzonitrile (0..486 mL, 2.48 mmol) and allowed to stir for 2 hours at room temperature. Reaction was monitored via TLC. The reaction mixture was extracted with EtOAc, washed with H_2O (3x) and brine, dried over $MgSO_4$ and concentrated in vacuo to form a crude product (0.331 g). The crude product was purified by flash column chromatography in DCM to yield 0.104 g, 65%.

2.5.5. Synthesis of (E)-3-cholor-N'-((1)-2-trifluoromethylbenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide (1)

To 3-((4-formyl-1H-indol-1-yl)methyl) benzotrifluoromethyl (0.195 g, 0.413 mmol) dissolved in dry DMSO (1 mL) was added 4-hydroxy-3-cholorbenzohydrazide (0.077 g, 0.413 mmol) and 1 drop of AcOH under N_2 and stirred overnight at 50 ºC. The resulting mixture was allowed to dry over MgSO4, and concentrated in vacuo. The resulting solid was recrystallized from EtOAc/Hexanes to yield 0.095 g (49%) of (E)-3-cholor-N'-((1)-2-trifluoromethylbenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide as a dark brown solid.

2.5.6. Synthesis of (E)-3-cholor-N'-((1)-3-trifluoromethylbenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide (1)

To 4-((4-formyl-1H-indol-1-yl)methyl) benzotrifluoromethyl (0.100 g, 0.384 mmol) dissolved in dry DMSO (1 mL) was added 4-hydroxy-3-cholorbenzohydrazide (0.072 g, 0.384 mmol) and 1 drop of AcOH under N_2 and stirred overnight at 50 ºC. The resulting mixture was allowed to dry over MgSO4, and concentrated in vacuo. The resulting solid was recrystallized from EtOAc/Hexanes to yield 0.035 g (18%) of (E)-3-cholor-N'-((1)-3-trifluoromethylbenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide as a dark brown solid.

2.5.7. Synthesis of (E)-3-cholor-N'-((1)-3-cyanobenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide (1)

To 3-((4-formyl-1H-indol-1-yl)methyl) benzonitrite (0.100 g, 0.384 mmol) dissolved in dry DMSO (1 mL) was added 4-hydroxy-3-cholorbenzohydrazide (0.072 g, 0.384 mmol)

and 1 drop of AcOH under N_2 and stirred overnight at 50 ºC. The resulting mixture was allowed to dry over MgSO4, and concentrated in vacuo. The resulting solid was recrystallized from EtOAc/Hexanes to yield 0.104 g (64%) of (E)-3-cholor-N'-((1)-3-cyanobenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide as a dark brown solid.

2.6. NMR Data

^{1}H NMR Analysis of 3-Chloro-4-Hydrobenzohydrazide (3)
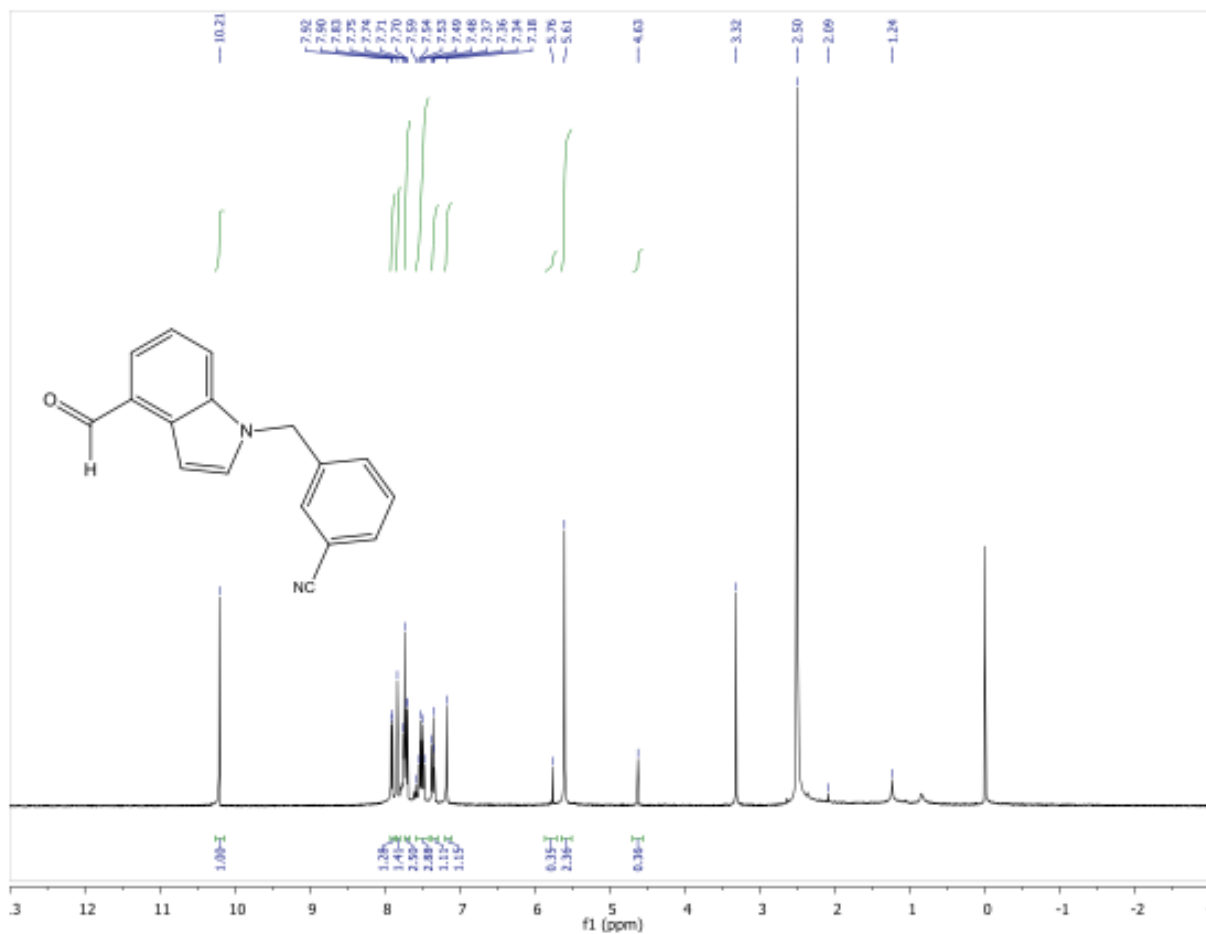
^{1}H NMR Analysis of 3-((4-formyl-1H-indol-1-yl)methyl) benzonitrite (4)
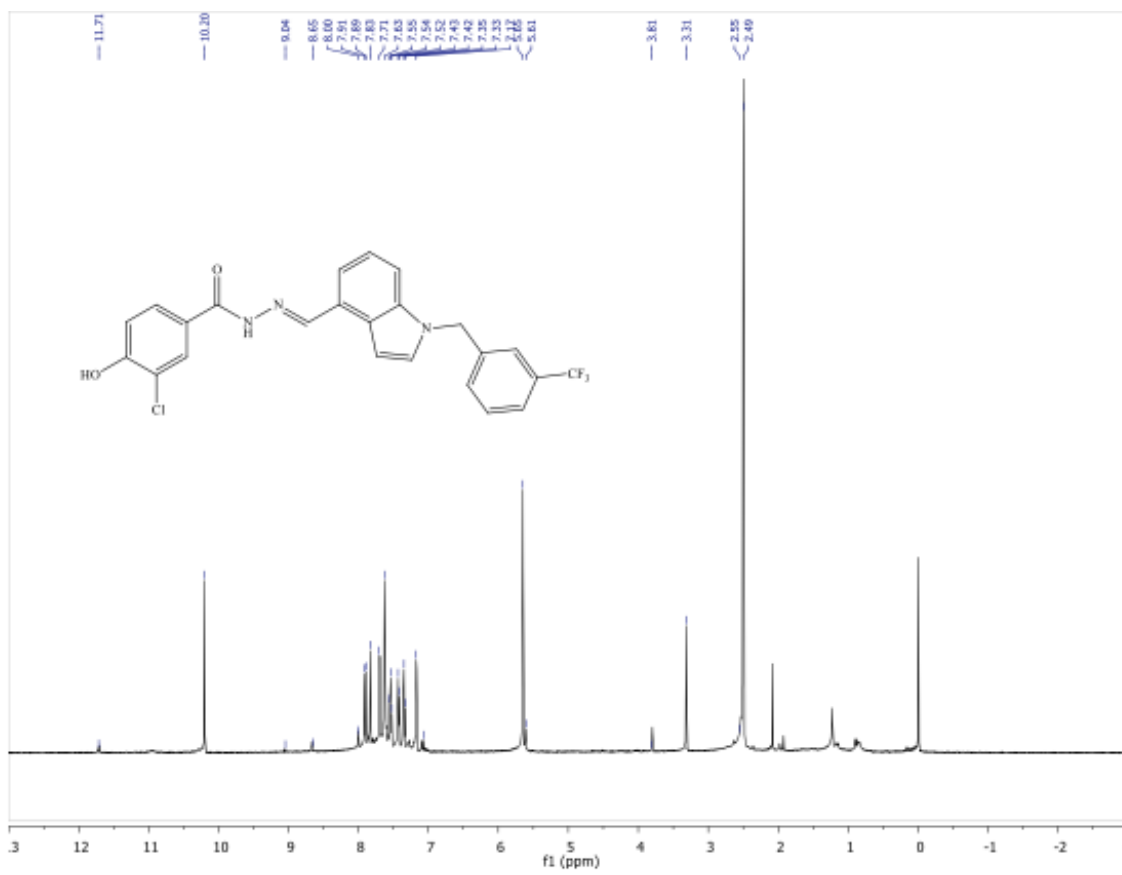
^{1}H NMR Analysis of (E)-3-cholor-N'-((1)-2-trifluoromethylbenzyl)-1H-indol-4-yl)methylene)-4-hydroxybenzohydrazide (1)

CHAPTER 3: HOMOLOGY MODELING AND THE DOCKING OF LIGANDS TO PITUITARY ANDENYATE CYCLASE 1 RECEPTOR (PAC1R)

3.1. Class B G-Protein Coupled Receptor Glucagon (GCGR) Crystallization

Many have attempted comparisons between well-known class A GPCRs and less understood class B GPCRs based on the known structures of the former group and the known sequences of the latter group. However, because there is generally less than 15% sequence homology between Class A and Class B GPCRs, these modelling studies are arguably inconsequential to furthering our understanding of Class B GPCRs. In July of 2013, as briefly discussed in the introduction, Siu and coworkers reported in *Nature* a crystal structure for the 7TM helical domain of the human glucagon class B G-protein-coupled receptor (Siu et al., 2013). This groundbreaking crystallization of the glucagon receptor has ignited greater strides in the quest to understanding Class B GPCRs and provided a path towards the elucidation of possible small-molecule drug candidates in the treatment of PTSD. Using the newly crystallized glucagon receptor (GCGR) as a template, we can now use computer software to simulate the three dimensional structure of the PAC1 receptor and glean from this model information pertinent to the development of small molecule PAC1R antagonists.

3.2. Comparison of PAC1R Amino Acid Sequence with GCGR Sequence

Human PAC1R consists of 468 amino acids and is encoded by the gene

adenylate cyclase activating polypeptide 1 (pituitary) receptor type I (ADCYAP1R1) (Ogi et al, 1993, Stoffel et al., 1994). The primary amino acid sequence for PAC1R, as well as for GCGR, was obtained from UniProtKB. Utilizing the UniProt Clustal Omega Align Tool (www.UniProt.org) (Sievers et al., 2011, Goujon et al., 2010, McWilliam et al., 2013), the PAC1R sequence (UniProt identification P41586) and GCGR sequence (UniProt identification P47871), were compared via a pairwise alignment. As illustrated in Figure 3-1, UniProt Align Tool results indicate 160 identical positions between PAC1R and GCGR yielding an identity of 30.948% with an additional 122 similar positions. Highlighted amino acids represent the 7TM portions of each protein. An asterisk (*) below the matched pair indicates positions which have a single, fully conserved residue. A colon (:) indicates conservation between groups of strongly similar properties and a period (.) indicates conservation between groups of weakly similar properties. A 30.9% identity, combined with a large number of amino acid pairings with strongly similar properties within the 7TM regions, indicates that GCGR is an appropriate model upon which to build a tertiary PAC1R structure as the basis for SAR analyses.

```
P41586 PACR_HUMAN      1   MAGVV--HVSLAALLLLPMAPAMHS----DCIFKKEQ---AMCLEKIQRANELMGFNDSS        51
P47871 GLR_HUMAN       1   MPPCQPQRPLLLLLLLLACQPQVPSAQVMDFLFEKWKLYGDQCHHNL-----SLLPPPTE       55
                           *       :  *  ****    *  :  *    *  :*:*  :     *  .::      :   :.

P41586 PACR_HUMAN     52   PGCPGMWDNITCWKPAHVGEMVLVSCPELFR------------IFNPDQVWETETIGESD       99
P47871 GLR_HUMAN      56   LVCNRTFDKYSCWPDTPANTTANISCPWYLPWHHKVQHRFVFKRCGPDGQWVRGPRG---      112
                            *    :*: :**  :  .   . :*** :              ** *      *

P41586 PACR_HUMAN    100   FGDSNSLDLSDMGVVSRNCTEDGWSEPFPHYFDACGFDEYESE-TGDQDYYYLSVKALYT      158
P47871 GLR_HUMAN     113   --------------------QPWRDA-----SQCQMDGEEIEVQKEVAKMYSSFQVMYT      146
                                               : * :      . * :*  * *  :    * *.:.:**

P41586 PACR_HUMAN    159   VGYSTSLVTLTTAMVILCRFRKLHCTRNFIHMNLFVSFMLRAISVFIKDWILYAEQDSN-      217
P47871 GLR_HUMAN     147   VGYSLSLGALLLALAILGGLSKLHCTRNAIHANLFASFVLKASSVLVIDGLLRTRYSQKI      206
                           **** ** :*   *:.**   : ******* ** ***.**:*:* **:: * :* :. ..:

P41586 PACR_HUMAN    218   ------HCF---ISTVECKAVMVFFHYCVVSNYFWLFIEGLYLFTLLVETFFPERRYFYW      268
P47871 GLR_HUMAN     207   GDDLSVSTWLSDGAVAGCRVAAVFMQYGIVANYCWLLVEGLYLHNLLGLATLPERSFFSL      266
                                  :    :.. *:.. **:*: :*:** **::*****..**  : :*** :*

P41586 PACR_HUMAN    269   YTIIGWGTPTVCVTVWATLRLYFDDTGCWDMNDSTALWWVIKGPVVGSIMVNFVLFIGII      328
P47871 GLR_HUMAN     267   YLGIGWGAPMLFVVPWAVVKCLFENVQCWTSNDNMGFWWILRFPVFLAILINFFIFVRIV      326
                           *  ****:*  : *. **.:: *::. **  **. ..:**::: **. :*::**.:*: *:

P41586 PACR_HUMAN    329   VILVQKLQSPDMGGNESSIYLRLARSTLLLIPLFGIHYTVFAFSPENV----SKRERLVF      384
P47871 GLR_HUMAN     327   QLLVAKLRARQMHHT--DYKFRLAKSTLTLIPLLGVHEVVFAFVTDEHAQGTLRSAKLFF      384
                           :** **:: :*  .   .  :***:*** ****:*:* .****  ::    : :*.*

P41586 PACR_HUMAN    385   ELGLGSFQGFVVAVLYCFLNGEVQAEIKRKWRSWKVNRYFAVDFKHRHPSLASS------      438
P47871 GLR_HUMAN     385   DLFLSSFQGLLVAVLYCFLNKEVQSELRRRWHRWRLGKVLWEERNTSNHRASSSPGHGPP      444
                           :* *.****:;********* ***:*::*:*: *:: : :  : :  .    :**

P41586 PACR_HUMAN    439   ------GVNGGTQLSILSKSSSQIRMSGLPADNLAT-      468
P47871 GLR_HUMAN     445   SKELQFGRGGGSQDSS----AETPLAGGLPRLAESPF      477
                                 *  **:* *     :.    .***    :
```
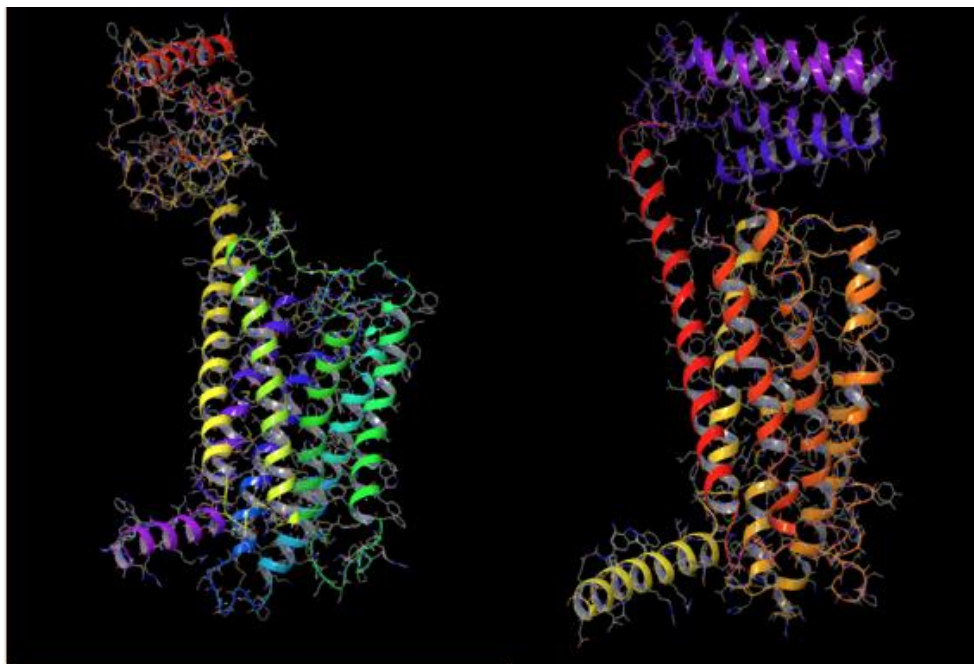
Figure 3-1: UniProt Clustal Omega Align Tool Results

3.2.1. Use of GCGR Model to Fold PAC1R

Having determined GCGR a viable model for folding PAC1R, the protein sequence was submitted to Swiss-Model, a web-based service for protein structure prediction, with instructions to fold PAC1R using the crystal structure determined for GCGR as a template (Biasini et al., 2014, Bordoli et al., 2006, Kiefer et al., 2009, Guex et al., 2009). This web-based software was able to recognize GCGR as a template, align the primary structures of GCGR and PAC1R while making any necessary alignment corrections, generate a backbone for PAC1R, model PAC1R loops and side-chains, and finally optimize the PAC1R model. A .pdb file was generated and opened via PyMOL.

Figure 3-2: Side By Side Comparison of PAC1R (Left) and GCGR (Right)

The resultant folded PAC1R is depicted in Figure 3-2 side by side with the GCGR template.

As is consistent with the GCGR template, the PAC1R model consists of an extracellular region, depicted at the top portion of Figure 3-3 in red, and seven membrane-spanning, or transmembrane, helices (the 7TM region) depicted in multiple colors in the bottom portion of the figure. Also present is an eighth helix in the intracellular domain which is nearly perpendicular to the 7TM helices. Figure 3-3 represents PAC1R with helices labeled.
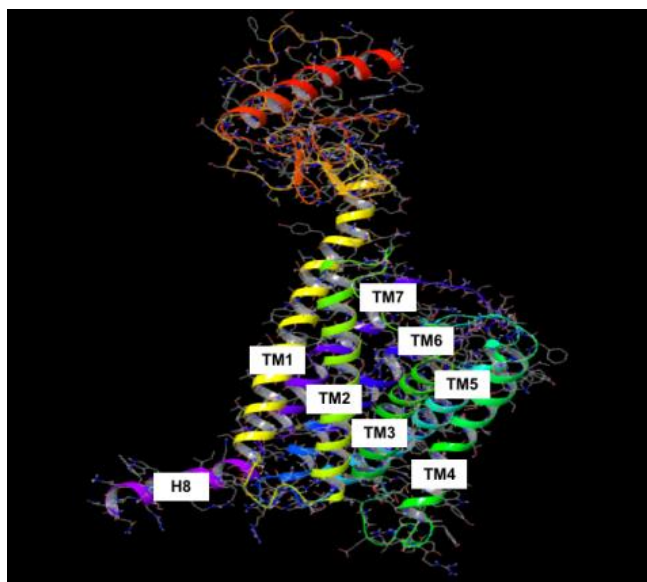
Figure 3-3: Folded Model of PAC1R

3.2.2. Homology Model Validation

The Swiss-Model software reported a Z-score for this PAC1R model of -4.56, and a sequence homology between template (GCGR) and target (PAC1R) of 42.94%. Z-score assignment reflects estimation of the local and global quality of the model. The local geometry is analyzed by a torsion angle potential over three consecutive amino acids. Two pairwise distance-dependent potentials are used to assess all-atom and C-beta interactions. A solvation potential describes the burial status of the residues. (Benkert et al., 2008). Figure 3-4 provides an visual estimate of the absolute quality of the model by relating it to reference structures from the PDB that were experimentally solved by X-ray crystallography or NMR analysis. This comparison is meant as an estimate of the "degree of nativeness" of the structural features observed in the model by describing the likelihood that this model is of comparable quality to high-resolution experimental

structures.

In Figure 3-4, the area built by the circles colored in different shades of grey in the plot represent the Z-scores of the reference structures from the PDB. The PAC1R model's Z-score (black dot) is compared to the scores obtained for experimental structures of similar size (model size +/- 10%). Although the PAC1R Z-score is within range of those from experimentally determined structures, these quality estimates for membrane-bound proteins need to be treated with caution: membrane proteins may receive very high Z-scores since their physico-chemical properties differ considerably from those of soluble proteins. A Z-score version with separate potentials optimized for membrane proteins is under development (Benkert et al., 2014)
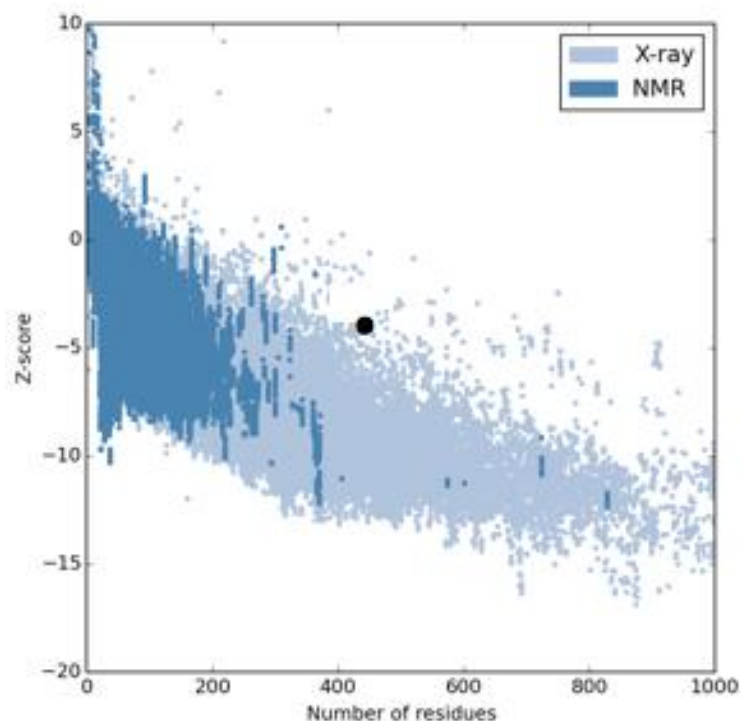

Figure 3-4: Z-Score Comparison of PAC1R (Black Dot) to Experimentally Resolved Proteins

Schrodinger's PyMol was used to align the GCGR template and PAC1R target structures and calculate the root-mean-square deviation (RMSD) of atomic positions. Although the extracellular domains do not align (Figure 3-5), the J-domains do align with an RMSD of 1.807 Å for 226 atoms. Although a good model is generally considered one with an RMSD of up to 1.5 Å, transmembrane proteins are permitted a bit more leeway in acceptable RMSD values (Arnold et al., 2006). Furthermore, a sequence homology of 30.9%, though within acceptable limits for a template-target modeling task, would not be expected to generate a terribly low RMSD score. Therefore, this calculated value of 1.8 Å for the 7TM domain is quite acceptable.
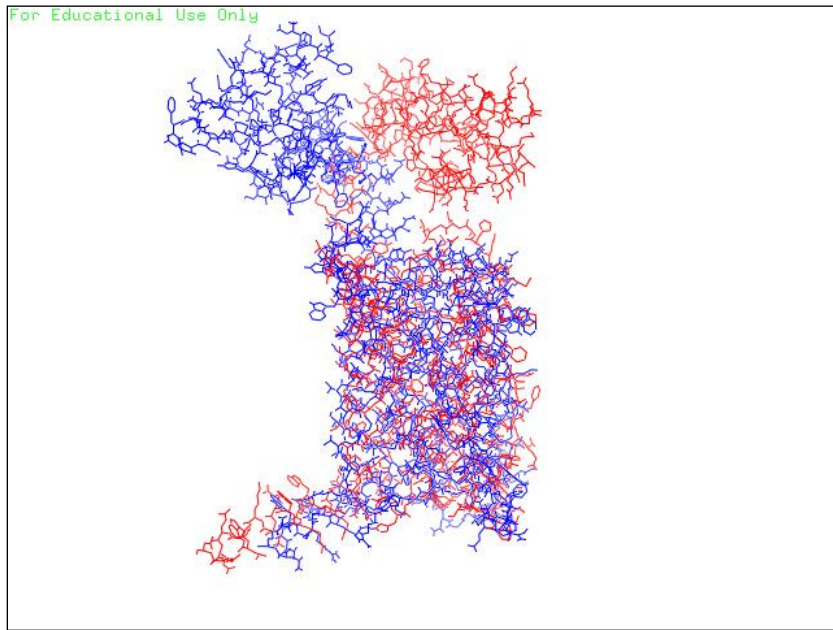

Figure 3-5: GCGR (Blue) and PAC1R (Red) RMSD Alignment

The Ramachandran plot, shown in Figure 3-6, generated using PDBSum PROCHECK software (de Beer et al., 2014), depicts 90% of the residues situated in the most energetically favorable regions of the plot (red colored low energy A, B and L areas). Only 0.6% (or 2) of the 398 plotted residues are situated within disallowed areas of the graph. For each type of the secondary structure elements there is a characteristic range of torsion angle values, which can clearly be seen on the Ramachandran plot of PAC1R: the region marked A is for alpha-helices and B is for beta-sheets. These results suggest that this PAC1R model is theoretically reliable and may actually reflect the true three dimensional structure of the protein (Morris et al., 1992).
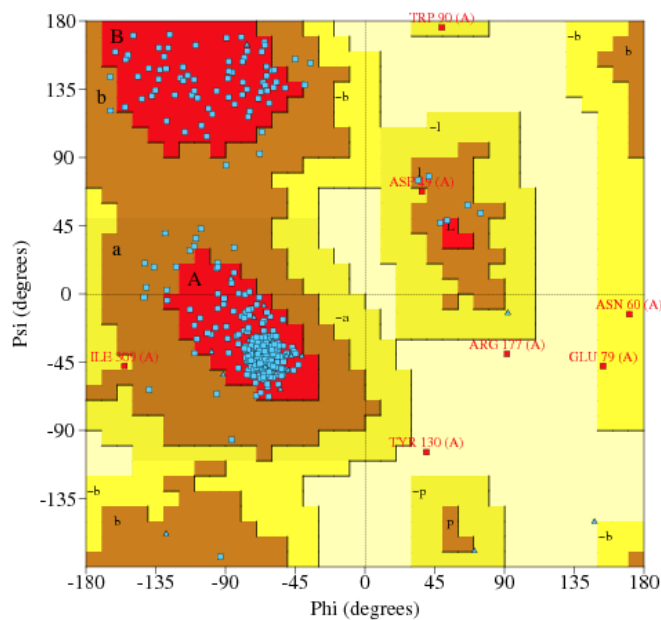

Figure 3-6: Ramachandran Plot of PAC1R Residues

41

3.3. Analysis of Theoretical PAC1R Structure

Knowledge of the three-dimensional structure of this GPCR is important for understanding molecular mechanisms underlying diseases and syndromes associated with PAC1R. In addition to primary structure similarities discussed in section 3.2 above, there are a number of valuable comparisons between the PAC1 receptor and other class B GPCRs that may elucidate structure-function relationships and inform the design of small molecule drug antagonists in the treatment of post-traumatic stress and other disorders that may be regulated by PAC1R.

There are a number of conserved features among Class B GPCRs that may contribute significantly to structure-activity relationships of the PAC1 receptor. Although this receptor is a model based closely on the template of GCGR, it is important to determine the presence of expected Class B GPCR features, including conserved residues and anticipated non-bonding interactions, as a means for validating the accuracy of this model.

3.3.1 Disulfide Bridges

Disulfide bridges, or covalent bonds between the sulfur atoms of cysteine residues, play important roles in the folding and stabilization of many proteins. Class B GPCRs are generally thought to possess 3 disulfide bridges in the extracellular domain (ECD) of the receptor (Siu et. al., 2013). An examination of the PAC1 receptor model confirms the expected presence of these cysteine-based sulfur bridges. As illustrated in

Figure 3-7, disulfide bonds indeed form between CYS 77 and CYS 134, CYS 34 and CYS 63, and between CYS 54 and CYS 118. These bonds are highly responsible for the three dimensional structure of the ECD.
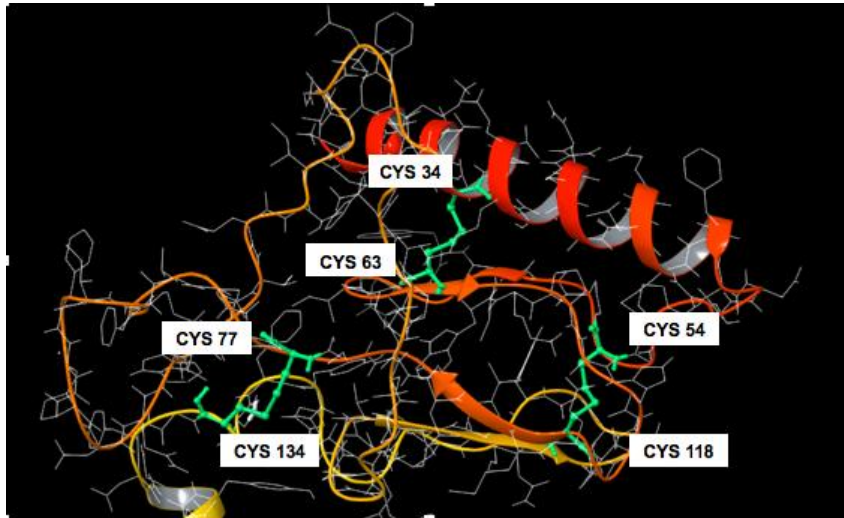

Figure 3-7: Disulfide Bridges in the Extracellular Domain of PAC1R.

Additionally, Class B GPCRs exhibit a conserved disulfide bond between the extracellular loop that connects TM2 and TM3 (ECL1) and the extracellular loop that connects TM4 and TM5 (ECL2). This bond, illustrated in Figure 3-8, is present in this PAC1R model and spans between CYS 226 on ECL1 and CYS 296 on ECL2. This disulfide bridge is thought to stabilize the receptor's 7TM fold (Siu et. al., 2013).
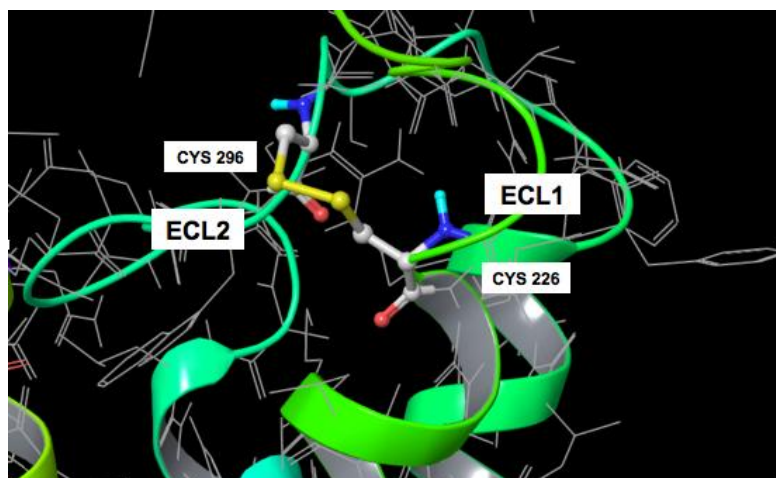
Figure 3-8: Disulfide Bond Between CYS 296 and CYS 226.

3.3.2. Conserved Tyrosine Residue

A tyrosine residue located on TM6 (Figure 3-9) is thought to be conserved among Class B GPCRs. TYR 348 is postulated to play an important role during receptor activation by moving into the space created by the shift of TM6 upon G-protein binding
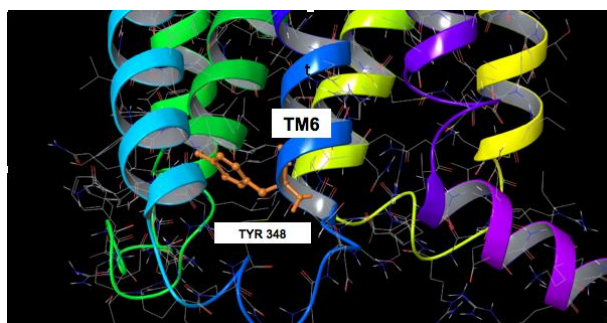

Figure 3-9: Mechanistically Important TYR 348

to the intracellular portion of the receptor (Worth et. al., 2013). Deemed a microdomain, this receptor residue highly conserved in the class is believed to have a mechanistically important role for protein function/activation (Venkatakrishnan et. al., 2014). Studies of

44

other GPCRs indicate that TM6 necessarily moves toward cell lipids to facilitate the docking of a G-protein (Siu et. al., 2013). The movement of TYR 348 in to the space created by the pre-docking shift of TM6 is thought to stabilize the G-protein coupling conformation (Bortolato et. al., 2014).

3.3.3. GWGxP Motif

An additional feature of PAC1R that is generally conserved among GPCRs is a residue ordering motif consisting of GLY-TRP-GLY-x-PRO. This residue sequence is located on TM4 and is proposed to play a key role in stabilizing the configuration of
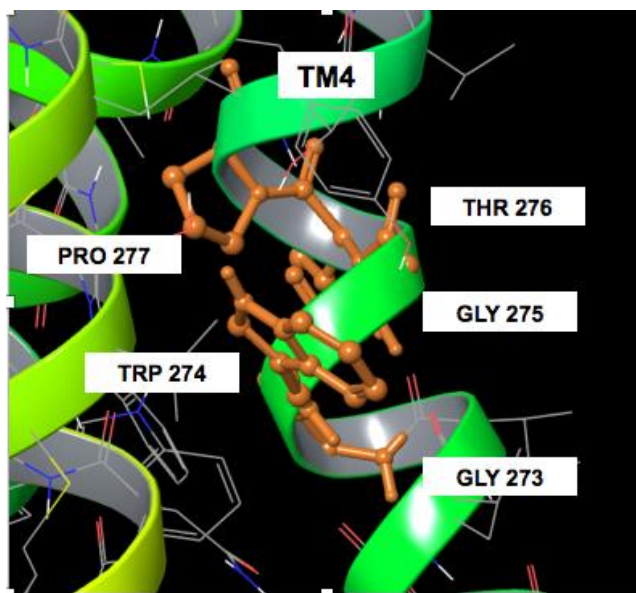

Figure 3-10: GWGxP Motif on PAC1R TM4

TM2, TM3 and TM4 relative to each other. Illustrated in Figure 3-10, this motif is demonstrated in PAC1R by GLY273, TRP 274, GLY 275, THR 276 and PRO 277. The stabilizing effect between TM2, TM3 and TM4 occurs as the TRP 276 residue engages in

hydrogen bonding with ASN 191 on TM2 and TYR 235 on TM3 (Bartolato et. al., 2014). The residues involved in this non-bonding stabilization interaction are depicted in Figure 3-11.
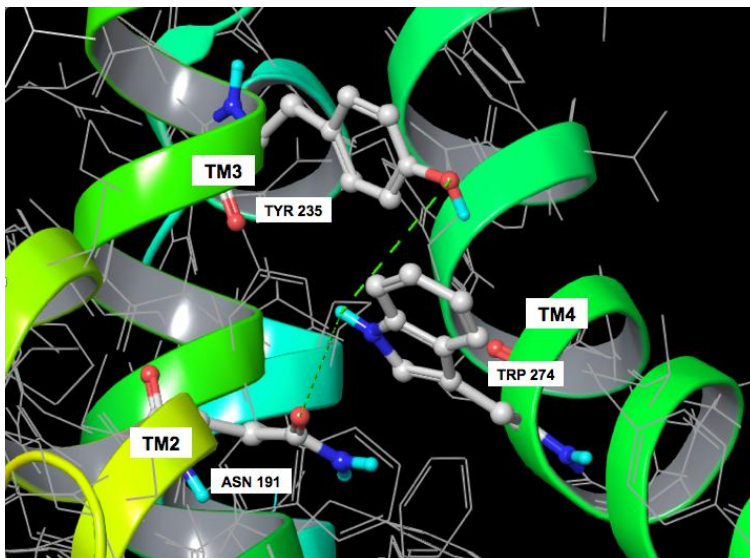

Figure 3-11: Non-Bonding Interactions Between PAC1R TM2, TM3 and TM4

3.3.4. Conserved Glutamic Acid Residue

As TRP 274 non-bonding interactions with TYR 235 and ASN 191 stabilize TM2, TM3, and TM4, a conserved glutamic acid residue (GLU 406) of Helix 8 (H8) forms a Class B-specific ionic network with ARG 185 on TM1 and ARG 350 on TM6. Two interhelical salt bridges form between GLU 406 and the two arginine residues. This interaction may contribute to the tilt of H8 approximately 25º away from the membrane, which is in turn thought to impact the coupling of G-proteins (Siu et al., 2013). The ionic interactions of this highly conserved Class B GPCR 3-residue relationship are depicted in Figure 3-12.
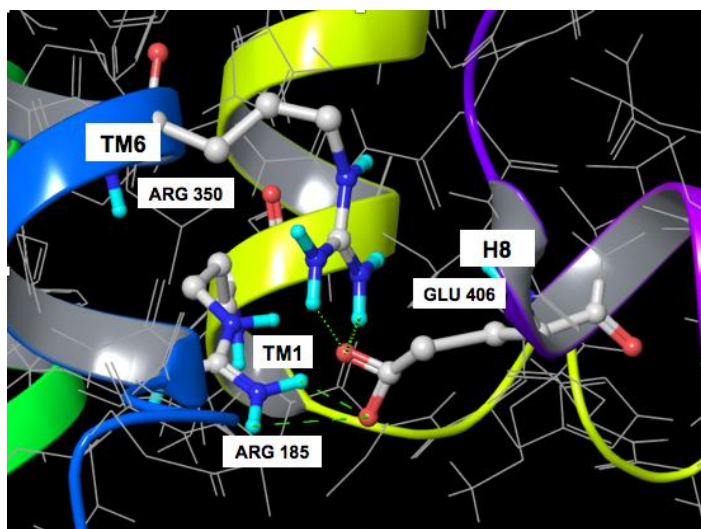
Figure 3-12: GLU 406 Ionic Network with ARG 185 and ARG 350

3.3.5. TM1 Stalk Extension

Finally, the extension of the GCGR TM1 stalk into the extracellular domain has been denoted an important feature of GCGR structure, thought to have a role in the positioning of the extracellular domain in relation to the TM helical bundle (Hollenstein et al., 2013). This stalk extension is also evident in the PAC1R model (Figure 3-13). Although perfectly reasonable for this feature to be present in GCGR and PAC1R, as the latter was modeled from the former, the stalk extension may also be critical to PAC1

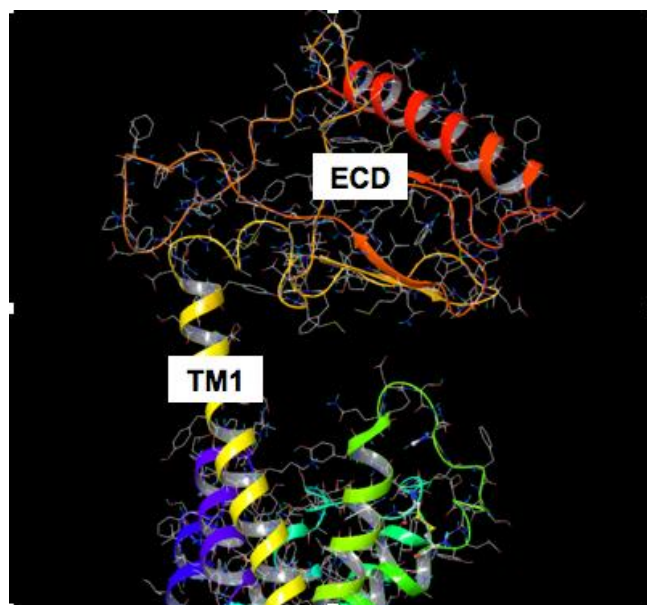**Figure 3-13: Extension of PAC1R TM1 into the ECD**

receptor stabilization. Unfortunately, the full extent and purpose of the PAC1R TM1 extension into the extracellular domain may not be fully understood until the receptor is experimentally resolved.

3.4. Molecular Docking of Hydrazide Antagonists to PAC1R

Ligand docking to the PAC1R homology model was accomplished using Autodock/Vina and PyMOL (Seelinger et al., 2010). The two hydrazides used for this docking study were the lead compounds identified by Beebe and coworkers as depicted in Figure 1-6 (Beebe et al., 2008), referred to herein as Hydrazide 1 (H1) and Hydrazide 2 (H2). Because there is no experimentally solved structure for PAC1R, and therefore the binding site is completely unknown, this research constituted a "blind" docking.

3.4.1. Ligand Preparation

AutoDock Tools (ADT), the free graphical user interface developed for use with AutoDock programs, was used to prepare the ligands and receptor for docking (Morris et al., 2009). Ligands H1 and H2 were prepared using the "Ligand" tool in ADT. The desired flexibility of the ligand is set by choosing molecule torsions. Based on atom and bond types, ADT automatically reads the number of torsions occurring in the ligand. Although bonds which appear red, as shown in Figure 3-14, are completely unrotatable, green bonds (rotatable) and magenta bonds (non-rotatable) can be adjusted to set desired torsions in the ligand. Initially, both ligand torsions were set to provide the most flexible molecule for docking. H1, therefore, had 9 rotatable bonds, while H2 had 5 rotatable bonds. Ligand files are then written as .pdbqt files, which store atomic coordinates, partial charges and AutoDock atom types for use in AutoDock Vina.
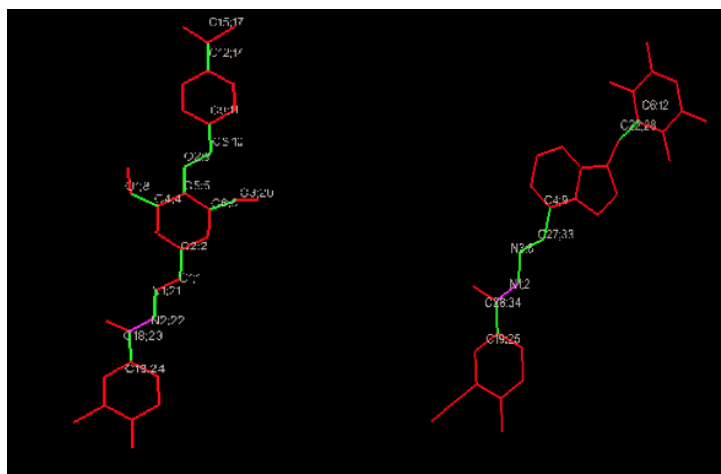

Figure 3-14: H1-Flex (Left) and H2-Flex (Right) Torsions

49

Although initial dockings were performed with the most flexible versions of H1 and H2 possible, most of the initial docking results (to be discussed momentarily) produced what appeared to be impossibly twisted molecules fraught with steric inconsistencies. Furthermore, initial binding affinities reported by AutoDock Vina were significantly higher than experimental expectations. To remedy this twisting, and in attempt to lower the binding affinity by developing as many non-bonding interactions as possible between ligand and binding site residues, torsions were slightly adjusted on each ligand to form more rigid molecules. These torsion adjustments are illustrated in Figure 3-15. H1_Flex was adjusted to a torsion of 8 rotatable bonds to form H1_Rigid and H2_Flex was adjusted to 4 rotatable bonds as H2_Rig.

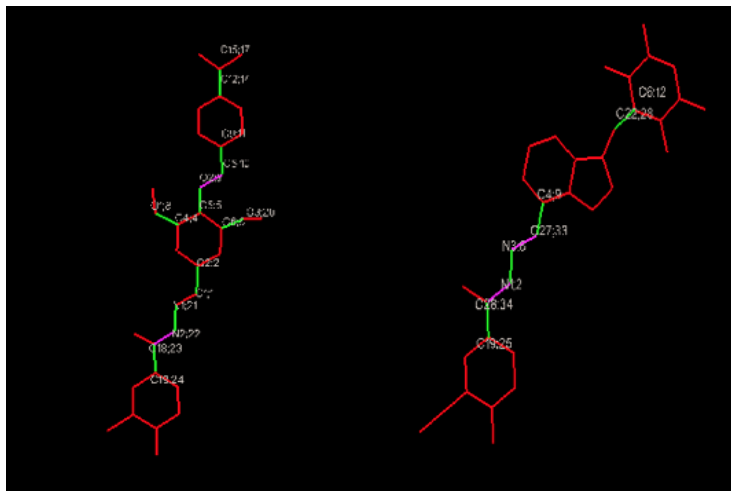

Figure 3-15: H1-Rigid (Left) and H2-Rigid (Right)

3.4.2. Receptor Preparation

To prepare the Swiss-Model-generated PAC1 receptor for docking studies, all water molecules were removed while charges and polar hydrogens were added via ADT

(Figure 3-16). The .pdb file was then written as a .pdbqt file. Because initial docking

results did not correlate to the experimental binding affinities for this ligand-receptor

complex, receptor torsions were revisited. ADT allows for the selective rendering of

residues as flexible. After initial docking studies were complete, 2 additional versions of

the PAC1R were created with flexible residues. The first contained as flexible only 5

residues which had demonstrated non-bonding interactions between the rigid PAC1R

(PAC1R_Rig) and both the rigid and flexible renderings of H1 and H2.

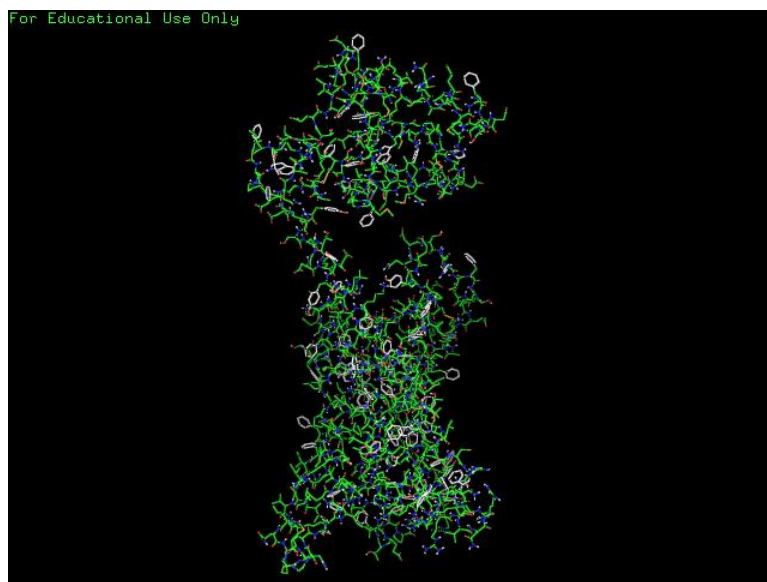**Figure 3-16: Docking-Prepared PAC1 Receptor**

This receptor was denoted as PAC1R_Flex5. To account for any non-bonding

interactions that may have been excluded by the initially rigid binding pocket, 10 residues

were selected as flexible in the second torsion-adjusted version of PAC1R, including the

5 residues with demonstrated non-bonding interactions in early trials plus an additional 5

residues in the immediate vicinity of the non-bonding 5. This receptor was denoted as PAC1R_Flex10.

3.4.3. Binding Site Determination

Siu et al. reported an experimental binding affinity of K_i equal to 56 nM for H1 and 73 nM for H2 (Siu et al). As AutoDock Vina reports binding affinity results in kcal/mol (ΔG), experimental results were converted to kcal/mol by solving for ΔG using the Gibbs formula for binding affinity:

$$K_i = \exp\ (\Delta G/RT)$$

where R = 1.986 cal/mol·K and T = 298 K (Klotz, 1997). In terms of ΔG, experimental results translate to -9.887 kcal/mol for H1 and -9.730 kcal/mol for H2. It was with these experimental values as a guide that binding site determination was undertaken. To hone in on a possible binding site for use in docking studies, the entire PAC1R was screened in two parts using the ADT grid generating feature with maximum box sizes. The ADT grid instructs AutoDock Vina where to focus ligand binding activity. These initial grids are illustrated in Figure 3-17. H1 was used as the ligand in the initial search for potential binding sites. Once the AutoDock Vina docking process is complete, possible bindings, or poses, are reported by the software with the calculated binding affinities in kcal/mol. Results significantly greater than the experimental values were discarded as unlikely indicators of potential binding sites. Bindings of the ligand to the outside/periphery of

52

the protein were also discarded as unlikely scenarios regardless of calculated binding affinities. Finally, it was generally expected that the acyl phenol portion of the hydrazide pharmacophore was a main player in binding pocket interactions. Therefore, poses which
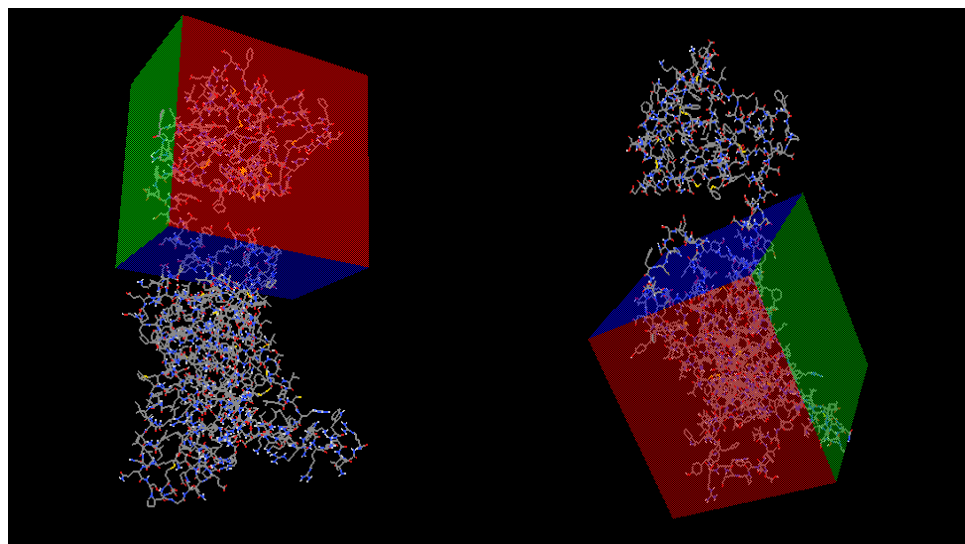
Figure 3-17: Binding Site Search Grid

placed the sterically cumbersome distal ring deep in the binding pocket were discarded as highly unlikely scenarios.

As discussed above, class B GPCRs are thought to bind peptides across two domains with the initial complex forming between the C-terminus of the peptide and the ECD of the receptor. This interaction facilitates the binding of the free N-terminus of the peptide to the juxtamembrane region of the 7TM domain of the receptor, resulting in a receptor-activating conformational change (Hoare, 2005). Although one initial binding of the H1 ligand to the PAC1R extracellular domain was somewhat consistent with this trend in binding to the receptor ECD, it is unlikely that small molecule antagonists would

behave thusly or be capable of exacting any conformational changes in the receptor owing to a lack of interaction with the 7TM region. Therefore, potential small molecule binding pockets were supposed to exist within the 7TM region of the receptor.

Although adjustments to the flexibility of the ligands to more rigid versions had the unwanted effect of greatly increasing the energy of the binding system, subsequent docking attempts with more flexible PAC1R models yielded promising results in which the ligand did bind within the 7TM portion of the receptor exhibiting respectable binding affinities. A number of parameters were experimented with to optimize docking results. These parameters included adjusting the centering coordinates for the binding site search area as well as altering the three dimensional size of the search area "box." Repeated adjustments to the grid box based on binding affinities led to an optimal grid box determination, as depicted in Figure 3-18.
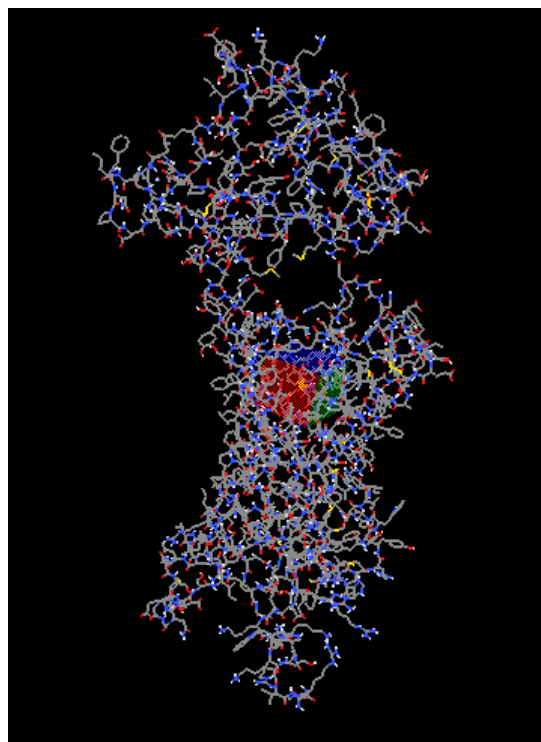
Figure 3-18: Final Binding Cite Determination

As mentioned above, initial binding results of flexible H1 and H2 ligands to the rigid PAC1R in the optimized grid area fell quite short of experimental results. The margin of error for AutoDock Vina-calculated binding affinities is given as ± 2.5 kcal/mol (Trott, 2009). Given this error range, docking study results of -12.387 to -7.4 kcal/mol and -12.23 to -7.0 kcal/mol for H1 and H2, respectively, potentially meet experimental values of -9.887 kcal/mol for H1 and -9.730 kcal/mol for H2. However, because the ligands in these early docking trials appeared impossibly bent and because it is preferred that docking results resemble experimental values as closely as possible, the ligand and receptor torsions were adjusted (as discussed above) to optimize non-bonding interactions.

3.4.4. Docking Results

Results from docking both flexible and rigid H1 and H2 ligands to variations of the PAC1 receptor (including the rigid receptor, PAC1R-Rigid, the receptor with 5 flexible residues, PAC1R-Flex5, and the receptor with 10 flexible residues, PAC1R-Flex10) are presented in Table 3-1. Results with H1-Rigid and H2-Rigid binding to PAC1R-Rigid were rather dismal with binding affinities rising to -1.8 kcal/mol for H1-Rigid and -2.6 kcal/mol for H2-Rigid. Generally speaking, rigid ligand results were notably higher in energy than results for their flexible counterparts. Furthermore, rigid receptor dockings were generally higher in energy than dockings which employed semi-flexible versions of the PAC1 receptor.

In attempt to maximize non-bonding interactions and lower binding affinities, and as discussed above, residues involved in non-bonding interactions with ligands in docking configurations 1 through 4 were used to create a semi-flexible receptor, PAC1R-Flex5. These residues, depicted in Figure 3-19, include TYR 161, LYS 206, ASP 207, ASP 298, and GLU 385. Both the rigid and flexible versions of H1 and H2 were docked to PAC1R-Flex5. The binding affinities of the rigid ligands to the Flex5 receptor were quite similar to results for rigid ligand bindings to the rigid PAC1R, and still much higher than experimental results (see Entries 5 and 6 in Table 3-1). The dockings of H1-Flex and H2-Flex to PAC1R-Flex5 yielded results closer to experimental values, but still with notably higher energies. H1 binding affinity was -6.6 kcal/mol versus an

experimental -9.887 kcal/mol. The H2 docking results, -7.6 kcal/mol versus -9.730 kcal/mol experimental, were slightly lower than the H1 results.

Table 3-1: Docking Results with Varying Ligand and Receptor Flexibilities

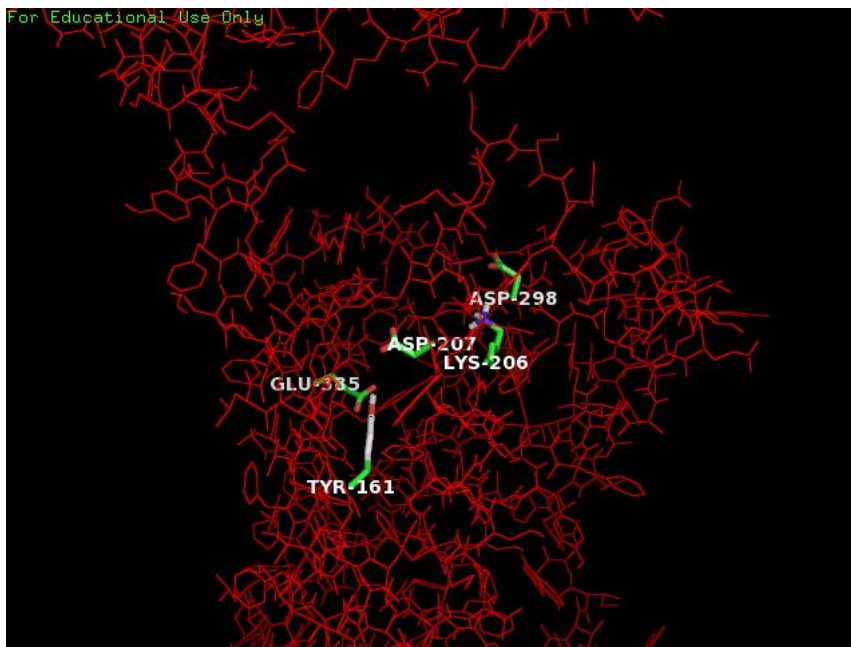| Entry	Ligand	Receptor	Binding Affinity (kcal/mol)
1	H1-Flex	PAC1R-Rigid	-5.9
2	H2-Flex	PAC1R-Rigid	-6.7
3	H1-Rigid	PAC1R-Rigid	-2.8
4	H2-Rigid	PAC1R-Rigid	-4.3
5	H1-Rigid	PAC1R-Flex5	-2.6
6	H2-Rigid	PAC1R-Flex5	-4.1
7	H1-Flex	PAC1R-Flex5	-6.6
8	H2-Flex	PAC1R-Flex5	-7.6
9	H1-Rigid	PAC1R-Flex10	-2.9
10	H2-Rigid	PAC1R-Flex10	-4.9
11	H1-Flex	PAC1R-Flex10	-9.9
12	H2-Flex	PAC1R-Flex10	-11.6

Figure 3-19: PAC1R with 5 Flexible Residues For Docking

Both the rigid and flexible versions of H1 and H2 were docked to a PAC1 receptor with expanded flexibility. As discussed above, an additional 5 residues in the immediate vicinity of the PAC1R-Flex5 residues were added to account for any non-bonding interactions that may have been excluded by the initially rigid binding pocket. These residues, depicted in Figure 3-20, include TYR 161, ARG 199, LYS 206, ASP 207, ASN 240, TYR 241, ASP 298, HIS 365, GLU 385, and GLN 392. Both the rigid and flexible versions of H1 and H2 were docked to PAC1R_Flex10.
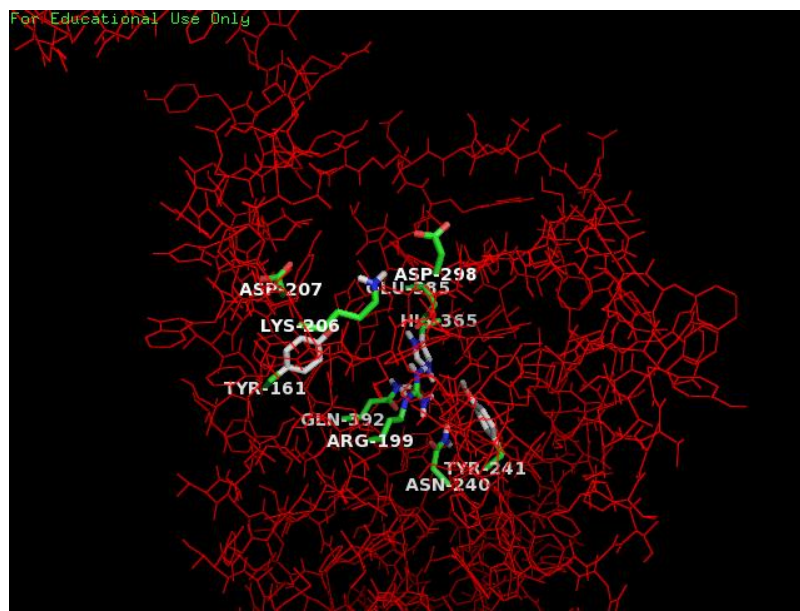
Figure 3-20: PAC1R with 10 Flexible Residues For Docking

As expected, because protein residues are not strictly rigid but naturally flexible and dynamic, the binding affinities for both H1/H2-Rigid and H1/H2-Flex were lower than PAC1R-Rigid and PAC1R-Flex5 docking results. The addition of 5 more flexible residues to form PAC1R-Flex10 predictably lowered binding affinity energy results as well. Interestingly, although H1-Flex was used for optimal docking to PAC1R-Flex10, the ligand does not appear impossibly bent as occurred in the dockings of H1-Flex to PAC1R-Rigid which had prompted the development of H1-Rigid. The same lack of bizarre ligand bending was noted in the optimal H2-Flex docking to PAC1R-Flex10 as well. It is likely that unusual conformations of the ligand were not due to the flexibility of the ligand, but rather a product of the inflexibility of the receptor. The use of both flexible ligands and a flexible receptor to achieve optimal docking study results reinforces the reality of the natural fluidity of both hydrazide compounds and the PAC1R.

However, due to software (and possibly hardware) limitations, full flexibility in the critical binding areas of the receptor was not possible. The general docking appearances of H1-Flex and H2-Flex are presented in Figure 3-21.
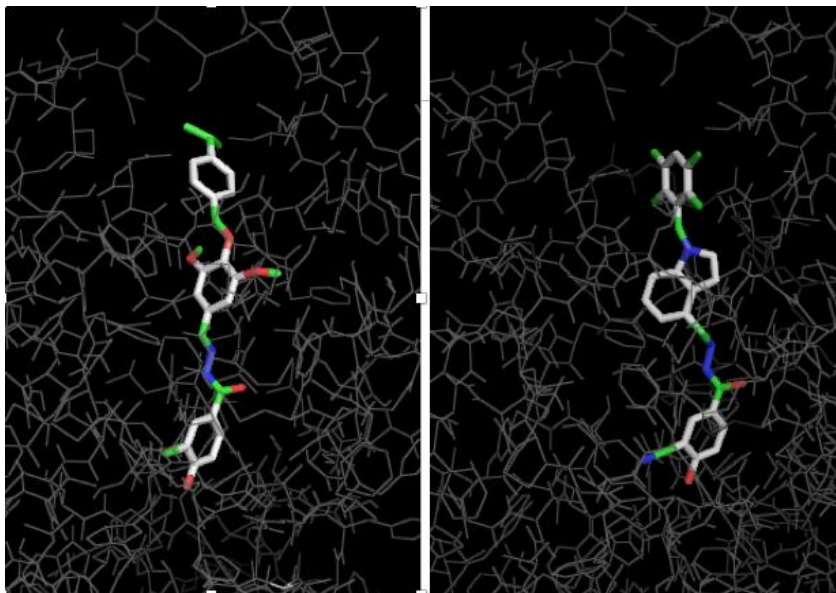

Figure 3-21: Optimized Dockings of H1-Flex (Right) and H2-Flex (Left)

H1 binding data closely mimicked the lower energy of the Beebe et al. experimental data (-9.9 kcal/mol compared to the expected value of -9.887 kcal/mol). H2 binding affinity results actually exceeded experimental values in recording binding affinities as low as -11.6 kcal/mol (see Table 3, Entry 12). Again, this may be due to the AutoDock Vina error margin in calculating binding affinities or it is also possible that the flexibility in the H2-Flex/PAC1R-Flex10 system exceeds the ligand/receptor flexibility occurring in nature, thereby artificially lowering system energy. PAC1R model

limitations and/or errors can also account for inaccuracies in binding affinity determinations.

3.4.5. Hydrogen Bonding

Weak intermolecular interactions play an important role in stabilizing a ligand energetically at the interface of a protein structure. Following the dockings of H1/H2-Rigid and H1/H2-Flex to both PAC1R-Flex5 and PAC1R-Flex10, non-bonding interactions were noted and a short list of possible binding pocket residues was devised based on these interactions. Three residues were found to appear repeatedly in the docking results for both rigid and flexible H1 and H2 ligands. These residues, depicted in Figure 3-22, form hydrogen bonds with the acyl phenol on H1 and include ASN 240, TYR 241 and HIST 365. No non-bonding interactions with surrounding PAC1R residues were found by the PYMOL software to engage the electron withdrawing chlorine group in the meta position on the H1 acyl phenol. However, it was found that the HIS 365 and TYR 241 residue orientations appear to be stabilized by hydrogen bonding to each other. Furthermore, the TYR 241 residue may achieve increased stability via hydrogen bonding to the backbone of PHE 362.
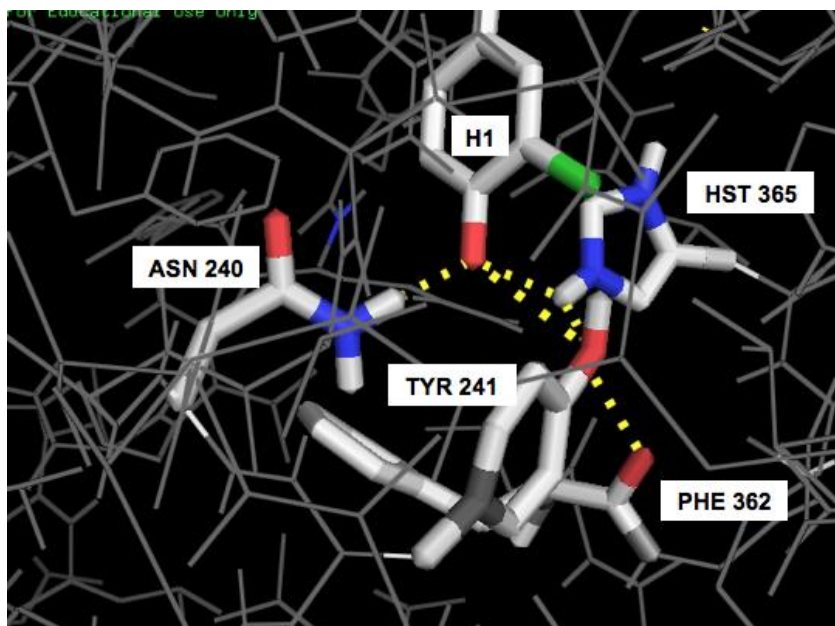
Figure 3-22: H1 Binding with 4 Interacting Residues

Interestingly, the H2 docking experiments also indicated ASN 240, TYR 241 and HIS 365 as key residues to the docking of Beebe's Hydrazide 2 to the PAC1R. However, 2 additional residues also appeared repeatedly in docking exercises. These residues included TYR 161 and GLN 392. Illustrated in Figure 3-23, while the acyl phenol on H2 also hydrogen bonded to ASN 240 and HIS 365. TYR 241 hydrogen bonds to the electron withdrawing nitrile group in the meta position. Additional hydrogen bonding occurs between the acyl phenol and GLN 392. Although not hydrogen bonded to the acyl phenol, TYR 241 engages HIS 365 in what may be orientation stabilizing hydrogen bonding with each other as occurred in the H1 docking depicted in
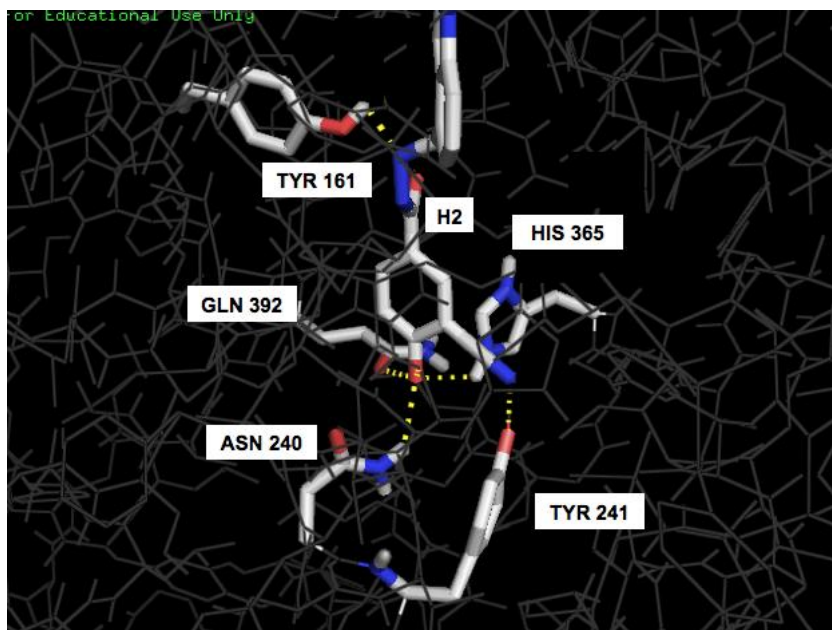
Figure 3-23: H2 Binding with 5 Interacting Residues

Figure 3-22 between TYR 241 and PHE 362. Although no non-bonding activity was detected between surrounding PAC1R residues and the oxygen atom located in the beta position to the acyl phenol on H1, this same oxygen on H2 did demonstrate hydrogen bonding interactions with the TYR 161 residue. Though only 3 residues interact directly with both H1 and H2 (ASN 240, TYR 241 and HIS 365), it is necessary to consider the role/importance of PHE 362, GLN 392, and TYR 161 in further investigations of this ligand-receptor complex. Proposed key residues to both H1 and H2 binding within the context of the 7TM domain are presented in Figure 3-24.
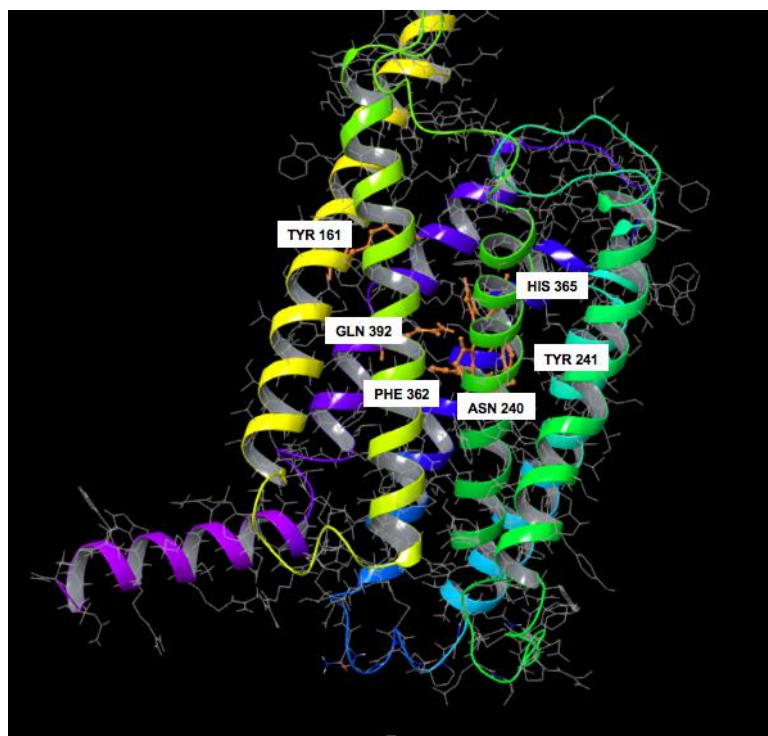
Figure 3-24: Proposed Key Binding Residues

While quite instrumental in determining a potential binding site in PAC1R, available software for this study was unfortunately limited by its inability to locate non-bonding interactions beyond hydrogen bonding. The inflexible nature of most of the receptor residues rendered quite difficult the exploration of docking result models for hydrophobic effects, π-effects and halogen bonding. However, it is possible to suggest some possibilities for these other types of non-bonding interactions based on the nature of the residues present and the characteristics of the ligand.

3.4.6. Hydrophobic and π- π Interactions

Although non-bonding interactions in addition to hydrogen bonding were not identified by the docking software, and the rigidity of the majority of the binding pocket residues rendered key interactions difficult to visually discern, the presence of particular residues in the binding pocket are suggestive of hydrophobic interactions that may take place between the receptor and ligand which would stabilize the binding of the ligand in the pocket by contributing to an energetically favorable complex. The interactions between ligands and the hydrophobic side chains of proteins contribute significantly to the binding free energy. Hydrophobic residues mutually repel water and other polar groups, resulting in a net attraction of the non-polar groups of the ligand. In addition, the apolar and aromatic rings of tryptophan, phenylalanine, and tyrosine participate in π-"stacking" interactions with aromatic moieties of the ligand. For interactions between two aromatic systems, two geometries are predominant: one, where two rings are parallel to each other, and, two, with a perpendicular, edge-to-face arrangement (Bissantz et al., 2010).

While depicting the already identified hydrogen-bonding interactions discussed herein (specified by dotted blue lines), Figure 3-25 also indicates possible hydrophobic/π-stacking interaction, between the general Beebe pharmacophore and binding pocket
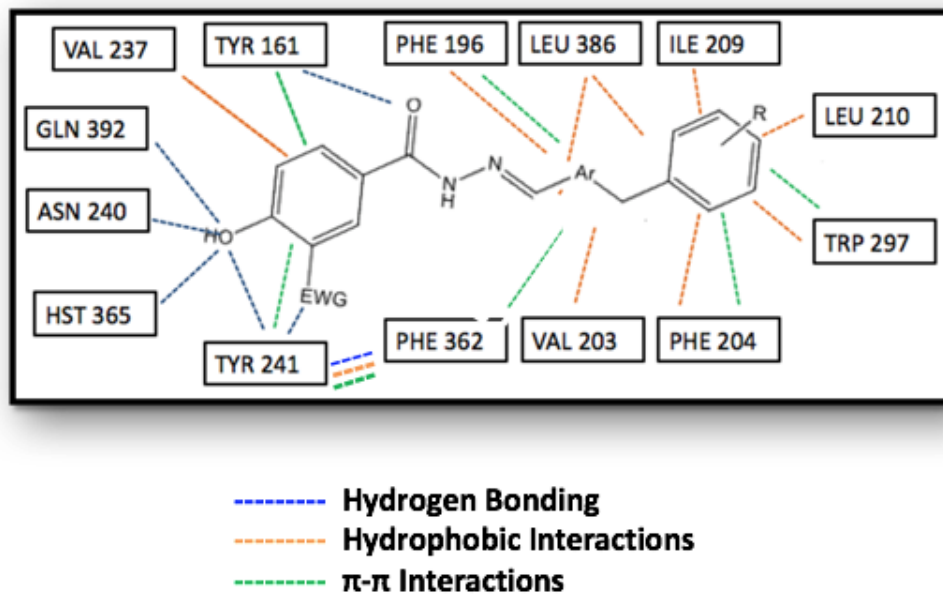
Figure 3-25: Possible Non-Bonding Binding Pocket Interactions

residues (specified by dotted orange lines). Due to their proximity to the bound ligand,

the hydrophobic residues PHE 196, VAL 203, PHE 204, ILE 209, LEU 210, VAL 237,

TRP 297 and PHE 362 are likely candidates for these energy-minimizing interactions.

The aromatic phenylalanine and tryptophan residues may participate in π-π interactions

as well (represented by green dotted lines). Additionally, the aromatic rings of TYR 161

and TYR 241 may interact with the pharmacophore's acyl phenol ring to stabilize ligand

docking in that region of the pocket.

3.4.7. Halogen Bonding

Halogen bonding refers to short contacts, previously referred to as charge-

transfer bonds, between a polarizable halogen (Lewis acid) and an oxygen, nitrogen or

sulfur atom (Lewis bases) in which negative charge is transferred from the base to the

acid (Metrangolo et al., 2001). Furthermore, electrophilic halogen atoms can interact with the nucleophilic oxygen and carbon atoms of C=O groups. Backbone nitrogen atoms are also thought to participate in halogen bonding. These interactions are known to contribute to ligand-protein stabilization (Fischer et al., 2008). Halogens, with the exception of fluorine, have unique electronic properties when bound to aryl or electron withdrawing alkyl groups. They show an anisotropy of electron density distribution with a positive area (called a "σ-hole") of electrostatic potential opposite the carbon-halogen bond (Clark et al., 2007). Essentially, a patch of negative charge is formed around the central region of the bond between the carbon and halogen atom, leaving the outermost region positive (hence the "hole"). This "hole" facilitates halogen bonding.

Beebe's Hydrazide 1 lead compound posses a chlorine atom meta to the hydrazide linker on the aryl phenol. Although docking software does not identify
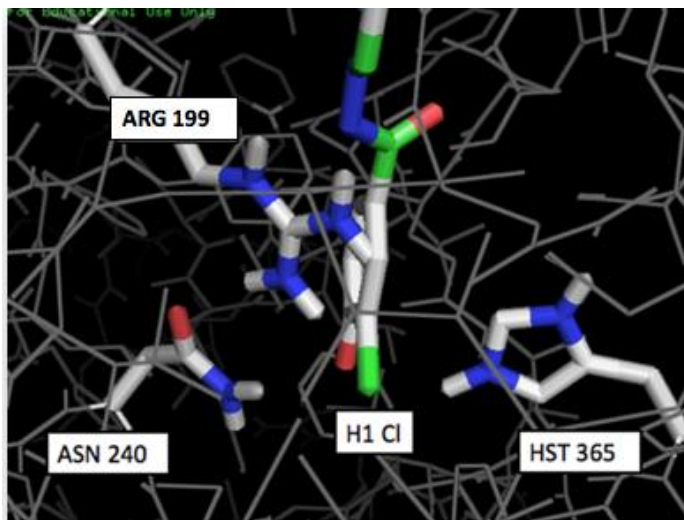

Figure 3-26: Potential Halogen Bonding Sites

potential halogen bonding interactions, some possibilities are discernible from docking models. Interestingly, ARG 199 is located in close proximity to the chlorine atom on H1. Arginine, a known Lewis base with its side chain guanidino group, could possibly form a halogen bonding interaction with the Lewis acidic chlorine atom. Furthermore, the nitrogen atoms of the imidazole ring of HIS 365, another binding pocket residue situated close the H1 chlorine atom, make it a good Lewis base and possible source of halogen bonding as well. ASN 240, also a Lewis base, constitutes a third possible site for halogen bonding. Figure 3-26 illustrates these three possibilities: ARG 199 (although this residue is situated awkwardly close to the acyl phenol in the figure), ASN 240 and HST 365. Of course, carbonyls and amides in neighboring residue backbones present further possibilities for halogen bonding. Although suggestions have been provided herein for hydrophobic, π-π , and halogen bonding interactions, these other forms of non-covalent interactions critical to small-molecule design and drug-to-receptor docking constitute grounds for further investigation into the binding pocket of PAC1R.

4.0. CONCLUSIONS

The homology model and ligand docking results herein presented are but one possible solution to the mystery of PAC1R architecture and binding. Though some promising leads have been established and possible key residues to this process identified, no definitive answers can be provided due to the theoretical nature of this undertaking. This data is heavily dependent on reputable software with obvious and discussed limitations. Although the reliability of homology and docking methods is not nearly as remarkable as empirical experimental data, these methods can provide new suggestions for protein-ligand interactions that might be otherwise overlooked.

In order to fully understand the mechanisms of ligand binding and activation of Class B GPCRs in general and PAC1R in specific, experimentally resolved structures of these receptors in both the active and inactive conformational states will be needed. Unfortunately, to date only two Class B GPCRs have been experimentally determined, corticotropin releasing factor-1 receptor (CRF1R) and the glucagon receptor (GCGR). CRF1R was crystallized in complex with an antagonist, enabling a thorough study of the CRF1 receptor small molecule binding site. GCGR was not crystallized with an antagonist ligand present in the binding pocket, thus hampering further elucidation of GCGR binding activities.

In lieu of available crystal structures or clearly defined binding pockets for the remainder of the Class B GPCRs, the challenge to researchers forward is to conduct thorough structural comparisons between the two known Class B receptors and

theoretical Class B receptors in order to feasibly illuminate the features and locations of small molecule binding sites and inform the treatment of illnesses associated with these receptors' activities.

5.0. BIBLIOGRAPHY

Alberts B, Johnson A, Lewis J, Ra M, Roberts K, Walter P. *Molecular Biology of the Cell.* 5th ed. New York: Garland Science; **2008**.

Arimura A. *Regul Pept*. **1991**, 37**,** 287–303.

Arnold K., Bordoli L., Kopp J., and Schwede T. *Bioinformatics*. **2006**, 22, 195-201.

Beebe X.; Darczak D.; Davis-Taber R.; Uchic M.; Scott, V.; Jarvis, M.; Stewart, A. *Bioorganic & Medicinal Chemistry Letters.* **2008**, 18, 2162-2166.

Benkert, P.; Tosatto, S.C.E.; Schomburg, D. *Proteins: Structure, Function, and Bioinformatics*. **2008**, 7, 261-277.

Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T.G.; Bertoni, M.; Bordoli, L.; Schwede, T. *Nucleic Acids Research.* **2014**, 42, 252-258

Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem*. **2010**, 53, 5061-5084.

Bortolato, A.; Dore, A.S.; Hollenstein, K.; Tehan, B.G.; Mason, J.S.; Marshall, F.H. *Brit. Journ. Of Pharmacology*. **2014**, 171, 3132-3145.

Chartrel N, Tonon MC, Vaudry H, and Conlon JM (1991) *Endocrinology.* **1991**, 129, 3367–3371.

Choi D., Furay A., Evanson N., Ostrander M., Ulrich-Lai Y., Herman J. *J Neurosci.* **2007**, 27, 2025–34.

Chothia, C; Lesk, A.M. *EMBO.* 1986, 5, 823–826.

Clark, T.; Hennemann, M.; Murray, J.S.; Politzer, P. *J. Mol. Model*. **2007**, 13, 291–296.

de Beer, T.A.P.; Berka, K.; Thornton, J.M.; Laskowski, R.A. *Nucleic Acids Res*. **2014**, 42, D292-D296.

Engelmann, M.; Landgraf, R.; Wotjak, C. *Front Neuroendocrinol.* **2004**, 25, 132–49.

Eyrich, V.A.; Marti-Renom, M.A.; Przybylski, D.; Madhusudhan, M.S.; Fiser, A.; Pazos, F.; Valencia, A.; Sali, A.; Rost, B. *Bioinforamtics.* **2001**, 17, 1242-1243.

Finkelstein, Y.; Loffler., B; Rabey, J.; Gilad, G. *Brain Research*. **1985**, 343, 314-319.

Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schioth, H. B. *Molec. Pharm.* **2003**, 63, 1256-1272.

Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. *Nucleic Acids Research.* **2010**, 38 695-699.

Guex, N.; Peitsch, M.C.; Schwede, T. *Electrophoresis*, **2009**, 30, 162-173.

Haas, J.; Roth, S.; Arnold, K.; Kiefer, F.; Schmidt, T.; Bordoli, L.; Schwede, T. *Database.* **2013**, 31, 156-162.

Hammack, S.E.; Cheung, J.; Rhodes, K.M.; Schutz, K.C.; Falls, W.A.; Braas, K.M.; May, V. *Psychoneuroendocrinology.* **2009**, 34, 833–843.

Hammack, S.; Roman, C.; Lezak, K.; Kocho-Shellengerg, M.; Grimmig, B.; Falls, W.; Braas, K.; May, V. *J. Mol. Neurosci.* **2010**, 42, 327-340.

Hoare, S. *DDT.* **2005**, 10, 417-427

Hollenstein, K.; Kean, J.; Bortolato, A.; Cheng, R.K.Y.; Dore, A.S.; Jazayeri, A.; Cooke, R.M.; Weir, M.; Marshall, F.H. *Nature.* **2013**, 499, 438-443.

Ishihara, T., Nakamura, S, Kaziro, Y., Takahashi, T., Takahashi, K., and Nagata, S. *EMBO J.,* **1991**, 10**,** 1635–1641.

Jacoby, E. *ChemMedChem.* **2006,** 8, 761-82

Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. *Nucleic Acids Research.* **2009**, 37, 387-392.

Klotz, I.M.. *Ligand-Receptor Energetics: A Guide for the Perplexed. **Wiley.** 1997.*

Köves*, K.;* Arimura*, A.;* Görcs*, T.G.;* Somogyvári-Vigh*, A. Neuroendocrinology* **1991**, *54*, *159–169.*

Krieger, E.; Nabuurs, S.B.; Vriend, G. *Structural Bioninformatics* 1st ed., New York: Wiley-Liss, Inc.; **2003**.

Kuntal, P.; Melcher K.; Xu, H.E. *Acta Pharmacologica Sinica.* 2012, 33, 300–311.

McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y.M.; Buso, N.; Cowley, A.P.; Lopez, R. *Nucleic Acids Research.* **2013**, 41, 597-600.

Metrangolo, P. & Resnati, G. *Chem. Eur. J.* **2001**, 7 , 2511–2519.

Miyata A.; Sato, K.; Hino, J.; Tamakawa, H.; Matsuo, H.; Kangawa, K. *Ann N Y Acad Sci.* **1989**, 865**,** 73–81.

Morris, A.L.; MacArthur, M.W.; Hutchinson, E.G.; Thornton, J.M. *Proteins*. **1992**, 12, 345-364.

Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. *J. Computational Chemistry.* **2009**, 16, 2785-91.

Nussdorfer, G.G.; Malendowicz, L.K. *Peptides*. **1998**, 19, 1443-67.

Ogim, K.; Miyamoto, Y.; Masuda, Y.; Habata, Y.; Hosoya, M.; Ohtaki, T.; Masuo, Y.; Onda, H,; Fujino, M. *Biochem. Biophys. Res. Commun.*, **1993**, 196, 1511-1521.

Olejniczak, E.T. *PNAS*. **2007**, 104, 7875-7880.

PyMOL Molecular Graphics System, Version 1.8 Schrodinger, LLC

Schrödinger Release 2015-1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.

Seelinger, D.; de Groot, B.L.; *J. Comput Aided Mol Des.* **2010**, 24, 417-422.

Sievers, F.; Wilm, A.; Dineen, D.G.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J.D.; Higgins, D.
Molecular Systems Biology. 2011, 7, 539-545.

Siu, F.Y.; de Graaf, C.; Han, G.W.; Yang, D.; Zhang, Z.; Zhou, C.; Xu, Q.; Wacker, D.; Joseph, J.S.; Liu, W.; Lau, J.; Cherezov, V.; Katritch, V.; Wang, M-W.; Stevens, R.C. *Nature*, **2013**, 499, 444–449.

Sousa, S.F.; Fernandes, P.A.; Ramos, M.J. *Proteins.* **2006**, 65, 15-26.

Stoffel M.; Espinosa R.; Trabb J.B.; Le Beau, M.M.; Bell, G.I. *Genomics*. **1994**, 23, 697-699.

UniProt Consortium. *Nucl. Acids Res*. **2015**, 43, D204-D212.

Venkatakrishnan, A.J.; Flock, T.; Prado, D.E.; Oates, M.E.; Gough, J.; Babu, M.M. *Current Op. in Struct. Bio*. **2014**, 27, 129-137.

Warne, T.; Moukhametzianov, R.; Baker, J.G.; Nehme, R.; Edwards, P.C.; Leslie, A.G.; Schertler, G.F.; Tate, C.G. *Nature*. **2011**, 469, 241-244.

Worth, C.L.; Kleinau, G.; Krausse, G. *Plos.* **2009**, 9, 1-15.

Worth, C.L.; Kleinau, G.; Krause, G. *Plos*. **2013**, 17, 62-68.

Yehuda, R. .*N Engl J Med.* **2002**, 346, 108-114.

Yehuda R.; Halligan, S.; Golier, J.; Grossman, R.; Bierer, L. *Psychoneuroendocrinology.* **2004**, 29, 389–404.

Yon, L. *Microsc Res Tech.* **1992**, 54, 137-157.