UNIVERSITY OF NAIROBI

THE 3-D STRUCTURE OF NEMATODE AF2 RECEPTOR AND PEPTIDE INTERACTIONS

BY:

Lynett K. Jones

I56/60282/2010

A THESIS SUBMITTED AS PARTIAL FULFILLMENT FOR THE AWARD OF A MASTER OF SCIENCE DEGREE IN ENVIRONMENTAL CHEMISTRY, UNIVERSITY OF NAIROBI.

JULY 2013

DECLARATION

I declare that this thesis, "**THE 3-D STRUCTURE OF NEMATODE AF2 RECEPTOR AND PEPTIDE INTERACTIONS,***"* is my own work that was performed under the supervision of Professor G.E. Jackson and Dr. D. N. Mbui, and that it has not been submitted before for a degree or examination at this or any other university.

.....

Lynett K. Jones

I56/ 60282/2010

Herewith our approval as supervisors

G E Jackson

....

Professor G.E. Jackson

Chemistry Department, University of Cape Town, SA

.....

Dr. Damaris N. Mbui

Chemistry Department, University of Nairobi

DEDICATION

This thesis is dedicated to my beloved husband Dr**. Joseph Ivivi** and my son **Gift Ivivi** for the endless encouragement, inspiration, unconditional love, guidance, and support that they have always given me. I love you all!

ACKNOWLEDGEMENT

I owe my gratitude to all the people who have made this thesis possible and who have made my graduate experience one that I will cherish forever.

I would like to dedicate my thanks to my supervisor **Prof. Graham E. Jackson** and his colleague **Prof. Kelly Chibale** for offering me the opportunity to do my research in their respective research groups at University of Cape Town, South Africa and also providing me with the funds. I would also like to express my deepest gratitude to **Dr. Damaris N. Mbui** not only for her valuable guidance, technical support and helpful suggestions during the entire work but also for the wonderful experience.

I would especially like to thank **Grace Mugumbate** and **Gadzikano Munyuki**, of University of Cape Town, for spending endless hours helping me in running my simulations.

Thank you all!

ABSTRACT

The AF2 receptor belongs to the super family of G - protein coupled receptor that are of great significance in the development of potent drugs for the treatment of parasitic nematodes affecting people, especially school going children in developing countries. An accurate evaluation of the peptide binding site into the receptor, at molecular level, may play a key role in the design of new molecules with desirable properties and reduced side effects. In this study, 3-D structure of AF2 receptor was developed using the X-ray crystal structure of beta 2- adrenergic and Adenosine A_{2A} receptors and MODELLER. The initial structure of the receptor was refined computationally with energy minimization and molecular dynamic simulations using GROMACS software. The AF2 peptide structure Lys-His-Glu-Tyr-Leu-Arg-Phe-NH_2 was generated using Insight II software and energy minimized using the force field available in Discover studio. NMR restrained molecular dynamics was used to determine the solution conformation of AF2. The resulting structure was docked into the optimized model and the critical amino acids responsible for binding were identified. Docking calculations of the AF2 peptide showed that the N-terminus, helix 7and the extracellular domains of the AF2 receptor defined its binding pocket. The receptor-ligand complexes were stabilized by formation of hydrogen bonds and hydrophobic interactions. Virtual screening studies were also performed on the receptor using ZINC compounds and they were found to dock inside the receptors binding side. Therefore, knowledge of the 3D structure and the binding pocket of the receptor, from parasitic nematodes, could lead to structure-based design of non- peptide antagonists that prevent binding of AF2 molecules. This would pave the way for the development of new drugs to control nematodes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abbreviations

GPCRs - G-protein coupled receptors

MD - Molecular dynamics

AF2R- Activated Functions 2 Receptor

β2AR - beta2-adrenergic receptor

7TMs - seven transmembrane helices

ECL - extracellular loop

ICL - intracellular loop

(EGF) - Epidermal growth factor

Amino Acid Residues

Name	3-letter code	1-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenyalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

1.0 INTRODUCTION

Cell signalling is vital for cell cooperation in the body. Cells signal to each other through compounds in the extracellular region, which the cells detect and respond to selectively. There are many endogenous signals, which include hormones, growth factors and cytokines and the cells responses may be changes in gene expression, cellular activity or ion channel activity. When the ligands are large or hydrophilic such that they cannot penetrate the lipid plasma membrane; proteins in the plasma membrane act as signal transducers by coupling the external signal to the biochemical responses of the cell. Hydrophobic ligands are capable of penetrating the lipid bilayer directly to act on the intracellular region. The cloning of genes expressing receptors and ion channels, together with pharmacological studies, has shown that the diversity among the target proteins is great. This molecular diversity of the target proteins raises the possibility of discovering drugs that act selectively on the different structures. Selectivity is very important from a pharmacological viewpoint, as it means that one target structure can be targeted without affecting other structures, and thereby be used as a tool in determining the function and distribution of the different isoforms of the proteins and later on, reducing the side effects of a drug by making the drug selective for only one isoform.

1.1 G Protein-Coupled Receptors (GPCRs)

1.1.1 GPCR structure and functions

G protein-coupled receptors (GPCRs) also known as seven-transmembrane (TM) domain receptors, 7TM receptors, heptahelical receptors, serpentine receptor, or G protein-linked receptors (GPLR) are a class of transmembrane proteins that constitute a large superfamily targeted by many drugs (Lagerström et al. 2008). Approximately 50% of the top 100 best-selling

drugs of most modern medicines act on GPCRs (ThiChamberlain, A.K. and Bowie J.U., 2004). This has led to growth in both pharmaceutical and academic research that focuses on drug discovery (Overington et al. 2006). Drug discovery is the process of discovering and designing drugs that includes target identification, target validation, lead identification, lead optimization and introduction of the new drugs to the public. GPCRs play a vital role in signal transduction and can be activated by a wide variety of ligands, including photons, amines, hormones, neurotransmitters and proteins. GPCRs are integral membrane proteins sharing a common global topology that consists of seven trans-membrane alpha helices (TM1-TM7) which are connected through six alternating extracellular and intracellular loops. The N- terminus is located on the extracellular side of the membrane, whereas the C- terminus occupies the intracellular side. The trans-membrane domain is constituted by seven alpha helices, which are known to adopt a common folding pattern, and thus, the 7TM domain represents the most conserved region among the GPCR family of proteins (Baldwin, 1994). **Figure 1.1** shows a schematic representation of the helical bundle of rhodopsin in the plasma membrane. Several highly conserved functional microdomains of amino acids are identified in the TM helices of class A GPCRs (Bissantz et al., 2004). They include: the N motif in TM1, the LAx(2)D motif in TM2,the D/ERY motif in TM3, W motif in TM4, F/Yx(2)Px(7)Y motif in TM5, F/Yx(3)WxPYY motif in TM6 and the NPx(2)Y motif in TM7 where x is any other amino acid residue (Kristiansen, 2004, Lundstrom, 2005). The extracellular loop 1 and 2 contain two highly conserved cysteine residues that form disulfide bonds to stabilize the receptor structure.

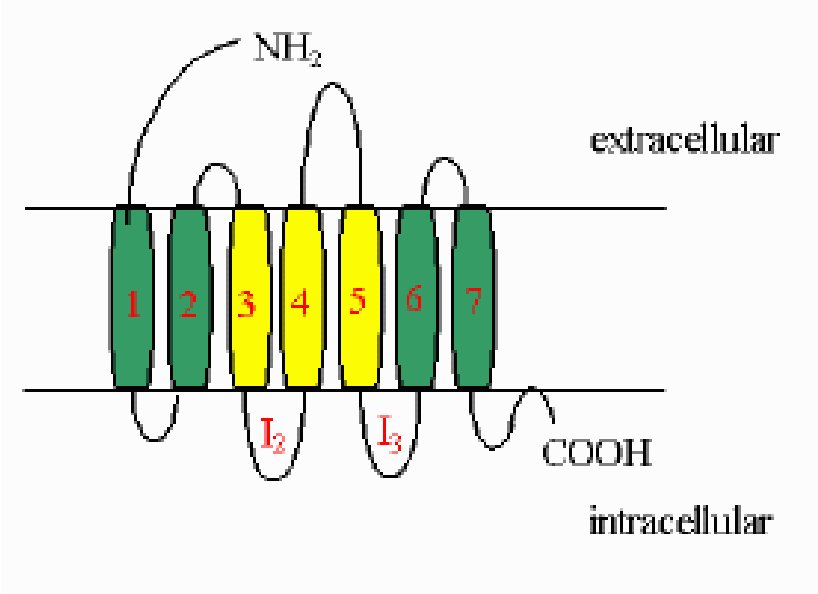


Figure 1.1: A schematic representation of the helical bundle of rhodopsin in membrane.

GPCRs are single polypeptide chains having seven hydrophobic transmembrane - spanning segments that couple in the presence of an activator to an intracellular effector molecule through a trimeric G protein complex (Svoboda et. al., 1999). G protein complex originates from the interaction of GPCRs with guanine nucleotides. GPCRs vary in sequence, length and function of the N-terminus domain, C-terminus domain and intracellular loops with each domain providing specific properties to the different receptor proteins. Of importance to note is that significant sequence homology is found within subfamilies. Researchers have also found out that the binding of agonistic ligands to the receptors elicits conformational changes of the receptor and activates the G protein. In this manner the receptors transfer extracellular signals to intracellular targets (Lomize et al, 1999).

1.2 Nematodes

Parasites from the phylum nematoda cause numerous diseases in humans, animals and plants placing major burdens on agricultural production and global health. Infections by these pathogens cause extensive suffering in humans and veterinary animals and major losses are incurred in agricultural production due to disease and the cost of implementing control programs (Jasmer et al., 2003). Current methods used to control or reduce the impact of nematode infections heavily rely on anthelmintics. The first drugs to be used were benzimidazoles (BZ), which were introduced in the early 1960s (Brown et al., 1961). This was followed by the introduction of imidothiazoles-tetrahydropyrimidines, avermectins and milbemycin (Chabala et al., 1980). Cyclodepsipeptide was commercially introduced to control nematodes in cats in 2006 and resistant parasites have started to appear a few years after the introduction hence the need for a better drug. Other methods for controlling control parasitic nematodes include pasture management, biological control and nutrient supplementation. These have been employed because of the anthelmintic efficacy whereby farmers monitor the faecal egg counts to determine the level of contamination. Other biological control methods include the use of predatory fungi that kill many nematode species hence reducing the intensity of infection (Larsen, 1999). One of the nematode killing fungi, *Dudding-tonia fragrans* has been used in New Zealand and has been found to have a very good trapping efficiency of approximately 78% and activity for up to 90 days on pastures, thus providing a promising alternative to reduce animal mortality from parasitic infections (Waghorn et al.,2003). On the other hand supplementing animals with dietary proteins, selenium and minerals can be applied to counter infections by enhancing host immunity (Islam et al., 2006; Koski and Scott, 2003). Although short-comings of chemical based methods are well recognized, the general approach has provided enormous benefits to human health and agricultural production and the use of anthelmintics is likely to remain a major factor in

integrated methods of parasite control. Deficiencies of current anthelmintics include: (1) the increasing and now widespread occurrence of nematode strains that have anthelmintic resistance (Kaminsky, 2003; Roos, 1997); (2) serious occupational exposure and environmental impacts presented by some anthelmintics (Risher et al., 1987; Schneider et al., 2003; Spratt, 1997); (3) the relatively poor efficacy of available anthelmintics against some nematode pathogens of humans (Stepek et al., 2006). The control of nematode parasites in plants and mammals is limited and thus it requires good hygiene and regular use of anthelmintics. Use of anthelmintics has led to resistance which may be attributed by the following: changes in drug translocation, receptor modification, changes in receptor numbers, receptor modification or post receptor modification (Sangster, 2001).

Hookworm

The major human soil transmitted helminthes (STH), *Ascaris lumbricoides*, *Trichuris trichiura* and the hookworms occur in 1221 million, 795 million, and 740 million people, respectively (De Silva et al., 2003), and are among the most common pathogens of humans in developing countries. There is considerable epidemiological overlap among the soil transmitted helminthes infections and Disability-Adjusted Life Year (DALY) estimates indicate that the disease burden is almost equivalent to better-known conditions such as malaria and tuberculosis (Bethony et al., 2006). Human hookworm infection is caused by the nematode parasites *Necator americanus* and *Ancylostoma duodenale* and infection is a leading cause of anaemia and protein malnutrition, afflicting an estimated 800 million people in the developing nations of the tropics. The largest numbers of cases occur in impoverished rural areas of sub-Saharan Africa, Latin America, South-East Asia and China. *Necator americanus* is the most common hookworm

worldwide, while *Ancylostoma duodenale* is more geographically restricted. *Necator americanus* is an obligate skin-penetrating hookworm (Hawdon et al, 1993), which resides in the duodenum (Pritchard *et al*, 1995). School-aged children (including adolescents) and preschool children tend to harbour the greatest numbers of intestinal worms and schistosomes and as a result experience growth stunting and diminished physical fitness as well as impaired memory and cognition. These adverse health consequences impair childhood educational performance; reduce school attendance and future wage-earning capacity. Hookworms are highly host specific; obligate blood-feeders causing gastro-intestinal bleeding, anaemia and iron deficiency in most chronic infections (Hotez and Pritchard, 1995; Albonico et al, 1998; Devaney, 2005). They have broad ranging effects on pregnancy, childhood growth, nutrition, and cognitive and intellectual development (Evans et al., 1995). Since hookworms do not replicate in the human host, the intensity of *Necator americanus* infection is estimated using a quantitative faecal egg count as a practical marker for worm burden (Hotez et al, 2005). Hookworm infections are defined by the World Health Organization (WHO) to be moderate in individuals producing 2,000 to 3,999 eggs per gram of faeces (epg) while heavily infected individuals produce over 4,000 epg. At high intensity, hookworm infections can cause significant morbidity and mortality (Hotez and Pritchard, 1995). The life cycle of nematodes is shown in **figure 1.2**. However, there is increasing evidence suggesting that a tolerated *Necator americanus* infection (at ~ 50 epg), as described by Mortimer et al (2006), might be associated with therapeutic benefits to autoimmune conditions such as Crohn's disease (Reddy and Fried, 2007, 2009), allergic disorders including asthma (Pritchard et al, 1995; Quinnell et al, 2004; Falcone and Pritchard, 2005) and atopic dermatitis (Cooper et al, 2003; Wordemann et al, 2008; Flohr et al, 2009). Hookworms are transmitted when hookworm eggs excreted in human faeces are incubated in appropriate soil

conditions, hatch into larvae and enter a human host through the skin. Adult hookworms are blood feeders and attach to the mucosa of the small intestine and cause intestinal blood loss. Typically, symptoms can arise from infections with as little as 50 to 150 of adult parasites. Infection induces iron-deficiency anaemia, which is in direct correlation with the number of parasites (as measured by quantitative egg counts). In children, chronic hookworm infection impairs physical and intellectual development, reduces school performance and attendance, and adversely affects future productivity and wage-earning potential (Evans et al., 1995). Control is achieved by population-wide treatment with benzimidazole-based anthelmintics albendazole and the 2001 World Health Assembly advocated the anthelminthic treatment of 75% of all at-risk school-aged children by 2010. However, drug treatment does not prevent re-infection and the frequent use of drugs required to maintain control would lead to drug resistance. Therefore, there is a drive to develop a safe and cost-effective vaccine.

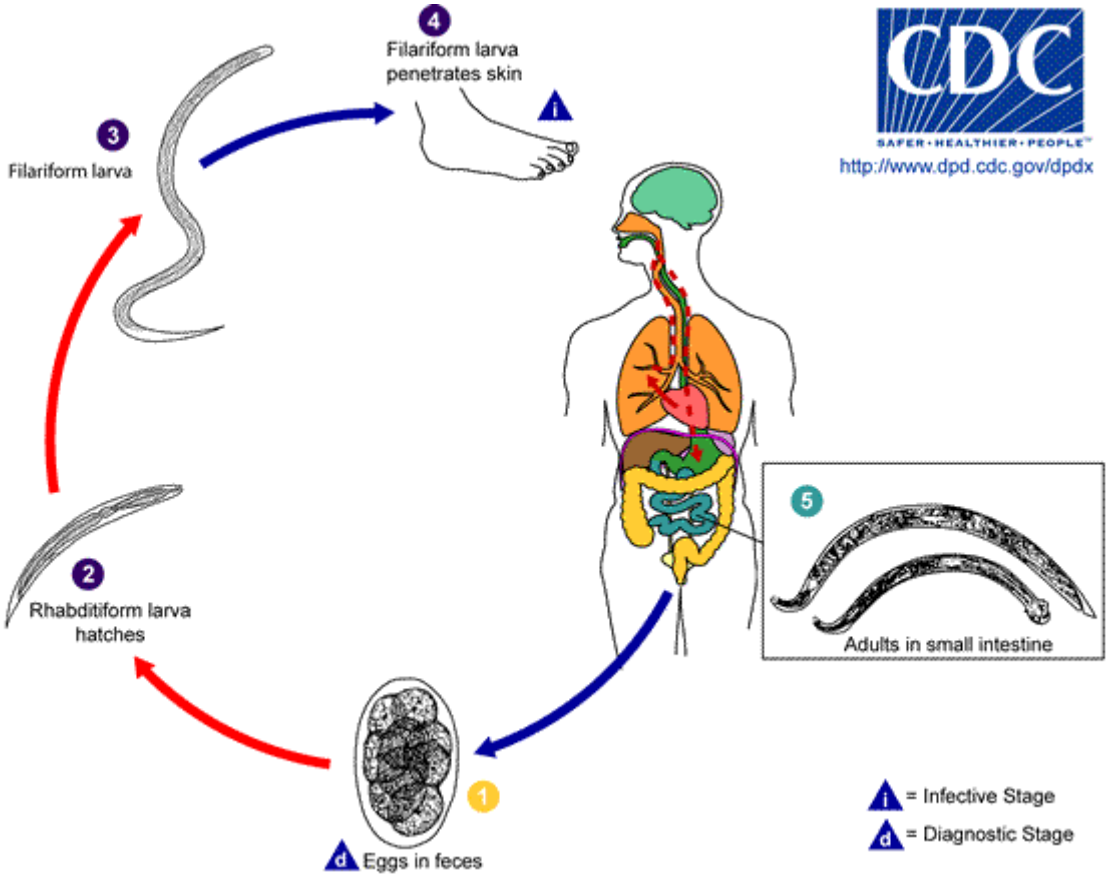


Figure 1.2: Life cycle of nematodes. Eggs are passed in the stool. Larvae hatch in 1 to 2 days, and grow in the faeces and or the soil, to become infective (third-stage) larvae within 5 to 10 days. On contact with the human host, the larvae penetrate the skin and are carried through the blood vessels to the heart and then to the lungs. They penetrate into the pulmonary alveoli, ascend the bronchial tree to the pharynx, and are swallowed. Once they reach the small intestine, larvae reside and mature into adults. Adult worms attach to the intestinal wall and feed on intestinal mucosa and blood. Female worms then start laying eggs. Source: http://www.dpd.cdc.gov/dpdx/html/hookworm.htm.

1.3 Overview and Significance

Parasitic nematodes have shown a dramatic impact on many aspects of the human condition. For example, lymphatic filariasis is a major health problem, striking people of all ages, inflicting both economic and social burdens on many tropical and subtropical countries (Anosike et al. 2005). *Wuchereria bancrofti*, *Brugia timori* and *B. malayi* infect 150 million people in 73 different countries, with 1 billion living in endemic areas at risk of infection (Williams, Lizotte-Waniewski et al. 2000; Hoerauf 2006). The visual manifestations of the infection can be severe leading to profound physical and psychological disabilities for the individual and for society (Wamae, 1994). Nematode infections are a major cause of human morbidity and contribute to a significant increase in Disability Adjusted Life Years (DALY, a summary measure that combines the impact of illness, disability and mortality on population health) with lymphatic filariasis having a 5.55 million DALY burden (Molyneux, 2003). Effective chemotherapy is needed to control and potentially eliminate filarial infections, especially lymphatic filariasis caused by *W. bancrofti*, and river blindness caused by *Onchocerca volvulus*.

However, nematodes exhibit enormous diversity and significant physiological, biochemical and molecular differences between nematode species have been clearly demonstrated (Gomez-Escobar, Lewis et al. 1998; Murray, Manoury, et al. 2005). During a typical career in medicinal chemistry, most chemists will run across a G-protein coupled receptor (GPCR) as a drug target. Almost invariably, the interest in the receptor will lead to the construction of a three-dimensional model as an aid in interpreting ligand binding and molecular biological data. The quality and usefulness of this model will depend on the assumptions made in its construction and by its supporting biophysical and molecular biological data. These models

are utilized for three functions: 1) to visualize the protein interior and to propose modes of ligand binding, 2) to plan mutagenesis experiments, and 3) to support ligand design. This perspective will critically assess the recent trends in model construction, the reliability of models themselves, and their impact on the drug discovery process. GPCRs are important both in understanding and in treating diseases. Because of their integral role in cellular signalling, GPCR dysfunction can lead to illness. Reversal of these aberrant effects can often contain, if not cure, many forms of disease. The GPCRs are therefore the primary target for a great number of drugs. Recent estimates suggest that up to 60% of the modern pharmacopoeia is targeted to GPCRs. Information and models that help us understand how these molecules interact with the receptor can therefore be of great practical interest to the medicinal chemist.

1.4 Problem Statement and Research Motivation

One-sixth of the world's population suffer from Neglected Tropical Diseases caused by parasitic nematodes, which are comparable to HIV/AIDS, tuberculosis and malaria. These diseases cause disability and death, mainly in developing countries. Parasitic nematodes cause elephantiasis, physical and mental retardation, and anaemia to school-going children. This is attributed to resistance of nematodes to most used drugs. Though, the nematode genome has more than 1500 predicted GPCRs which participate in a tremendous diversity of functions by playing a key role in mediation extracellular signalling, there is inadequate research and ground-breaking investigations done to develop new and innovative approaches to battling and eradicating diseases caused by these parasitic nematodes. Therefore, insight into peptide-receptor interaction is of great importance for designing new ligands with therapeutic potential. In order to study these interactions, three-dimensional structural information about the receptor is of great

importance. By determining the 3-D structure of the nematode GPCRs and identifying the binding sites, the receptors can be used in designing species-specific anthelminthes and this will go a long way in addressing poverty in the developing countries. This thesis focuses on the generation of the 3-D structure and the determination of the binding site of the parasitic nematode.

1.5 Aims and Objectives

1.5.1 Overall Objective

To determine the 3D structure of the nematode AF2 receptor, and characterize the binding pocket of the receptor using the AF2 peptide.

1.5.2 Specific Objectives

1. To use homology modelling to construct a 3-D model of the AF2 GPCR receptor of the parasitic nematode
2. To perform molecular dynamic studies on the AF2R in vacuo
3. To perform conformational searches of AF2 peptide in water using NMR distance restrained molecular dynamics
4. To use the optimized solution structure of AF2 to identify its binding site on AF2R.
5. To perform docking calculation of the AF2 peptide and the AF2R
6. To carry out Virtual Screening studies on the AF2 receptor.

2.0 LITERATURE REVIEW

2.1 G Protein-Coupled Receptors (GPCRs)

G protein-coupled receptors (GPCRs) form a super family of transmembrane (TM) embedded receptors that convey an extracellular signal exerted by a hormone or neurotransmitter to an intracellular response through G proteins. The G proteins are composed of α, β and γ subunits. These subunits dissociate from each other and separately activate several classical effectors, including adenylyl cyclases and phopholipases. They also regulate the activity of ion transporters, several kinases and ion channels. G-protein-coupled receptors (GPCRs) are the most important classes of transmembrane (TM) proteins involved in cell communication processes and in mediating such senses as vision, smell, taste, secretion, neurotransmission, metabolism and pain. The signals that activate these proteins are usually chemical in nature, for example for the opsin family, the signal is "visible" light (electromagnetic radiation).They all have a similar structure, with an extracellular N-terminus, 7 hydrophobic transmembrane α-helical regions connected by three extracellular loops (ECL1–3) and three intracellular loops (ICL1-3), and an intracellular C-terminus as shown in **Figure 1.1** above. The GPCR super-family consists of five main classes, of which the class A rhodopsin-like GPCRs form the largest subfamily (Fredriksson, R. et al. 2003). Next to the N- and C-termini, the extracellular loops of GPCRs are the most variable structural elements of the receptor, differing greatly in length and sequence.

The malfunction of GPCRs are implicated in the pathology of many diseases and their progression such as ulcers, allergies, migraine, anxiety, psychosis, nocturnal heartburn, hypertension, asthma, congestive heart failure, Parkinson's, schizophrenia, and glaucoma (Schoneberg, T., et al., 2002, Pierce, K. L., et al., 2002). This makes GPCRs an essential target

for drug development. Unfortunately, despite their importance there is insufficient structural information on GPCRs for structure-based drug design. This is because these membrane-bound proteins are difficult to crystallize, and the atomic-level structure has been solved only for bovine rhodopsin (Palczewski et al., 2000; Teller et al., 2001). In fact, GPCRs only account for about 3-4% of the human genome (Palczewski et al., 2000), and yet are targets for more than 50% of the drugs in the current market (Teller, D., et al., 2001). Due to the difficulty in generating 3-D structures using high resolution X-ray diffraction data or NMR data for GPCRs, it is widely accepted that theory and computation to predict the 3-D structures of GPCRs from first principles can aid the structure-based drug design for many GPCR targets (Strader 1994, Parrill et al., 2000).

Successful protein structure prediction methods for globular proteins generally utilize homology to known structures (John B., and Sali A., 2003). To understand the functional mechanism of GPCRs, it is important to study the amino acid residues conserved in GPCR sequences. Conserved residues at the extracellular side affect ligand-binding selectivity, whereas those at the cytoplasmic side affect G-protein coupling selectivity. Many articles have reported the positions of conserved residues and the mutation experiments of key residues that have significant influence on the ligand binding or G-protein coupling selectivity (Nygaard et al 2009, Karmik et al 2003, and Mizadegan et al, 2003).

Mutagenesis experiments, medicinal chemistry and molecular modelling have suggested that an agonist binding pocket for cationic neurotransmitter receptors is located in the extracellular part of the transmembrane region between helices 3,4,5,6 and 7 (Strader et al., 1987; Hibert et al., 1991; Trumpp-Kallmeyer et al., 1992). Because GPCRs exist in the membrane environment, crystallization is difficult and the number of 3D structures available is

very small. However, in the last decade the number of 3D structures of GPCRs has increased rapidly. β_1- and β_2-adrenergic receptors (Cherezov et al ,2007; Rasmussen et al ,2007 ; Warne et al, 2008), A_{2A} adenosine receptors (Jaakola et al, 2008) and squid rhodopsin (Murakami, M.; Kouyama, T. et al,2008; Shimamura et al, 2008) were crystallized and their 3D structures completely solved. In 2010, the structures of CXCR4 chemokine (Wu et al., 2010), D3 dopamine receptors (Ellen et al., 2010) and adipokinetic Hormone receptor, AKHR (Mugumbate et al., 2010) were successfully determined. The structural data from these studies not only provides critical information on the molecular nature of drug interaction with GPCRs, but also provides approaches that allow for crystallization of other GPCRs, revealing more rational methods for drug development.

2.1.1 Structure of G Protein- Coupled Receptor

All G protein- coupled receptors share a common molecular structure which consists of an extracellular amino-terminus, seven transmembrane α-helices and an intracellular carboxyl-terminus. The seven transmembrane domains are linked with 3 intracellular loops (ICL 1, 2 and 3) and 3 extracellular loops (ECL 1, 2 and 3), all with variable length (Palczewski et al., 2000).The N-terminal end is often glycosylated and contains disulphide bridges, whereas the C terminal end contains several putative phosphorylation sites that can play a role in receptor desensitization and internalization. The ligand specific binding pocket can contain amino acid residues from the N-terminal extracellular end and / or residues from the extracellular loops. All GPCRs share a common signalling mechanism in that they interact with G-proteins, heterotrimeric GTP-binding proteins. Inactive GPCRs differ in the outer TM domains compared to the inner TM domains with TM6 and TM7 being the most structurally conserved helices. (Ishii et al, 2010).

2.1.2 GPCRs Classification

GPCRs superfamily is divided into six classes (Horn *et al.,* 2001). The first is class A, commonly known as the Rhodopsin-like GPCRs which accounts for over 80% of family members across species. Most Class A GPCRs share structural features such as an Asparagine residue in transmembrane (TM) 1, an Asparagine residue in (TM) 2, Proline residues in TMs 4-7, 3 consensus sequences for N-linked glycosylation at the N-terminus and 2 Cysteine residues in ECL1and ECL2 which are responsible for receptor stability. They also have phosphorylation and palmitoylation sites located at the C-terminal site which are involved in signal transduction and a DRY-motif at the end of the TM3 which is suggested to be involved in G-protein interaction (Howard *et al.,* 2001; Lee *et al.,* 1999). Class A receptors mostly bind to peptides, biogenic amines and lipids. The ligands bind in a cavity formed by TM3 and TM4. Research has found out that short peptides which activate class A receptors interact with the extracellular loops and the N- terminus (Bockaert and Pin, 1999). Receptors which bind to peptides are important in mediating neurotransmitters, hormones and paracrine signals, whereas biogenic amine binding receptors are good targets for drugs. Receptors in this class include; bovine rhodopsin, $\beta_1$and β_2 adrenoceptors and A_{2A} adenosine receptors .The second class is Class B also termed as Secretin-like receptors. These receptors bind large endogenous peptides such as glucagon and have a large N-Terminal on the extracellular domain, which binds the ligands. The third class is C and it's made up of Metabotropic glutamate-like receptors (mGluRs) also termed as excitatory neurotransmitters. The fourth is Class D, which comprises of highly-diverged receptors for peptide binding. This class of receptors lack many characteristic features found in class A receptors. They lack the ERY motif in TM3, disulfide bond between loop 2 and TM3 and NPxxY motif in TM7. Class E comprises of cAMP receptors found in protozoan amoeba

Dictyostelium discoideum which help in chemotactic signalling systems (Pin et al, 2003). The last Class is F which contains Frizzled/ smoothened receptors found in *Drosophilla melanogaster.*

2.1.3 Crystal structures of GPCRs

Structure-based, rational drug discovery has been widely used to design compounds with optimal specificity and potency to molecular targets. The structural information is important because it gives the physical properties of the ligand binding site which helps in designing a pharmacophore structure of the ligand when docked to the GPCR. This is the foundation for virtual screening of molecules with predictable chemical characteristics as well as optimal functional properties to modulate the molecular target (Hubbard, R. 2011; Richard M. Eglen et al., 2011). Rhodopsin was the first receptor to be subjected to X-ray analysis and it belongs to the Class A GPCRs. GPCRs are difficult to crystallize because they are structurally unstable in purified form and this has been attributed to the unordered structure of the large third intracellular loops of the receptors (Kobilka and Deupi, 2007). Since all GPCRs constitute a single polypeptide chain that spans the plasma membrane seven times with seven alpha helical structures (Costanzi et al., 2009), it has been found that the helical bundle of most Class A GPCRs hosts a ligand binding cavity opened toward the extracellular loops, which provides access to diffusible ligands. However, the cavity of rhodopsin is sealed from the extracellular space by the second extracellular loop (ECL2). Thus there is likelihood that the hydrophobic ligands of this receptor make their way into the binding cavity through the transmembrane domain. For most GPCRs this ligand binding cavity is lined by transmembrane domains (TMs) 2, 3, 5, 6, and 7 and is deeper in proximity to TMs 5 and 6 while shallower in proximity to TMs

2, and 7 although, in some cases TMs 1 and 4 also form the pocket and can impact ligand binding.

2.1.4 G -Protein Activation

G-protein Coupled Receptors, (GPCRs) are activated by extracellular ligands that initiate signalling pathways in the cell through trimeric G-proteins. Their activation by extracellular ligands and the diversity of the receptors and the ligands make them highly interesting to the pharmaceutical industry and one of the richest targets for drug discovery. For about 80 % of the known human GPCRs the activating ligand is unknown (Schöneberg, T et. al, 2002). G-proteins transmit the signal to effector proteins, such as enzymes, hormones, chemokines, neuropeptides, neurotransmitters, photons and ion channels. When a specific ligand binds to its corresponding receptor, the ligand stimulates the receptor to activate a specific heterotrimeric guanine nucleotide- binding regulatory protein coupled to the intracellular region of the receptor. The G protein then transmits a signal to an effector molecule within the cell by stimulating or inhibiting the activity of the receptor molecule. The effector molecule includes adenylate cyclase, phospholipases and ion channels. Adenylate cyclase, phospholipases are enzymes involved in the production of the second messenger molecules cAMP, inositol triphosphate and diacyglyceral. This sequence of events makes an extracellular ligand stimulus to exert intracellular changes through the G protein coupled receptor. Since GPCRs play a vital role in the communication between cells and their environment, this makes them attractive targets for modern drugs. GPCRs having known ligands, agonists and antagonists can be identified to mimic and enhance the action of the ligand. However, not all receptors that activate G-proteins are members of the superfamily of GPCRs. Activation of G proteins has also been implicated in signal transduction

mediated by several tyrosine kinase receptors, including the receptors for epidermal growth factor (EGF).

2.1.5 Homology Modelling of GPCRs

Homology modelling is the generation, manipulation and representation of three-dimensional molecular structures and their associated physical, chemical, biological and pharmacological properties from its amino acid sequence and a three-dimensional structure of a related homologous protein (Ravna, 2006). Molecular modelling has become an essential tool in a number of fields of science, including chemistry, physics, drug discovery and biochemistry. Despite an increase in the number of protein structures in the PDB database their experimental structures are still unknown because they are too large for Nuclear Magnetic Resonance analysis and cannot be crystallized and characterized using X- Ray crystallography. Thus homology modelling is employed to address the problem. The constructed three- dimensional model is used as an aid in interpreting ligand binding and molecular biological data and the quality and usefulness of the model will depend on the assumptions made in its construction and its supporting biophysical and molecular biological data. The model will be used to 1) visualize the protein interior and propose model of ligand binding 2) plan mutagenesis experiments and 3) support ligand design. The known protein structure is termed as the "template" structure, whereas the protein with unknown structure is the "target" protein.

Homology modelling relies on the identification of one or more known protein structures which are likely to resemble the structure of the query sequence, and on the production of an alignment that maps residues in the query sequence to residues in the template sequence. The homology modelling approach is based on the observation that structure is more conserved than

sequence, such that a known protein structure can be used to construct a model of a homologous protein. Related proteins have similar sequences and naturally occurring homologous proteins have similar protein structure. The sequence alignment and template structure are then used to produce a structural model of the target and the quality of the homology model depends on the quality of the sequence alignment and template structure. Homology modelling has produced high-quality structural models when the target and template are closely related leading to the formation of structural genomics group which produces experimental structures for all classes of protein folds (Williamson A.R., 2000). Improper template selection lowers the sequence identity hence making the model inaccurate. Both Molecular mechanical calculations, (calculations on atomic nuclei) and quantum mechanical calculations (calculations on the electronic systems of molecules) are included as techniques in molecular modelling. The prediction can subsequently be used as guideline for the construction and characterization of point mutations, studies of ligand-receptor interaction, and the design of new leads by application of flexible docking and virtual screening methods.

2.1.5.1 GPCR Template Structures

2.1.5.1.1. Bovine Rhodopsin Receptor

Bovine rhodopsin has been successfully used as a template for the generation of various GPCRs homology models since 2000 when the crystal structure was released (Palczewski, K. 2000). The crystal structure of bovine rhodopsin has been extensively used because it can easily be obtained in relatively high quantities.

2.1.5.1.2 Adenosine A_{2A} Receptor

The human adenosine A_{2A} has four disulfide bonds in the extracellular domain with a subtle repacking of the transmembrane helices relative to the rhodopsin and adrenergic receptor structures.

2.1.5.2 Steps in Homology Modelling

2.1.5.2.1 Searching structure database for homologous sequences.

This is the first stage of homology modelling done to identify a suitable template structure to base the prediction of the protein structure of interest. This is done by performing search using standard comparison tool such as BLAST and FASTA (Altschul et al, 1990; Pearson and Lipman, 1988) against a number of structures in the PDB database. These methods apply a scoring system for differences between sequences, with the substitution of chemically similar residues incurring a small penalty and insertions, deletions and substitutions of non-similar residues having larger penalties associated with them. The sequences with the lowest penalty scores are identified as possible templates. If areas of the template structure are poorly defined in one template then it is sometimes possible to use multiple structures and to take the most well defined areas from each (Krieger et al, 2003).

2.1.5.2.2 Aligning the target to the template

Sequence alignment is done to match the order of amino acids of the target protein against that of a suitable protein structure as the template (Mount, 2003). They allow the search for candidate template structures based on sequence similarity to the target sequence. Allowances are made because mutations are conservative in that changes in amino acid side chain only slightly alter the biochemical properties. In regions of low identity pairwise alignment can prove

to be hard and this can be improved by using a third intermediate sequence which is similar to the target in the low identity region. Programs such as CLUSTALW (Higgins et al, 1996) can perform these multiple sequence alignments i.e. comparisons of several homologous protein sequences allow conclusions about highly conserved, homologous regions and areas of rather insignificant functional residues. Sequence alignments can be used to find characteristic motifs and conserved residues in protein families and to improve prediction of secondary structure elements. This information can be used for manually adjusting an automatically created pairwise alignment for an optimal transfer of structured elements/parts. This can be particularly useful when aligning areas where there are insertions or deletions between the sequences.

2.1.5.2.3.Loop modelling

Loops are fragments of a protein chain that connect the secondary structure elements. They also play a vital role in protein folding, stability and determine the functional specificity of a given protein molecule. They mediate important biological processes as most of them are found on active and binding cavities. Due to their role in protein function, loops are also an important consideration in protein engineering (Thanki et al., 1997). Because loops are often on the surface of protein structures; they are susceptible to insertions and deletions of amino acids. Loop modelling involves fitting a generated loop conformation with a given protein structure so that the loop connects with the rest of the protein structure and completes it. This is achieved by aligning the regions of the target GPCR sequence that are not aligned to the template. This is a common source of modelling error and occurs with higher frequency when the target and template have low sequence identity.

2.1.6 Molecular Dynamics

Molecular dynamic simulation is a tool used to study the behaviour of a given system whereby the position of every atom is computed as a function of time by an algorithm that solves Newtonian classical equation of motion and results in trajectories for all atoms in the system. The aim of computer simulations of molecular systems is to compute macroscopic behaviour from microscopic interactions. According to the Laws of thermodynamics, molecules spontaneously seek the lowest energy. By performing energy minimization calculations on the model, the lowest energy conformation of the model is calculated. Molecular dynamics is the simulation of molecular motion during a short period of time and is used to refine a model further (Ravna 2006). Since chemistry concerns the study of properties of substances or molecular systems in terms of atoms, the basic challenge facing computational chemistry is to describe or even predict the structure and stability of a molecular system, the (free) energy of different states of a molecular system and reaction processes within molecular systems in terms of interactions at the atomic level. In MD simulations force fields are required in order to solve the task as they account for both the Born-Oppenheimer Approximation position of nuclei and electrons of each atom considered. MD refers to the use of classical mechanics in order to describe the physical basis behind the model and to calculate the potential energy of the system using force fields. This method is used to solve both biochemical and biophysical problems such as conformational analysis of proteins, ligand-receptor interactions and drug design. Factors that govern the outcome of MD simulations are as follows:

1. Force field parameters
2. Boundary conditions
3. Temperature and pressure control

4. How non-bonded interactions are treated
5. Time step integration
6. Choice of the degrees of freedom

2.1.6.1 Force Fields

A force field is a set of parameters and mathematical equations used to describe the properties of atoms and their bonded and non-bonded interactions. Force fields include descriptors of the atomic masses and charges for different atoms as well as bond lengths, bond angles and dihedral angles. These parameters define the behaviour and potential energy of the system. Non-bonded interactions are calculated between all pairs of atoms that are in different molecules or in the same but separated by at least three bonds. Non-bonded interactions comprise of both electrostatic and van der Waals interactions (Guvench, O and Mackerell, A. D., 2008). Molecular mechanics relies on the laws of Newtonian physics and experimentally derived parameters to calculate geometry as a function of potential energy based on a force field (**equation 1**).

$$\mathbf{E}_{tot} = \mathbf{E}_{bonds} + \mathbf{E}_{angle} + \mathbf{E}_{dihedral} + \mathbf{E}_{electrostatic} + \mathbf{E}_{van\ der\ Waals} + \mathbf{E}_{other} \quad \text{..... (equation 1)}$$

Where $\mathbf{E}_{tot}$= total energy

$\mathbf{E}_{bonds}$=the energy resulting from deforming a bond length from its natural value.

$\mathbf{E}_{angle}$ = the energy due to deformation of an angle from its natural value

$\mathbf{E}_{dihedral}$ = energy from deforming the dihedral angle

$\mathbf{E}_{electrostatic}$ = the energy arising from coulomb forces

$\mathbf{E}_{van\ der\ Waals}$= energy arising from van der Waals non-bonded interactions, and

$\mathbf{E}_{other}$= terms that are specific for a certain force field.

Intramolecular potential energy for bond stretching and angle bending are represented by a harmonic potential and the torsion potential is described by a periodic cosine function. On the

other hand, for non-bonded interactions the electrostatic interactions are computed based on the Coulomb potential and the Lennard-Jones potential is commonly used for the van der Waals interactions.

2.1.6.2 Energy Minimization

Energy minimization is a method, which minimizes the potential energy of a system's structure to find the local minimum starting from the initial conformation. Energy minimization is important because after sequence structure alignment, during loop regions insertion and amino acid side chain addition there exist some steric clashes and distorted bonds present in the model. Thus energy minimization relaxes the worst conflicts in the resulting structure and finds an energetically favourable conformation of the system in order to be able to start a simulation.

2.1.6.3 GROMACS

GROMACS is a widely used molecular dynamics simulation package developed at the University of Groningen. Information on GROMACS can be found at http://www.gromacs.org/. GROMACS is used for energy minimization and molecular dynamics simulations of complex organic compound with very high molecular weights like proteins, nucleic acids and peptides (Ohno et al., 2011; Lazar et al., 2010). For a simulation to run a number of parameters should be put in place which include: a) a coordinates file for all atoms b) information on the interactions of bond angles, charges and Van der Waals interactions and c) parameters to control the Molecular Dynamic simulation. The .pdb and .gro files contain the coordinates for all atoms and it is the input structure file for MD simulation while the interactions are listed in the topology file and the input parameters are put into an .mdp file. During MD simulation the protein data bank

(.pdb) structure file is converted to a GROMACS structure file (.gro) with the generation of a descriptive topology file. The topology files preparation for performing EM and MD simulations with GROMACS is very important in investigations of biological processes involving drug interactions with their molecular targets. The force fields available for biomolecules in the literature (Jorgensen et al, 1996 and Tiradorives, 1988) are generally based on macromolecules, such as proteins and nucleic acids. The main purpose of MD simulation is to generate trajectory molecules in a limited time period, become a bridge between theory and experiments and allow chemists to make simulation that can't be done in the laboratory. Energy minimization is done to the structure to release strain. GROMACS uses periodic boundary condition and group concept. Periodic boundary condition is a classical way used to reduce edge effect in a system whereby the atoms are placed in a box and surrounded by a copy of the same atoms **(Figure 2.1**). The model boxes used in GROMACS are either triclinic, cubic or octahedron. The other concept of group is used to show an action and each group can only have a maximum number of 256 atoms and each atom can only have six different groups (Van der Spoel, D. et al, 2005).

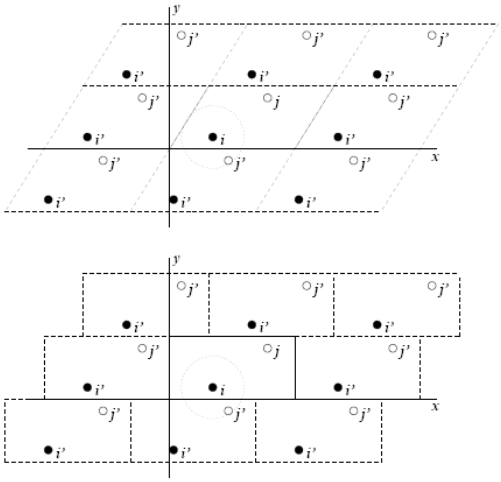


Figure 2.1: Periodic boundary condition in Two Dimensions

2.1.6.4 Simulated Annealing

Simulated annealing is a technique used to optimize molecules in order to overcome local minima during conformational space search.

2.1.6.5 Ramachandran plot

Ramachandran plots, where the psi and phi angles of amino acid residues are plotted against each other, may be constructed to evaluate a model built by homology modelling. Amino acids are linked together by covalent bonds at the Cα atoms and the only degrees of freedom they have are rotations around these bonds, i.e. around the Cα -C bond (psi, ψ) and the N-Cα bond (phi, φ), as the CO-NH (the peptide bond) is rigid and planar due to the double bond character. The angle pairs ψ and φ are plotted against each other in a diagram called a Ramachandran plot, which shows allowed combinations of the ψ and φ angles.

Most combinations of ψ and φ angles for an amino acid are not allowed because they cause steric collisions between the side chains and main chain. Each point in the Ramachadran plot represents psi and phi values for an amino acid residue (Branden 1999). α helices in proteins are found when a stretch of consecutive residues all have the φ ,ψ angle pair approximately -60° and -50°, corresponding to the allowed region in the bottom left quadrant of the Ramachandran plot. In the Ramachandran plot, left-handed α helices are seen in the upper right quadrant of Ramachandran plots, whereas β strands are seen in the upper left quadrant of the plot.

Residues that may be expected to be located outside the allowed regions of the Ramachandran plot are glycine and proline. The glycine side chain consists of only hydrogen atom and steric collisions do not occur that often as with the other amino acid residues that have

longer side chains. In the case of proline, the side chain is bonded to the main chain nitrogen atom and forms a ring structure, thereby preventing the nitrogen atom from participating in hydrogen-bonding and also providing some sterical hindrance to the α helical conformation.

2.1.7 Molecular Docking

In molecular docking approach, a ligand is automatically placed into a predetermined, predefined binding site of a 3D receptor structure model. Molecular docking has a wide range of application in computational structure biology such as preparing the ligand or receptor simulations, validating experimental data or evaluating predicted structural models. The initial step for docking is to define the binding site by using experimental information derived from mutagenesis or comparing homologous structures with known binding sites. This information is not always available; therefore structural methods have been developed to find the binding sites using the sequence conservation information.

Molecular docking aims at prediciting the structure of the intermolecular complex that is formed between two or more molecules and has become a useful tool in structure-based drug design and discovery (Sousa et al. 2006). When performing molecular docking, protein flexibility is a critical aspect and the best approach is to treat both the target protein and the ligand as flexible molecules. Proteins are not rigid structures and ideally the proteins should be flexible during the docking simulations. However, because of limitations in computing, Auto Dock and Auto Dock Vina treats the ligand as flexible and the receptor protein as rigid. Molecular docking is the most widespread method used to calculate protein-ligand interactions as it predicts all the potential ligand binding sites on the entire protein target. Dynamic studies on binding modes are useful in elucidating the structural characteristics and interactions in order in

provide relevant data for designing agonists and antagonists of the protein of interest as shown in **Figure 2.2.**

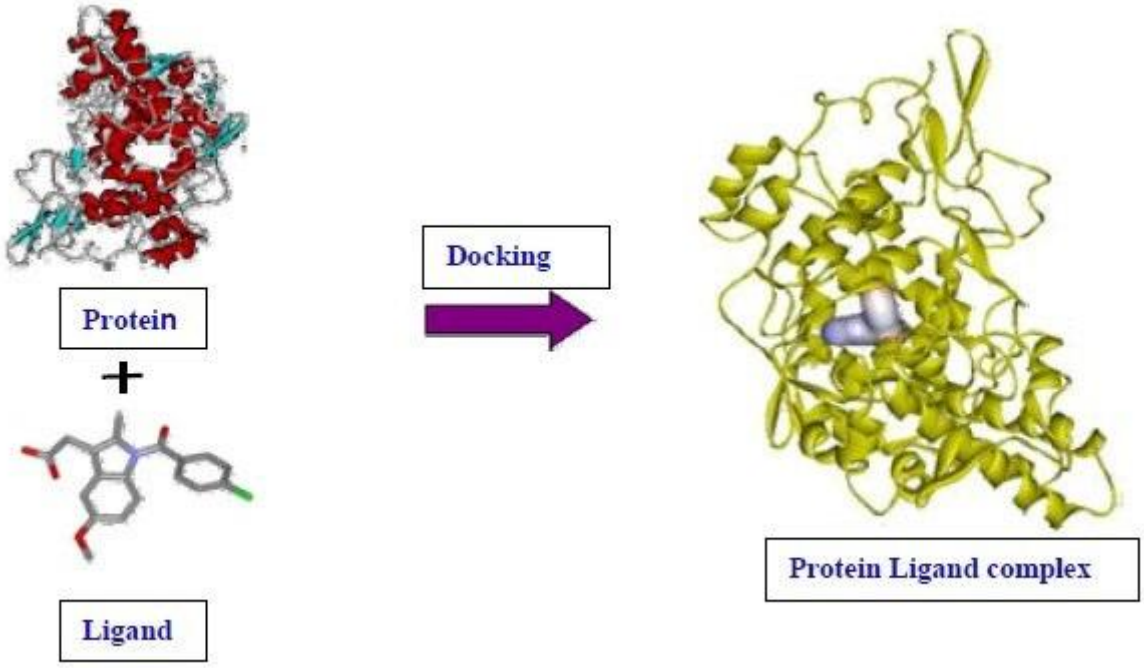


Figure 2.2: Schematic diagram of docking protocol

2.1.7.1 Docking Tools: ***Autodock 4.2***

AutoDock is a suite of automated docking tools which is designed to predict how small molecules like substrates or drug candidates, bind to a receptor of known 3D structure. It is used to perform computational molecular docking of small molecules to proteins. AutoDock actually consists of two main programs namely AutoDock and Auto Grid. AutoDock performs the docking of the ligand to a set of grids describing the target protein while Auto Grid precalculates these grids. It uses a scoring function based on the AMBER force field, and estimates the free energy of binding of a ligand to its target. AutoDock consists of three search Algorithms namely: 1) Global Search Algorithm 2) Local Search Algorithm and 3) Hybrid global- local Search

Algorithm which is the most powerful of all. AutoDock has been extensively used over the years because it's accurate, reliable and can model flexible ligands to proteins.

2.1.8 Virtual Screening

Many steps are employed in the drug discovery process: disease selection, target hypothesis, lead compound identification, lead optimization, pre-clinical trial and pharmacological optimization. Since it's not easy to speed up the preclinical tests then the only way to fasten the process is to act on the preclinical steps. High throughput screening (HTS) is the common technique used where millions of compounds on a target are screened to identify the hits (Mestres, 2002). Virtual screening has become an integral part of the drug discovery process in recent years. It uses computer-based methods to discover new ligands on the basis of biological structures to select compounds by evaluating their desirability in a computational model (Mestres, 2002) which comprises of high potency, selectivity towards the target protein, appropriate pharmacokinetic properties, and favourable toxicology. There are some important points to be considered for virtual screening: the availability of the compounds to be screened against the receptor, the knowledge about the structure of the receptor and the receptor ligand interactions, and the knowledge about drugs and drug characteristics (Lyne, 2002). Virtual screening is important because diverse hits can be identified leading to more diverse lead compounds.

There are two approaches used in virtual screening of compounds for GPCRs activity namely structure based screening and ligand based screening. Structure based screening requires knowledge of the 3D structure of the target protein's binding site to prioritize compounds by

their likelihood to bind to the protein whereas in ligand-based screening no information on the protein is required and compounds known to bind to the protein are used as a structural Query (Dror et al, 2004; Schneidman-Duhovny et al, 2004) and a compound's similarity to certain query features determine the likelihood for high affinity towards the particular receptor. Due to the limited structural data for GPCRs, ligand based screening is the main technique used in designing drugs for the GPCR family. This is achieved by using information from the compounds that are known to bind to the target GPCRs then the data is used to identify other molecules in the various databases with similar properties in order to improve their biological activities (Stahura, 2005; Bajorath, 2005; Lengauer et al, 2004). Structure-based drug design is a method used in rational drug design and pharmaceutical research. The 3D structure of the target is obtained by performing X-ray crystallography, NMR experiments or predicted by homology modelling (Evers, 2004; Klebe, 2004; Oshiro et al, 2004 and Wieman et al, 2004). The approach used in structure-based virtual screening is to identify the binding pose of each small molecule in a test library (docking), and from that identify the free energy of binding of that molecule (scoring). The set of hit compounds is then predicted by sorting all compounds in the test library by this score and deciding on a threshold score. Virtual screening is fast and allows identifying possibly active compounds with a completely different scaffold than the existing compounds, and it is thus a valuable tool in finding novel drug candidates. Homology modelling is a reliable method used to retrieve known antagonists via structure-based virtual screening from several compound databases (Wieman, et al, 2004; Evers, 2005; Klabunde, 2005 and Varady et al, 2003).

2.2 FMRFamide-like neuropeptides (FLPs)

FMRFamide-related peptides (FaRPs) are a diverse family of vertebrate and invertebrate neuropeptides possessing a C-terminal RF-amide amino acid sequence. FaRPs are identified by the first tetrapeptide identified with this defining terminal sequence, the molluscan cardioexcitatory tetrapeptide FMRFamide (Price and Greenberg, 1977), and compose an extensive neuropeptide superfamily with multiple physiological and behavioural functions in insects.AF2 peptide belongs to the FMRFamide-like neuropeptides distributed throughout the nervous system of nematodes, including *Ascaris suum* (Cowden et al., 1993), *Haemonchus contortus* (Keating et al., 1995), and *Caenorhabditis elegans* (Nelson et al., 1998a) and have been implicated in a wide-range of behaviours including locomotion and feeding (Nelson and Li, 1996; Brownlee et al.,1995). Neuropeptides are presumed to be either neurotransmitters or neurohormones involved in the regulation of both physiology and behaviour of nematodes (Halton et al. 1994). AF1 (KNEFIRF–NH_2) and AF2 (KHEYLRF–NH_2), were the first two nematode peptides structurally characterized after their isolation from the *ascaris suum* head (Cowden et al. 1980, 1993). AF1 and AF2 are found to have complex effects on dorsal muscles of *A .suum* which includes stimulation, relaxation and rhythmic activity (Walker et al. 1993; Bowman et al. 1996).

3.0 EXPERIMENTAL METHODS

3.1 AF2 PEPTIDE

3.1.1 AF2 structure generation

The starting AF2 peptide structure Lys-His-Glu-Tyr-Leu-Arg-Phe-NH_2 was generated using Insight II software (Accelrys Inc., San Diego, CA.) and energy minimized using the force field available in Discover studio (Accelrys Inc., San Diego, CA.). The energy minimized model was used as the starting structure for MD simulations using GROMACS software (Van Der Spoel D, *et al*, 2005). All these were performed on four processor computer clusters available at University of Cape Town, South Africa.

3.1.2 Molecular Dynamics Simulations *in vacuo*

Molecular dynamics were performed using GROMACS software version 3.3 installed on cluster computers available in University of Cape Town, South Africa (Van Der Spoel *et al*, 2005). This was done to search for lowest energy conformation of AF2 peptide using Optimized Potentials for Liquid Simulations- All Atom (OPLS-AA) force field (Jorgensen et al, 1996) which provides parameters for every type of atom in a system. A box size of $60.81nm^3$ was used. Distance restraints between protons were applied from the provided Nuclear Magnetic Resonance inter-proton distances. MD simulations were carried out *in vacuo* using simulated annealing protocol. This was done to overcome energy barriers. Structures were collected every 25000 ps and then cooled from 600 K to room temperature (300 K) in 10 ns. 200 structures were collected. Energy minimizations were carried out with the steepest descents, conjugate gradient and L-BFGS integrators, until gradient convergence was achieved. On clustering the structures into families of room mean square deviation< 0.15 Å by comparing the entire backbone atoms, one cluster was found.

3.1.3 Molecular Dynamics Simulations in water

The 200 structures obtained from *in vacuo* MD simulations were analysed and the structure with the lowest energy was selected as the starting structure for MD dynamics in water. The lowest energy structure was generated after 2950 ps and was found to have energy of -61219.51 kcal/mol. The same box dimensions were used and the peptide was solvated with 1296 water molecules using the SPC water model. The system was energy minimised and simulations were carried out with inter-proton distance restraints. The Nuclear Magnetic Resonance experiments of the peptide were performed on a DRX- 500 MHz machine at the Department of Analytical Chemistry, University of Debrecen, Hungary. Measurements were done in both water- acetate buffer and DPC- D38 membrane mimic.

3.2 AF2 RECEPTOR

3.2.1 Secondary structure, transmembrane topology and fold recognition prediction

The available primary sequence of the AF2 receptor (T19F4.1b) was used to predict the 3D structure of the receptor. PSIPRED, MEMSAT3 and pGenTHREADER software, available on http://globin.bio.warwick.ac.uk/psipred/, were used to predict the secondary structure, transmembrane topology and fold recognition of the receptor (Jones et al., 1994; Jones, 1999). PSIPRED predicts secondary structure by incorporating a two feed- forward neural network's by performing analysis obtained from PSI- BLAST (Buchan et. al, 2010). MEMSAT3 predicts the transmembrane topology (Lobley, et al., 2009) while pGenTHREADER predicts the fold recognition and identifies distant homologues by making use of profile- profile alignments and the predicted secondary structure using PSIPRED as the input (Nugent& Jones2009).

3.2.2 Sequence alignment and Homology Modelling

From the fold recognition results, four GPCR crystal structures were selected for multiple alignments. AF2R is a class A rhodopsin-like receptor because of its sequence homology and functional similarity found in this family. Four receptor crystal structures namely beta2-adrenergic (2RH1), rhodopsin (1U19), Histamine (3RZE) and adenosine A_{2A} (2YDV) from the Protein Data Bank (Berman et al., 2000) were used as templates to perform homology modelling of the entire receptor. Sequence multiple alignment of AF2R was performed with the templates obtained from PDB using ClustalW software tool (Higgins et al., 1996). Manual adjustments were done to remove gaps from the helices. Crystal structures with highest percent identity were used as templates during homology modelling of the target receptor using Modeller 9 .9 (Sali and Blundell, 1993). The AF2 receptor was automatically generated and constructed using the Modeller software. The model generated was complete showing the helices and the loop regions.

3.2.3 Molecular Dynamics Simulations of AF2R

The available primary sequence of AF2 receptor was subjected to Molecular Dynamics *in vacuo* using GROMACS software version 3.3 (Van Der Spoel et al, 2005) to search for lowest energy conformation of AF2 Receptor using a box size of 3243.98nm^3. Molecular Dynamic simulations were carried out with frozen helices for 50ps at 300K using a time step of 0.001 ps without distance restraints. The C-terminus amino acid residues were separated from the structure because it was thought to be too lengthy for MD simulations. The structure was energy minimized using steepest descents, conjugate gradient and L-BFGS until gradient convergence was achieved and the root mean square value at 0.10 Å. The minimized structure was further

refined using Simulated Annealing where the structure was heated to 1500K in 30ps followed by 10ps at a constant temperature of 300K. This was done to overcome the energy barriers.

3.2.4 Molecular Docking

The automated docking software Autodock Vina and Autodock tools were employed in the identification of the binding site of AF2 GPCR using AF2 peptide as the ligand (Morris, et al. 1998).Docking was performed using a rigid receptor and a flexible ligand. The number of rotatable bonds in the peptide was reduced from 32 to 28 and number of torsion degrees of freedom set at 28.Generic Algorithm docking parameter which gives energy and the different conformations assumed by the ligand in the receptor binding pocket was applied. The molecule was subjected to 10 trials with a population size of 150 and number of evaluations set to 2.5 million.

3.2.5 AF2R Virtual Screening

The compound library was obtained from ZINC database (Irwin et al. *2012),* which contains over 21 million chemical compounds from different vendors. It is an open source database and the structures are filtered according to the Lipinski's rule of five (Lipinski, et al, 1997) that describes molecular properties important for a drug's pharmacokinetics in the human body. The properties include: absorption, distribution, metabolism and excretion (ADME). Pipeline Pilot Software was used to filter the ZINC compounds using the Lipinski's based filter. Compounds were filtered based on their chemical description to remove unsuitable compounds due to undesired and toxic properties. The general filter 'Lipinski's rule of five' was used. The 'Lipinski's rule of five' states that a drug-like molecule should have a molecular weight less than

500, a log P value of less than 5, less than 5 hydrogen bond donors and less than 10 hydrogen bond acceptors. Duplicates were also removed and the molecular weight for the compounds screened was adjusted to 900 since the peptides molecular weight is 996.1862 atomic mass unit.

4.0 RESULTS AND DISCUSSION

4.1 AF2 PEPTIDE

4.1.1 NMR Chemical Assignments and Inter-proton Distances

NMR spectra were recorded on a DRX- 500 MHz spectrometer in a water-D_2O buffer solution. The proton chemical shifts of the AF2 peptide are given in **Table 4.1**

Table 4.1: ^{1}H chemical shift (ppm) for the AF2 peptide in water.

Residue	**N^H**	**H_α**	**$H_{\beta 2}$**	**$H_{\beta 3}$**	**$H_{\gamma 1}$**	**$H_{\gamma 2}$**	**$H_{\gamma 3}$**	**$H_{\delta 1}$**	**$H_{\delta 2}$**	**other**
Lys[1]	-	4.018	1.869	1.885	-	1.373		-	1.686	2.983
His[2]	-	4.636	3.130	-	-	-		7.594	7.065	8.503, 7.153
Glu[3]	8.696	4.297	1.941	1.847	-	2.219	2.221	-	-	-
Tyr[4]	8.484	4.555	3.013	2.943	-	-	-	7.129	7.129	6.800, 6.800
Leu[5]	8.247	4.285	1.557	1.449	1.447	-	-	0.902	0.841	-
Arg[6]	8.269	4.193	1.664	-	-	-		-	3.133	
Phe[7]	8.279	4.615	3.176	3.008	-	-		7.276	7.276	7.354

Inter-proton distances were calculated by measuring the cross-peak intensities in the NOESY spectra using the Isolated Spin Pair Approximation (ISPA) method. The limits for the lower and upper bound of the inter-proton distance was calculating by subtracting ~10% and adding 40% respectively as tabulated in **table 4.2**. These distances were used to perform distance restrained Molecular Dynamic simulation of AF2 peptide. In order to determine the solution conformation of the peptide NMR experiments were performed in water at different temperatures as shown in **Figure 4.1.**

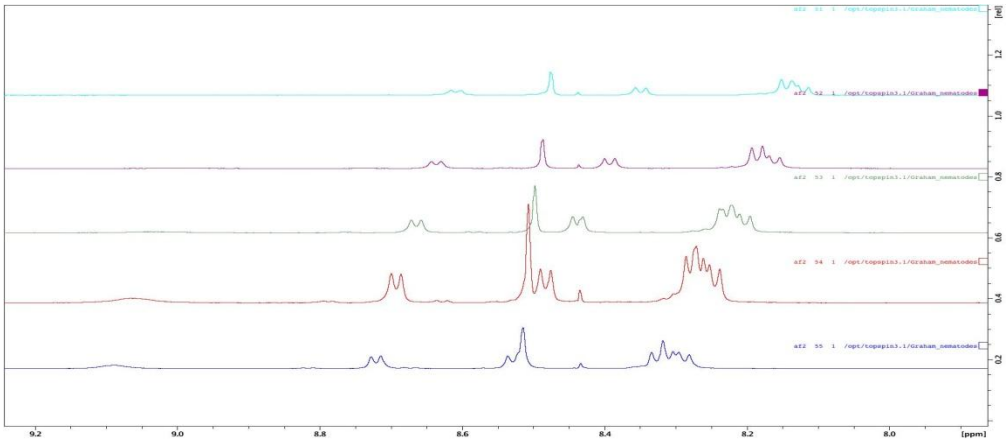


Figure 4.1: NMR spectra of AF2 in water at varying temperatures. Blue at 280°C, Red at 285°C, Grey at 290°C, violet at 295°C and arctic at 300°C.

Table 4.2: NOE derived inter-proton distances for AF2 in water, with lower and upper bound restraints.

Protons	**Distance (nm)**	**Lower bound (nm)**	**Upper bound (nm)**
$Glu^{3}H^{N}$ - His^{2} $H_{\beta 2}$	0.32	0.30	0.40
$Glu^{3}H^{N}$ - $His^{2}H_{\alpha}$	0.23	0.21	0.30
$Tyr^{4}H^{N}$ - $Glu^{3}H_{\beta 2}$	0.34	0.32	0.40
$Tyr^{4}H^{N}$ - $Glu^{3}H_{\alpha}$	0.22	0.20	0.30
$Phe^{7}H^{N}$ - $Arg^{6}H_{\alpha}$	0.25	0.23	0.30
$Phe^{7}H^{N}$ - $Arg^{6}H_{\beta 2}$	0.29	0.27	0.40
$Phe^{7}H^{N}$ - $Arg^{6}H_{\gamma 2}$	0.31	0.29	0.40
$Phe^{7}H^{N}$ - $Arg^{6}H_{\gamma 3}$	0.33	0.31	0.40
$Arg^{6}H^{N}$ - $Leu^{5}H_{\alpha}$	0.22	0.20	0.30

4.1.2. Molecular Dynamic Simulations of AF2 Peptide

Successful molecular dynamic simulation was performed in vacuo and 200 structures were collected. On performing cluster analysis only one cluster was found. The structure with the lowest energy was used as the starting structure for molecular dynamics in water, which was carried out using GROMACS software and employing distance restraints in order for the simulation to converge faster. When a Ramachandran plot of the phi and psi angles was plotted it was found that 5 residues were in the favoured region. Glutamic acid was the only residue outside the Ramachandran plot as shown in **Figure 4.2**. The experimental chemical shifts and the Ramachandran plot of the phi/psi angles, clearly shows the peptide has a beta turn with most residues in the favoured beta sheet region.

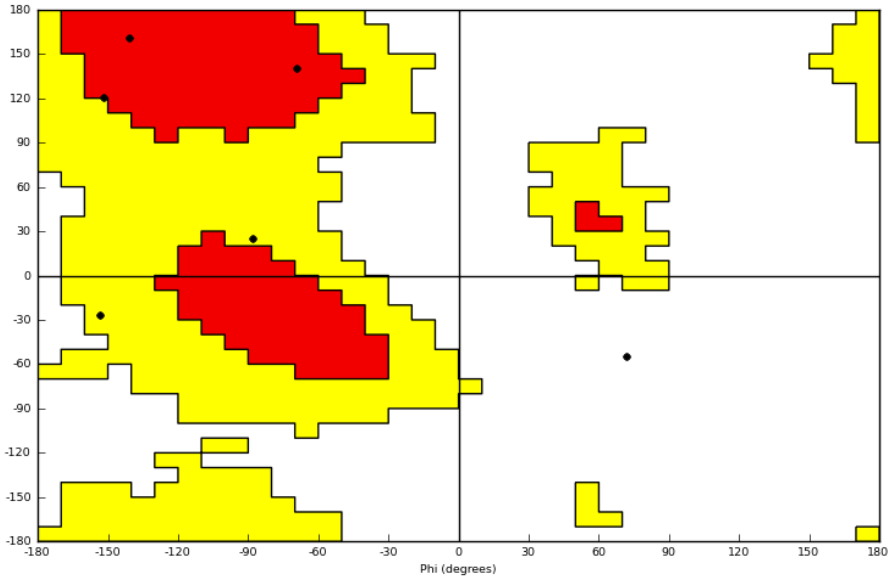
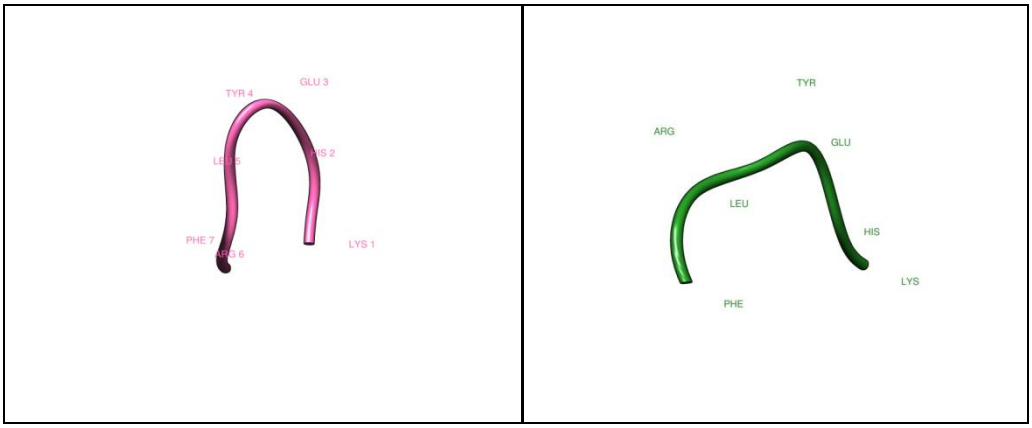


Figure 4.2: The Ramachandran plot of AF2 receptor. The most favoured regions are indicated as red, generously allowed are indicated as yellow and disallowed region indicated as white fields.

On doing cluster analysis only one cluster was found which shows that the peptide structure is highly conserved. **Figure 4.3A** shows the AF2 peptide structure with the lowest energy conformation before carrying out molecular docking. **Figure 4.3B** shows how the peptide opened after performing docking studies.



A **B**

Figure 4.3: (A) AF2 peptide conformation before docking and (**B**) AF2 peptide after docking.

The AF2 peptide was found to be closed after carrying out molecular dynamic simulations on it as shown in figure **4.2 A** above. There were no observable hydrogen bond interactions within the side chains of the peptide amino acid residues. However, the peptide was found to have a hydrophobic side composed of Tyrosine and Phenylalanine. This is more evident in molecular docking with the receptor as discussed in section 4.1.4.

4.2 AF2 RECEPTOR

4.2.1 Secondary structure, transmembrane topology and sequence alignment

The available primary structure of the AF2 Receptor (GenBank ID, CCD68155.1), from parasitic nematodes was used to predict the 3D structure. The secondary structure and

transmembrane region predictions by PSIPRED and MEMSAT3 software indicate that 179 residues make up the seven transmembrane alpha helices (**Figure 4.4**).

Figure 4.4: Predicted secondary structure of AF2R Pred: Predicted secondary structure (H=helix, E=strand and C=coil). AA=Target (AF2R). The highly conserved residues are in blue.

The helices were found to contain highly conserved residues found in class A GPCRs as follows: helix 1=N, helix 2=D, helix 3=ERW, helix 5=FLVP, helix 6=CNTLP and helix 7=NLLVVL.

These highly conserved residues were used as reference points for aligning each helix during the alignment. In contrast to other rhodopsin-like receptors the N-Terminus of AF2R is short with only six residues as indicated above. The AF2R C-Terminal was predicted to have 122 residues and consists of short helices. Multiple sequence alignments of the AF2 receptor transmembrane domains with those of the beta-2 adrenergic and adenosine A_{2A} receptors gave the best percent identity of 18.8% and 17.1% respectively. All gaps were manually removed from the helix region. Poor sequence identity in GPCRs is a normal phenomena in the GPCR superfamily. The GPCRs normally have very low sequence identity but high sequence similarity hence similar structural and functional features.

Table 4.3: Overall percent identities between the AF2, Beta-2 and Adenosine A_{2A} receptors

AF2 Receptor	Beta-2 adrenergic receptor	Adenosine A_{2A} receptor
Reference	18.8%	17.1%

4.2.2 Homology Modelling of AF2 Loops and the Helices

In general the transmembrane regions of GPCRs share a similar topology and the sequence alignment is governed by conserved residues in the helices. The extracellular loop 2 is the most variable region and it plays a vital role in homology modelling due to its involvement in ligand recognition, which is associated with the disulfide bond formed between TM3 and EL2. The presence of the disulfide bond is deemed critical in determining the conformation of the binding pockets for the ligands. It has been found that the conformation of the ECL2 is different in various GPCR crystal structures, leading to different binding pockets. MODELLER 9.0 Version, a spatial restraint method was used to construct the AF2 receptor model by satisfying

restraints derived from the template structure. The restraints include bond lengths and angles, van der Waals contact distances and dihedral angles. These were mapped onto the target structure based on the alignment. The sequence of the AF2 receptor was aligned to that of the templates in order to look for amino acids that form the seven transmembrane α-helices. Using MODELLER 9.0 Version the model was complete with the helices and loops. Five models were generated and the best of the models based on their DOPE scores was selected as shown in **Figure 4.5** (Min-Yi Shen and Sali, 2006). The homology models obtained were based on the X-ray structures of beta-2 adrenergic and adenosine A_{2A} receptors. From the prediction results the AF2R was divided into 7 transmembrane Helices, 6 loops, and N and C- Terminus segments. The C-terminus of the AF2 receptor was extremely long compared to other rhodopsin-like GPCRs. It was however separated from the entire structure i.e. from valine 311 to cysteine 432 amino acid to shorten the GPCR for easy dynamics. Therefore the model was found to have some structural features in the transmembrane domain.

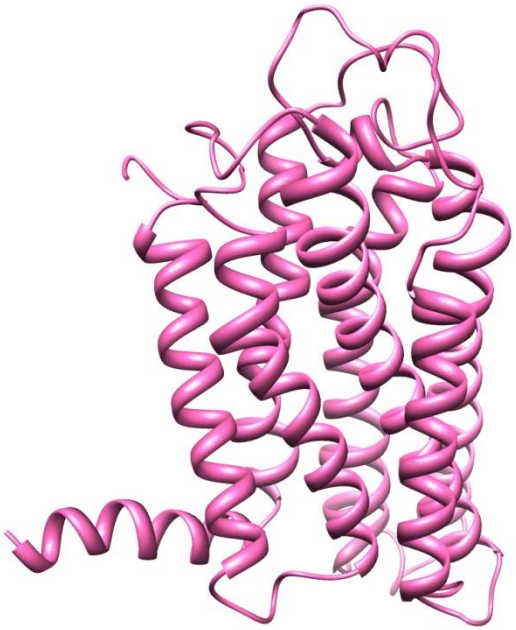


Figure 4.5: The ribbon representation of the AF2R model without the C-terminus

The final model was evaluated by use of Ramachandran's plot. As shown in **figure 4.6** it was found that most of the helical amino acids located in the region favouring a right- hand α-helix. Ramachandran plot shows that 95% of the AF2 receptor model has psi and phi angles in the favoured region hence the model is acceptable. Only 5% of all residues were found in the sterically disallowed region in the model and these residues are located in the loop domains thus not affecting the main region of the model and they don't have to be corrected to fit in the allowed region.

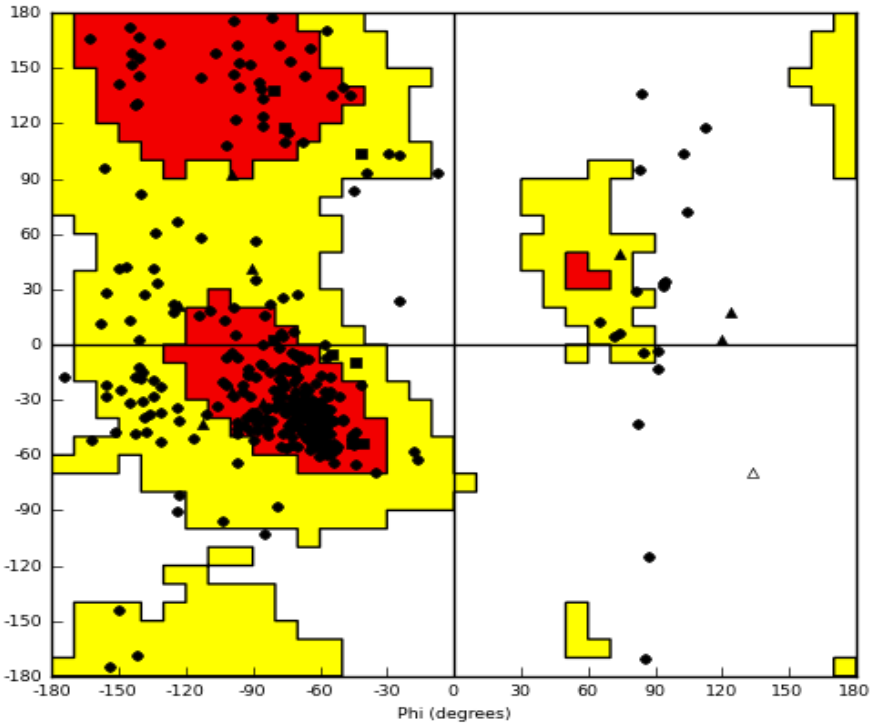


Figure 4.6: 2D graphical representations of the Ramachandran plot showing phi and psi angles of AF2 receptor. The most favoured regions are indicated as red, generously allowed are indicated as yellow and disallowed region indicated as white fields.

The AF2 receptor model was overlaid (**Figure 4.7**) with those of beta-2 and adenosine 3 D structures and gave a root mean square deviation (RMSD) of 0.819 A° and 1.046 Aº respectively.

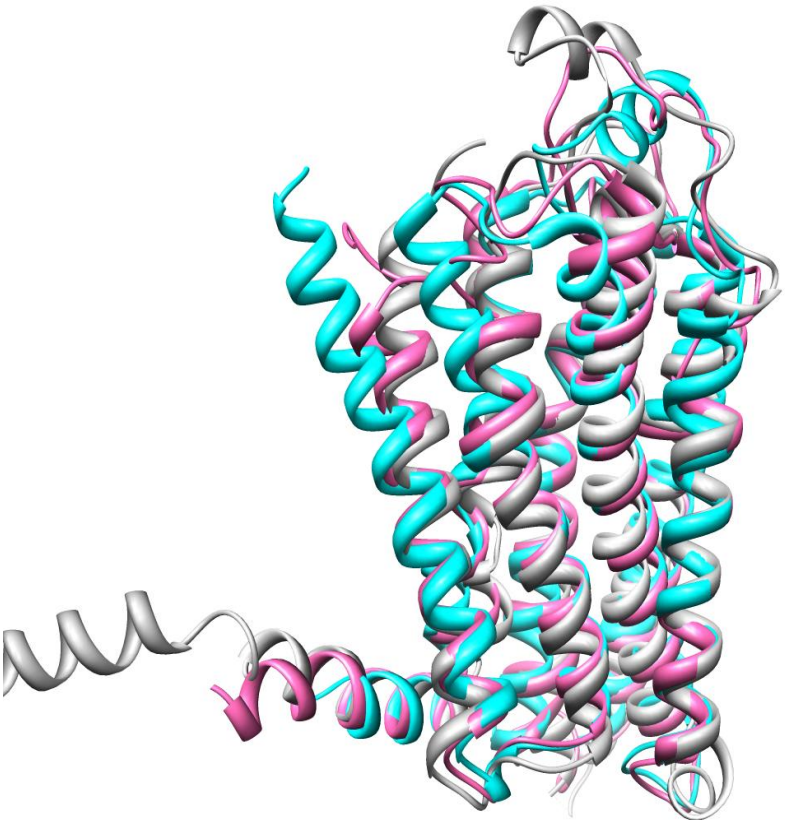


Figure 4.7: An overlay of AF2R model, beta-2 adrenergic and Adenosine A_{2A} receptors (Pink = AF2R, Grey =Adenosine A_{2A}receptor and cyan = Beta-2 adrenergic receptor).

From the superimposition of the AF2 model over the Adenosine A_{2A} and Beta-2 adrenergic receptor templates the R.M.S.D are relatively low indicating a valid structure for the whole model. This demonstrates that the model was reasonable and could be used for further MD simulations studies.

4.2.3 Molecular Dynamics Simulation of the AF2 Receptor

In vacuo Molecular Dynamic simulations were carried out on the peptide using GROMACS software with frozen helices. The structure was not subjected to any restraints. The structure of the AF2 receptor model obtained from the MD simulation was similar to the experimentally determined structure in terms of the 7 transmembrane helices orientations. The validation of the receptor model is based on the assumption that GPCRs share significant structural commonalities and the high helical content of the initial model structure correlates with the presence of 7 alpha helices in the transmembrane regions which incorporate non- helical deviations by proline residues. During the simulations, the alpha helical content of the AF2 receptor decreased marginally and this reflects the balance between the stability of the helices and the structural dynamics during the simulation. However, there were neither beta sheets nor disulphide bonds noticed in the extracellular loops which are common features of most rhodopsin-like GPCRs.

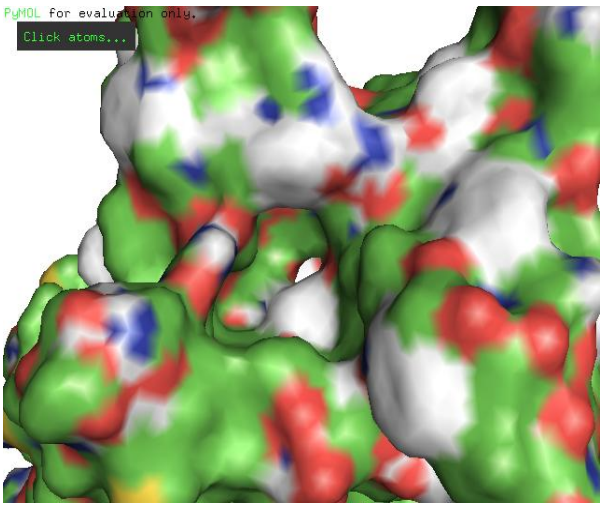
During MD simulations on the AF2R, one hundred (100) structures were collected. Cluster analysis was done and 4 clusters were identified differing in the number of conformations with a cut-off point of 0.2nm. Cluster 3 had the largest number of conformations with 98 structures falling under it. The structure with the lowest energy was used as the starting structure for molecular docking studies.

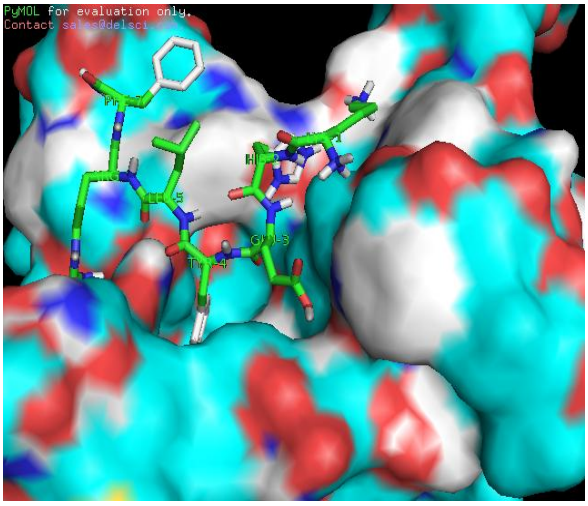
4.2.4 Molecular Docking

The AF2 peptide was docked into the homology model structure of AF2 receptor using Autodock Vina software and the best receptor-ligand pose was inspected for close intermolecular interactions of the binding residues. After molecular docking the residues of AF2 peptide were

found docked into the residues of the AF2 receptor which comprises of Gly^{5}, Pro ^{64}, Leu ^{66}, Leu ^{68}, Pro ^{167}, Pro ^{168}, Arg ^{171}, Ser ^{261}, Arg ^{263}, Tyr^{267}, Trp ^{268}, Asp ^{271} and Leu ^{272}. The AF2 receptor constituted the N- Terminus, extracellular loops and Helix 7 residues. The hydrophobic residues of the peptide, i.e. Tyr and Phe made van der Waals contacts with residues in the extracellular loop region as well as extracellular ends of N- Terminus and TM 7. The side chain hydroxyl group of Tyr ^{4} amino acid residue of AF2 peptide formed a hydrogen bond with the side chain amino acid group of Arg ^{171} of EL2. The side chain aromatic ring of the seventh Phenylalanine of the AF2 peptide was found to be entangled within a hydrophobic cleft formed by Asp ^{166}, Tyr ^{160} and His ^{170} of the of EL2 of the AF2 receptor. The side chain amino group, NH_2 of Lys ^{1} of AF2 peptide was assumed to form a salt bridge with the side chain carboxyl group of Asp ^{166} of the receptor, which contributes to additional stability of the peptide into the binding site. Helix 7 was particularly found to interact with the N- terminus of the peptide. Internal interactions within the receptor were also observed with Gly ^{264} interacting with His ^{170} and Asp ^{166} interacting with both Gly ^{168} and Tyr ^{160} amino acid residues.

The binding pocket for the AF2 receptor was found to be open to the extracellular region. This is similar with the beta-2 based AKHR whose binding site is easily accessible. However, the rhodopsin-based AKHR was found to be closed and could not allow free access to the binding site. (Mugumbate et al, 2010). However, researchers have found that the binding site of both rhodopsin and beta-2 adrenergic receptors are buried inside the intracellular region and does not allow free diffusion of ligands into the binding site.





(A) **(B)**

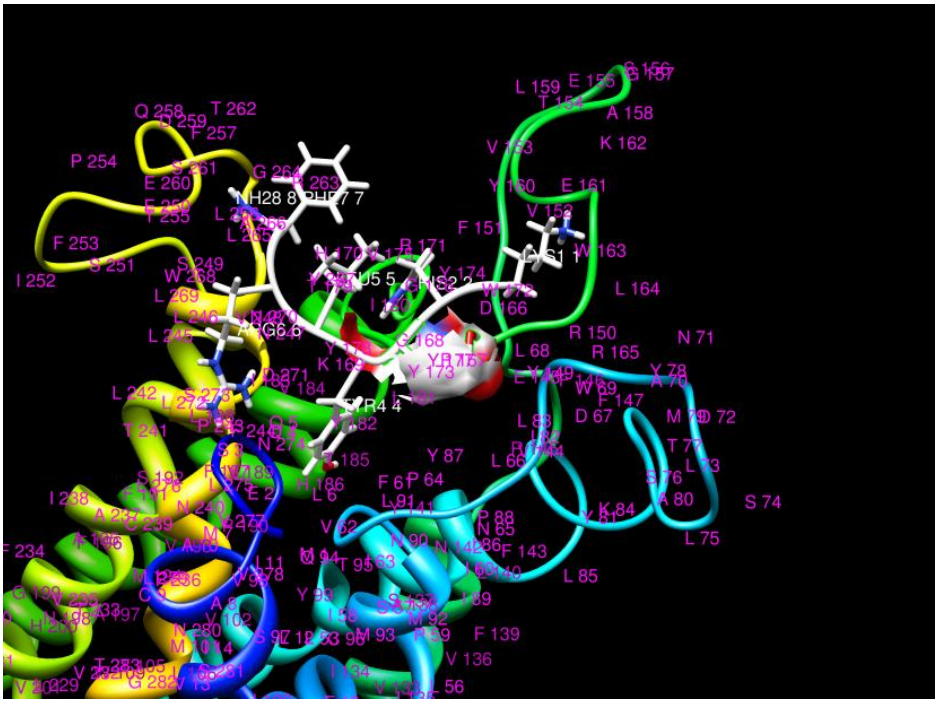


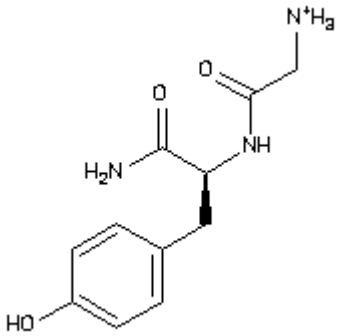
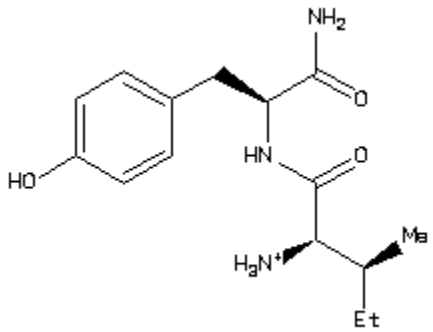
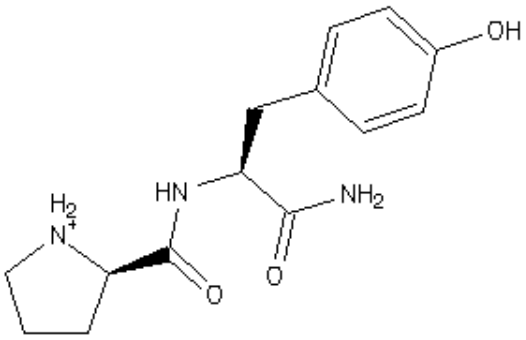
Figure 4.8: The surface representation of the AF2 receptor showing (A) the binding site, (B) AF2 peptide inside the receptor's binding pocket, ΔG (-7.9 kcal mol^{-1}) and (C) AF2 interactions with the receptor.

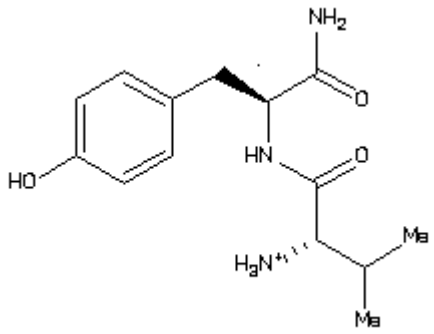
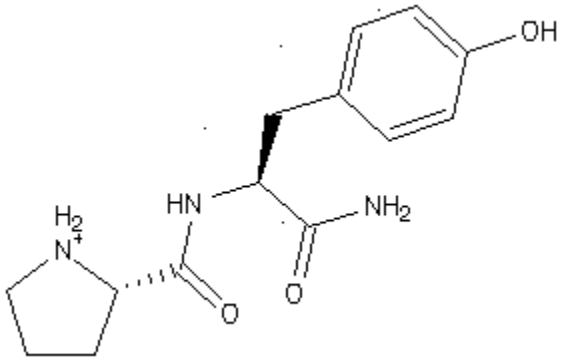
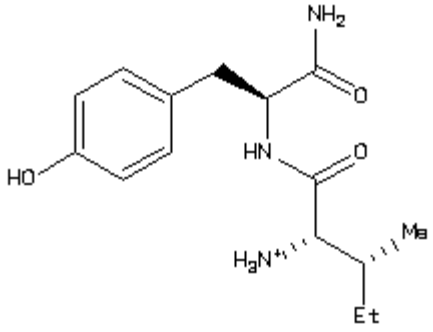
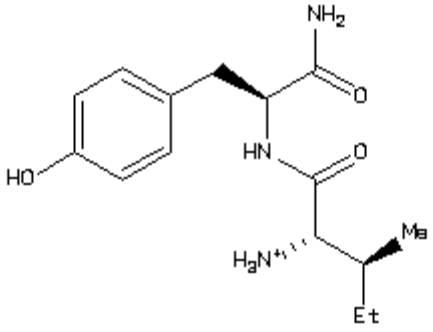
4.2.5 Structure- based Virtual (*in silico*) Screening

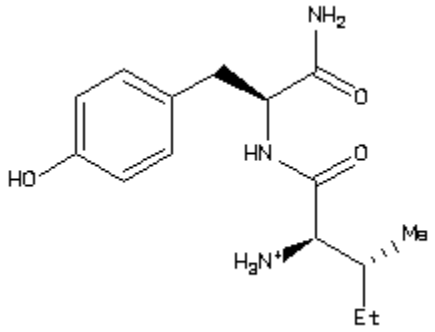
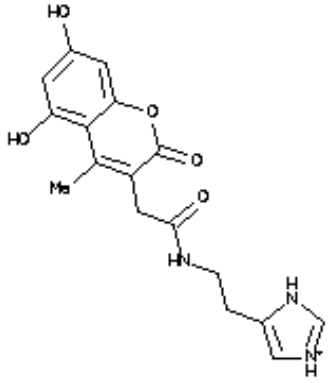
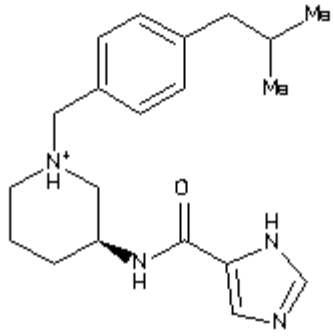
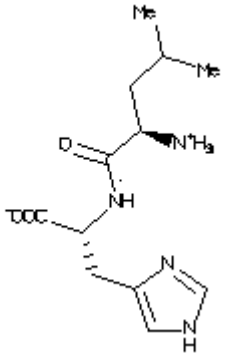
The generated AF2 receptor homology model was evaluated for its ability to dock various ZINC compounds in virtual screening experiment. The top 35 compounds were selected based on the free energy of binding. All the compounds were found to dock inside the receptor's cavity with reasonable binding energies. All the compounds were found to freely diffuse in the binding site; this was attributed to the open conformation of the binding site to the extracellular region.

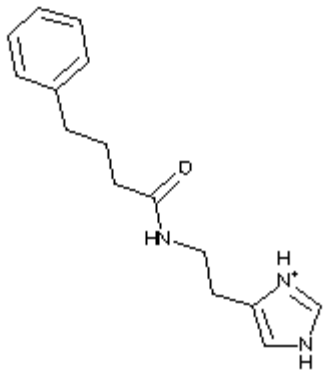
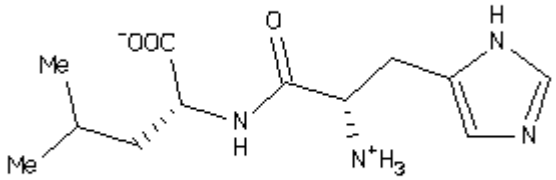
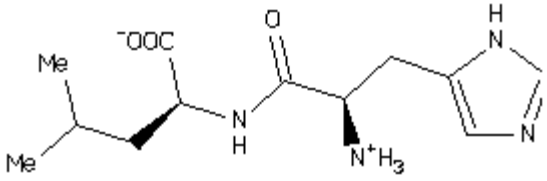
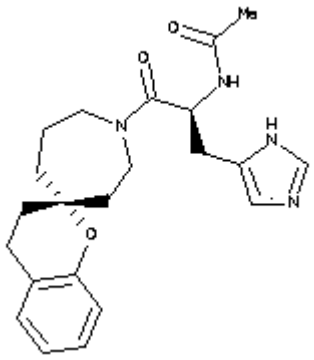
The following are the representative hits with the AF2 receptor model and the predicted free energy of binding (ΔG) in kcal/mol.

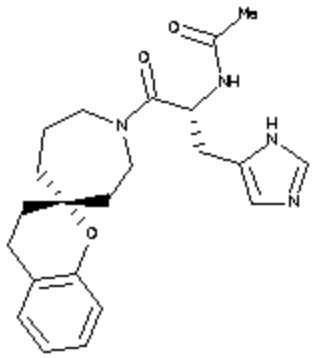
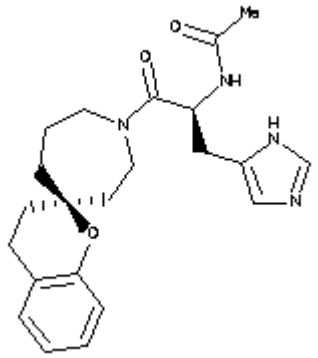
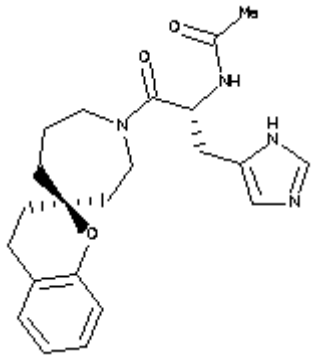
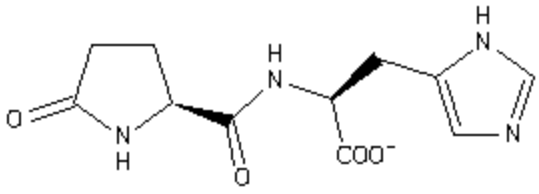
Table 4.4: Representative ZINC hits and their predicted free binding energy (kcal mol^{-1})

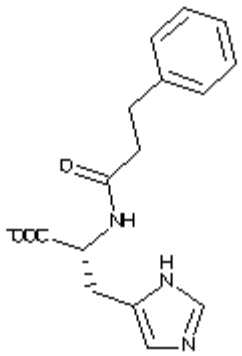
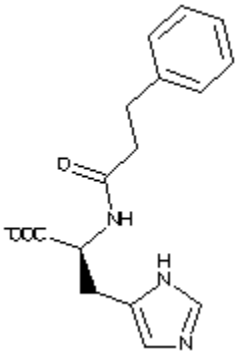
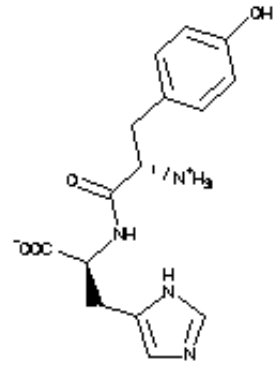
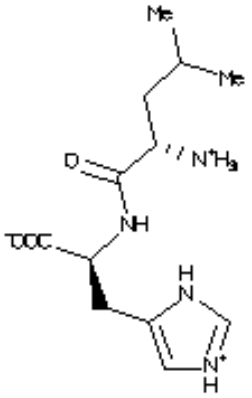
No.	Label (database)	Structure	Calculated ΔG (kcal mol^{-1})
1	ZINC02036139		-8.0
2	ZINC02172627		-7.9
3	ZINC02508274		-7.8

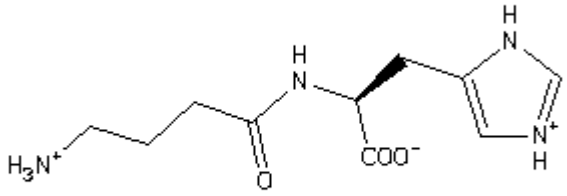
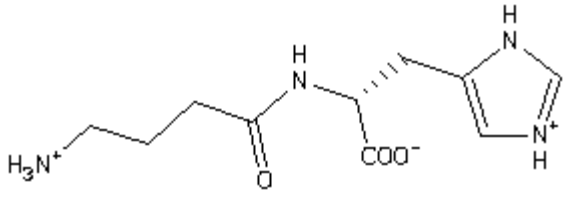
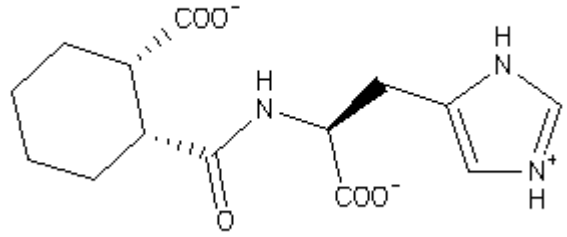
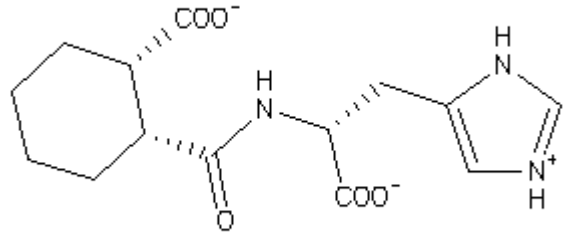
4	ZINC02561145		-7.7
5	ZINC04762808		-7.7
6	ZINC04763176		-7.7
7	ZINC04763179		-7.7

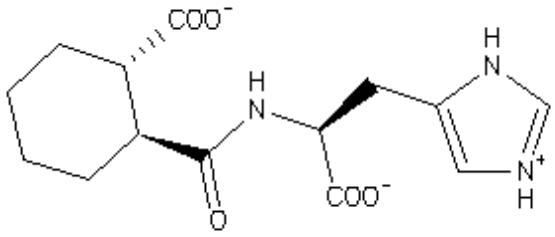
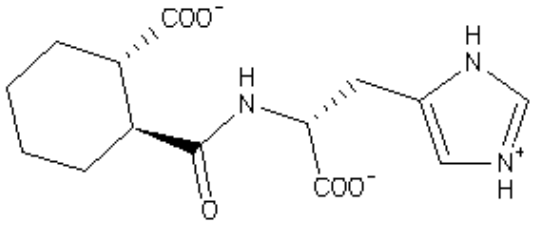
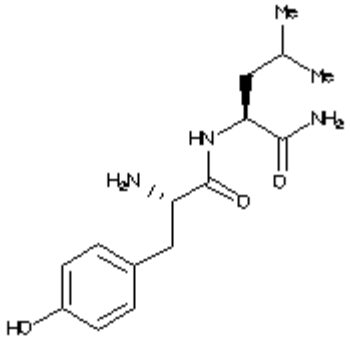
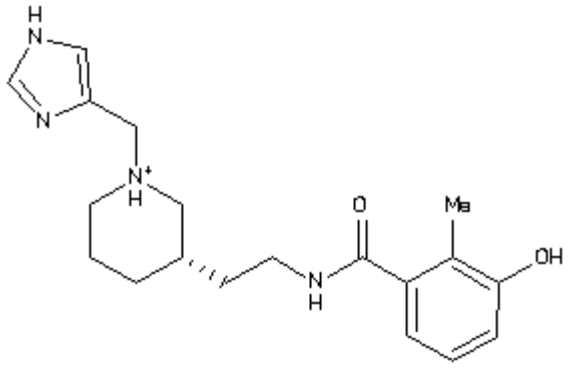
8	ZINC04763182		-7.7
9	ZINC08791907		-7.6
10	ZINC12302910		-7.6
11	ZINC13212372		-7.6

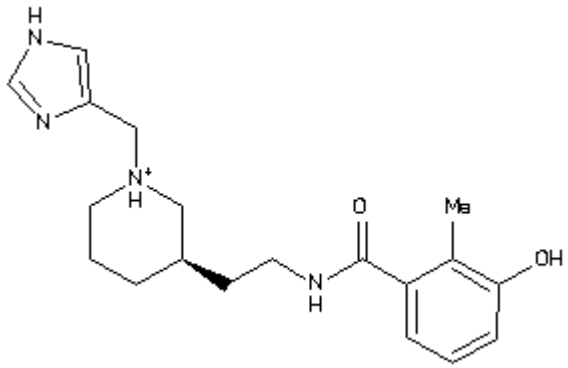
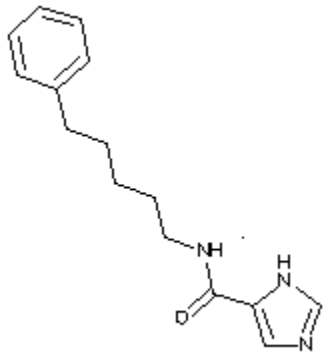
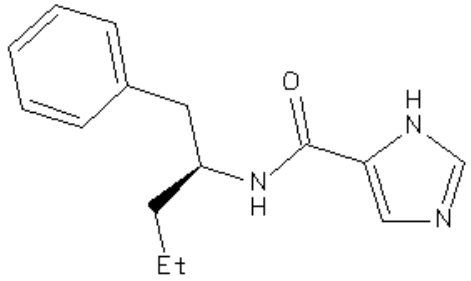
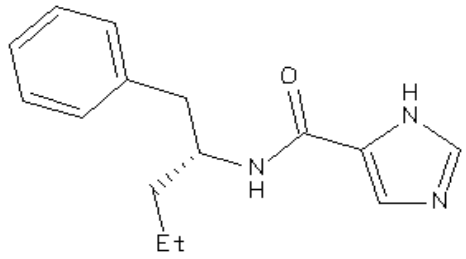
12	ZINC13591247		-7.5
13	ZINC19014734		-7.5
14	ZINC19014742		-7.5
15	ZINC19020120		-7.5

16	ZINC19020123		-7.5
17	ZINC19020126		-7.5
18	ZINC19020129		-7.4
19	ZINC19505337		-7.4

20	ZINC22222657		-7.4
21	ZINC22222658		-7.4
22	ZINC34349844		-7.3
23	ZINC38200487		-7.3

24	ZINC38418187		-7.3
25	ZINC38606076		-7.3
26	ZINC43755172		-7.3
27	ZINC43755173		-7.2

28	ZINC43755174		-7.2
29	ZINC43755176		-7.2
30	ZINC59968896		-7.1
31	ZINC67965450		-7.1

32	ZINC67965452		-7.0
33	ZINC69489001		-7.0
34	ZINC69540787		-7.0
35	ZINC69540788		-7.0

Comparing the molecular weight of the peptide and the ZINC compounds, the ZINC compounds have lower molecular weight. This was because of the ‘Lipinski's rule of five’ applied in virtual screening. The binding position of the ZINC compounds was found to similar to all, as they were

buried inside the cavity. This was similar to the AF2 peptide. The calculated binding energies of the ZINC compounds are almost the same. This may be attributed to the fact that the compounds are similar in structure. They are all amides and most of them have hydroxyl functional groups and an imidazole ring in their structures. Others are skeletal isomers of each other for example structures 2 and 4; and 6, 7 and 8. Imidazole is a polar and ionisable aromatic compound which improves the pharmacokinetic characteristics of lead compounds thus making the ZINC compounds potential hits.

5.0 CONCLUSION AND RECOMMENDATIONS.

5.1 CONCLUSION

In this thesis, Molecular dynamic simulations were carried out on the AF2 peptide. On cluster analysis only one cluster was found which meant that the peptide structure was highly conserved. The peptide had a closed conformation before docking but opened up on performing docking studies.

The available primary sequence of nematode AF2 receptor was used to predict the 3-D structure of the receptor based on the Adenosine A_{2A} and beta2-adrenergic receptors. The resulting conformation and structural features were found to be similar to the rhodopsin family receptors because of the existence of conserved residues in the transmembrane regions. However, the C- terminus was found to be extremely lengthy and N-terminus short compared to the rhodopsin family receptors. Successful homology modelling was carried out and an accepted AF2 model was generated though in this receptor the disulfide bond found between TM3 and EC2 was missing. The AF2 receptor was found to have an exposed binding site. The AF2 peptide was buried inside the receptor cavity. There was no observable hydrogen bonding in the peptide, whereas the docked structure had both hydrogen and hydrophobic interactions between the receptor and the peptide. The docked ZINC compounds were also found to dock inside the binding site of the receptor. The ZINC compounds can be used to design specific anthelminthes and this will go a long way in controlling nematodes and addressing poverty in developing countries.

5.2 RECOMMENDATIONS

I would like to recommend that molecular dynamic simulations of the AF2 G-protein coupled receptor be performed in water and in membrane and docking studies to be performed using the ZINC compounds above. The ZINC compounds should also be tested for their activity in order to design specific anthelminthes to be used in control of parasitic nematodes.

6.0 BIBLIOGRAPHY

Accelrys Software Inc., (2012) *Discovery Studio Modelling Environment, Release 3.1,* San Diego: Accelrys Software Inc.

Albonico, M., Stoltzfus R.J., Savioli L., Tielsch J.M , Chwaya H.M, Ercole E, (1998). Epidemiological evidence for a differential effect of hookworm species, Ancylostoma duodenale or Necator americanus, on iron status of children. *Int J Epidemiol;* 27(3):530-537.

Altschul, S.F., Gish, W., Miller, W., Myers E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol*, 215, 403-410.

Anosike, J.C., Nwoke B. E., (2005). Lymphatic filariasis among the Ezza people of Ebonyi State, eastern Nigeria. *Ann Agric Environ Med* 12(2): 181-6.

Baldwin, J. M. (1994). Structure and function of receptors coupled to G proteins.

Berman, H.M., Bhat, T.N., Bourne, P.E., Feng, Z., Gilliland ,G., Weissig, H.(2000). The protein data bank and the challenge of structural genomics. *Nat Struct Biol* 7:957–9.

Bethony, J., Brooker, S., Albonico, M. (2006). Soil-transmitted helminth infections. ascariasis, trichuriasis, and hookworm. *Lancet* 367: 1521-32.

Bissantz, C., Logcan, A., and Rognan, D. (2004). High-Throughput Modelling of Human G-Protein Coupled Receptors: Amino acid Sequence-alignment, Three Model building and Receptor Library Screening, *J. Chem,. Inf. Comput. Sci.* 44, 1162-1176.

Bockaert, J. and Pin, J. P. (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo J.* 18, 1723-1729.

Bowman, J.W, Friedman, A.R, Thompson, D.P, Ichhpurani, A.K, Kellman, AF, Marks, N, Maule, A.G, Geary, T.G. (1996). Structure-activity-relationships of KNEFIRFNH_2 (AF1), a nematode FMRFamide-related peptide, on *Ascaris suum* muscle. *Peptides.*

Branden C.T.J. (1999). *Introduction to Protein Structure*. Garland Publishing, Inc., New York.

Brown, H.D., Matzuk, A.R., Ilves, I.R., Peterson, L.H., Harris, S.A., Sarett, L.H., Egerton, J.R., Yakstis, J.J., Campbell, W.C., Cuckler, A.C., 1961. Antiparasitic drugs. IV. 2-(40-Thiazolyl)-benzimidazole, a new anthelmintic. *J. Am. Chem*. Soc. 83, 1764–1765.

Brownlee, D.A., Holden-Dye. L., Fairweather. I., Walker, R,J., (1995). The action of serotonin and the nematode neuropeptide KSAYMRFamide on the pharyngeal muscle of the parasitic nematode, *Ascaris suum*. *Parasitol* 111:379– 384.

Buchan, D.W., Ward, S.M., Lobley, A.E., Nugent, T.C., Bryson, K. & Jones, D.T. (2010) Protein annotation and modelling servers at University College London. *Nucl. Acids Res*. 38 Suppl, W563-W568.

Chabala, J.C., Mrozik, H., Tolman, R.L., Eskola, P., Lusi, A., Peterson, L.H., Woods, M.F., Fisher, M.H., Campbell, W.C., Egerton, J.R., Ostlind, D.A. (1980). Ivermectin, a new broad-spectrum antiparasitic agent. *J. Med. Chem.* 23, 1134–1136.

Chamberlain, A.K. and Bowie, J.U. (2004). Analysis of side-chain rotamers in transmembrane proteins. *Biophysical Journal*, 87(5): p. 3460-3469.

Cherezov, V.; Rosenbaum, D.M.; Hanson, M.A.; Rasmussen, S.G.; Thian, F.S.; Kobilka, T.S.;Choi, H.J.; Kuhn, P.; Weis, W.I.; Kobilka, B.K. (2007). High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. *Science, 318,* 1258-1265.

Cooper, P.J, Chico, M.E., Rodrigues, L.C., Ordonez, M., Strachan, D., Griffin and G.E. (2003). Reduced risk of atopy among school-age children infected with geohelminth parasites in a rural area of the tropics. *J Allergy Clin Immunol*; 111(5):995-1000.

Costanzi, S., Siegel, J., Tikhonova, I. and Jacobson, K. (2009) Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr Pharm Des* 15:3994-4002.

Cowden, S. and Stretton A.O.W. (1993). AF2, an *Ascaris* neuropeptide: isolation, sequence and bioactivity. *Peptides*; 14:423–30.

Cowden, S., Stretton, A.O.W. and Davies, R.E. (1980). AF1, a sequenced bioactive neuropeptide isolated from the nematode Ascaris suum. Neuron 2: 1465-73.*Curr Opin Cell Biol*, 6, 180-190.

De Silva, N., Brooker, S., Hotez, P., Montresor, A., Engels, D. and Savioli, L. (2003). Soil-transmitted helminth infections: updating the global picture. Trends in Parasitology 12, 547–551.

de Silva, N.R., Brooker, S., Hotez, P.J., Montresor, A., Engels, D. and Savioli, L. Soil-transmitted helminth infections (2003). Updating the global picture. *Trends Parasitol* 2003; 19: 547-51.

Devaney, E. (2005). The end of the line for hookworm? An update on vaccine development. *PLoS Med*; 2(10):e327.

Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J (2004). Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. *Curr. Med. Chem. 11*, 71-90.

Ellen, Y.T.C.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G.W.; Hanson, M.A.; Shi, L.; Newman, A.H.;Javitch, J.A.; Cherezov, V.; Stevens, R.C.(2010). Structure of the human dopamine D3 receptor in complex with a D2/D3 selective agonist. *Science*, *330,* 1091-1095.

Evans, A.C., Stephenson, L.S. (1995). Not by drugs alone. the fight against parasitic helminths. *World Health Forum;* 16(3): 258-61.

Evers, A.; Klabunde, T. (2005). Structure-based drug discovery using GPCR homology modelling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.*, *48*, 1088-1097.

Evers, A.; Klebe, G. (2004). Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med.Chem. , 47*, 5381-5392.

Falcone, F.H., and Pritchard, D.I., (2005). Parasite role reversal: worms on trial. Trends Parasitol; 21(4):157-160.

Flohr,C., Quinnell, R.J. and Britton, J. (2009). Do helminth parasites protect against atopy and allergic disease? *Clin Exp Allergy*; 39(1):20-32.

Fredriksson, R. (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol*. 63, 1256–1272.

Gomez-Escobar, N., Lewis, E. (1998). A novel member of the transforming growth factor-beta (TGF-beta) superfamily from the filarial nematodes *Brugia malayi* and *B. pahangi*. *Exp Parasitol* 88(3): 200-9.

Guvench, O.; Mackerell, A. D.(2008). Jr. Comparison of protein force fields for molecular dynamics simulation. *Methods Mol. Biol., 443*, 63-88.

Halton, D.W, Shaw C, Maule AG, Smart D. Regulatory peptides in helminth parasites. *Adv Parasitol*.1994; 34:163–217.

Hawdon, J.M., Volk, S.W., Rose, R., Pritchard, D.I., Behnke, J.M., Schad, G.A. (1993). Observations on the feeding behaviour of parasitic third-stage hookworm larvae. *Parasitology*; 106 (Pt 2):163-169.

Hibert, M.F., Trumpp-Kallmeyer, S., Bruinvels, A.T. and Hoflack, J. (1991). 3D-models of cationic neurotransmitter G-protein coupled receptors. *Mol. Pharmacol*. 40, 8-15.

Higgins, D. G., Thompson, J. D. and Gibson, T. J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods Enzymol*, 266, 383-402.

Higgins, D.G., Thompson, J.D. and Gibson, T.J (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol*, 266: 383–402.

Hoerauf, A. (2006). New strategies to combat filariasis. Expert Rev Anti Infect Ther 4 (2): 211-22.

Horn F., Vriend G., and Cohen F. E. (2001). Collecting and harvesting biological data: the GPCRDB and NucleaRDB information systems. *Nucl. Acids Res.* 29, 346-349.

Hotez, P.J., and Pritchard, D.I. (1995). Hookworm infection. *Sci Am*; 272(6):68-74.

Hotez, P.J., Bethony, J., Bottazzi, M.E., Brooker, S. and Buss, P. (2005). Hookworm: "the great infection of mankind". *PLoS Med*; 2(3):e67.

Howard, A.D., McAllister, G., Feighner, S.D., Liu, Q.Y., Nargund R.P., Van der Ploeg L.H. T., and Patchett A. A. (2001). Orphan G-protein-coupled receptors and natural ligand discovery. *Trends Pharmacol. Sci.* 22, 132-140.

Hubbard, R. (2011). Structure-based drug discovery and protein targets in the CNS. *Neuropharmacology, 60,* 7-23.

Irwin, Sterling, Mysinger, Bolstad and Coleman, (2012). *J. Chem. Inf. Model.*

Ishii J, Fukuda N, Tanaka T, Ogino C, Kondo A (2010) Protein-protein interactions and selection: yeast-based approaches that exploit guanine nucleotide-binding protein signaling. FEBS J 277: 1982–1995.

Jaakola, V.P.; Griffith, M.T.; Hanson, M.A.; Cherezov, V.; Chien, E.Y.T.; Lane, J. R.; IJzerman, A. P.; Stevens, R. C. (2008). The 2.6 Angstrom Crystal structure of a Human A2A adenosine receptor bound to an antagonist. *Science, 322*, 1211-1217.

Jasmer, D.P., Groverse, A., and Smant, G., (2003). Parasitic nematode interactions with mammals and plants. *Ann. Rev. Phytopathol*. 41, 245–270.

John, B., and Sali, A., (2003). Comparative protein structure modelling by iterative alignment, model building and model assessment, *Nucleic Acids Research*, 31, 3982-3992.

Jones D.T. (1999). GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *J. Mol. Biol*. 287: 797-815.

Jones D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol*. 292: 195-202.

Jones, D.T, Taylor, W.R., and Thornton, J.M. (1994). A Model Recognition Approach to the Prediction of All-Helical Membrane Protein Structure and Topology. *Biochem*. 33: 3038-3049.

Jorgensen, W.L., and Tirado-Rives, J. (1988). The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.*110 (6): 1657–1666.

Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*118 (45): 11225–11236.

Kaminsky, R., (2003). Drug resistance in nematodes: a paper tiger or a real problem? Curr. Opin. Infect. Dis. 16, 559–564.

Karmik, S.S.; Gogonea, C.; Patil, S.; Saad, Y.; and Takezako, T. (2003). Activation of G-protein-coupled receptors: a common molecular mechanism. *Trends Endocrinol. Metab., 14*, 431-437.

Keating CD, Holden-Dye L, Thorndyke MC, Williams RG, *Neuropeptide Actions in* C. elegans Mallett A, Walker RJ. 1995. The FMRFamide-like neuropeptideAF2 is present in the parasitic nematode *Haemonchu scontortus*. *Parasitol* 111:515–521.

Kobilka, B.K.; Deupi, X. (2007). Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci., 28,* 397-406.

Koski, K.G., and Scott, M.E., (2003). Gastrointestinal nematodes, trace elements and immunity. *J. Trace Elem. Exp. Med*. 16, 237–251.

Krieger, E., Nabuurs, S. B., and Vriend, G. (2003). Homology modelling. *Methods Biochem Anal*, 44, 509-523.

Kristiansen, K. (2004). Molecular mechanism of ligand binding, signalling, and regulation within the superfamily of G-protein-coupled receptors: molecular modelling and mutagenesis approaches to receptor structure and function. *Pharmac Therapeutics* 103, 21-80.

Lagerstrom, M. C., & Schioth, H. B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Reviews. Drug Discovery*, 7(4), 339–357.

Larsen, M., 1999. Biological control of helminths. *Int. J. Parasitol*. 29, 139–146.

Lazar, P.; Kim, S.; Lee, Y.; and Lee, K. W. (2010). Computational approach to ensure the stability of the favourable ATP binding site in *E. coli* Hfq. *Journal of Molecular Graphics and Modelling*, v. 29, n.4, 573-580.

Lee D. K. (1999) Discovery of a receptor related to the galanin receptors. *FEBSLett.* 446, 103-107.

Lengauer, T.; Lemmen, C.; Rarey, M.; Zimmermann, M. (2004). Novel technologies for virtual screening. *Drug Discov. Today*, *9*, 27-34.

Lipinski, C.A.; Lombardo, F., Dominy, B.W. and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv*. *Drug*. *Deliver*. *Rev*. **1997**, 23, 3-25

Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev*., *23*, 3-25.

Lobley, A., Sadowski, M.I. and Jones, D.T. (2009). pGenTHREADER and pDomTHREADER: New Methods For Improved Protein Fold Recognition and Superfamily Discrimination. *Bioinformatics*. 25, 1761-1767.

Lomize, A. L.; Pogozheva, I. D. and Mosberg, H. I. (1999). Structural organization of G-protein-coupled receptors. *Comput. Aided Mol. Des.* 1999, 13, 325-353.

Lundstrom, K. (2005). Structural genomics of GPCRs. *Trends Biotech*. 23(2), 103-108.

Lyne, P. D. (2002). Structure-based virtual screening: an overview. *Drug Discov. Today*, *7*, 1047-10455.

McGuffin LJ & Jones DT. (2003) Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics*, 19, 874-881.

Mestres, J. (2002). Virtual screening: a real screening complement to high-throughput screening. *Biochem. Soc. Trans*., *30*, 797-799.

Min-Yi Shen and A. Sali (2006). Statistical potential for assessment and prediction of protein structure. *Protein Science* 15, 2507-2524.

Mizadegan, T.; Benkö, G.; Filipek, S. and Palczewski, K. (2003). Sequence analysis of G-Protein-coupled receptors: similarities to rhodopsin. *Biochemistry*, *42*, 2759-2767.

Morris, G., M. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* 19(14) 1639-1662.

Mortimer, K., Brown, A., Feary, J., Jagger, C., Lewis, S., Antoniak M. (2006). Dose-ranging study for trials of therapeutic infection with Necator americanus in humans. *Am J Trop Med Hyg;* 75(5):914-920.

Mount. D. M. (2003). Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press, 2nd edition.

Mugumbate, G., Jackson, G.E., and David van der Spoel.(2010). Open conformation of adipokinetic hormone receptor from the malaria mosquito facilitates hormone binding. *Peptide*, 32: 553-559.

Murakami, M.; and Kouyama, T.(2008). Crystal structure of squid rhodopsin. *Nature*, *453*, 363-368.

Murray, J., and Manoury, B. (2005). Bm-CPI-2, a cystatin from Brugia malayi nematode parasites, differs from *Caenorhabditis elegans* cystatins in a specific site mediating inhibition of the antigen-processing enzyme AEP. *Mol Biochem Parasitol* 139(2): 197-203.

Nelson, L.S., Kim, K.Y., Memmott, R.E. and Li, C. (1998). FMRFamide-related gene family in the nematode, *Caenorhabditis elegans*. *Mol Brain Res* 58:103–111.

Nelson, L.S.,and Li, C. (1996). *Flp-1* deletion mutants suggest FMRF-amide like peptides are involved in multiple sensory and motor functions in *Caenorhabditis elegans*. *Soc Neurosci Abstr* 22:333.

Nugent, T. & Jones, D.T. (2009). Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics*. 10, 159. Epub.

Nygaard, R.; Frimurer, T.M.; Holst, B.; Rosenkilde, M.M.; Schwartz, T.W. (2009). Ligand binding and micro-switched in 7TM receptor structures. *Trends Pharmacol. Sci., 30*, 249-259.

Ohno, K.; Mitsui, T.; Tanida,Y.; Matsuura, A.; Fujitani, H.; Niimi T.; Orita. M. (2011). Docking study and binding free energy calculation of poly (ADP-ribose) polymerase inhibitors. Journal of *Molecular Modelling*, v. 17, n.2, p. 383-389.

Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lanctot, J. K.;Putta, S.; Stanton, R.; Grootenhuis, P. D. (2004). Performance of 3D-database molecular docking studies into homology models. *J. Med. Chem.*, *47*, 764-767.

Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lanctot, J. K.; Putta, S.; Stanton, R.; Grootenhuis, P. D. (2004). Performance of 3D-database molecular docking studies into homology models. *J. Med. Chem.*, *47*, 764-767.

Overington, J.P, Al-Lazikani, B., and Hopkins, A.L. (2006). How many drug targets are there? *Nature Rev. Drug Discov., 5*, 993-996.

Palcezewski, K. (2000). Crystal structure of rhodopsin: A G- protein coupled receptor. *Science*, 289 (5480), 739 – 45.

Palcezewski, K. (2000). Crystal structure of rhodopsin: A G- protein coupled receptor. *Science* 289 (5480), 739 – 45 (2000).

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C., Motoshima, H., Fox, B., Trong, I., Teller, D., Okada, T., Stenkamp, R., Yamamoto, M. and Miyano, M. (2000). Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* 289, 739-745.

Pearson, W.R. and Lipman, D.J. (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci* U S A, 85(8):2444–2448.

Pearson, W.R., and Lipman, D.J., (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci* U. S. A., 85, 2444-2448.

Pierce, K. L., Premont, R. T. and Lefkowitz, R. J. (2002). *Nat. Rev. Mol. CellBiol.* 3, 639–650.

Pin, J.P., Galvez, T., and L. Prezeau, L. (2003). Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. *Pharmacol Ther*, .98(3): p. 325-54.

Price, D.A., Greenberg, M.J. (1977). Structure of a molluscan cardioexcitatory neuropeptide. *Science* 197: 670-671.

Pritchard, D.I., Quinnell, R.J., Walsh, E.A. (1995). Immunity in humans to Necator americanus: IgE, parasite weight and fecundity. *Parasite Immunol*; 17(2):71-75.

Quinnell, R.J., Pritchard, D.I., Raiko, A., Brown, A.P., Shaw, M.A. (2004). Immune responses in human necatoriasis: association between interleukin-5 responses and resistance to reinfection. *J Infect Dis*; 190(3):430-438.

Rasmussen, S.G.; Choi, H.J.; Rosenbaum, D.M.; Kobilka, T.S.; Thian, F.S.; Edwards, P.C.;Burghammer, M.; Ratnala, V.R.; Sanishvili, R.; Fischetti, R.F.; (2007). Crystal structure of the human β_2-adrenergic G-protein-coupled receptor. *Nature*, *450*, 383-387.

Ravna A. W. (2006). Three-dimensional models of neurotransmitter transporters and their interactions with cocaine and S-citalopram. *World J Biol Psychiatry* 7, 99-109.

Reddy, A., and Fried, B. (2007). The use of Trichuris suis and other helminth therapies to treat Crohn's disease. *Parasitol Res*; 100(5):921-927.

Reddy, A., and Fried, B. (2009). An update on the use of helminths to treat Crohn's and other auto immune diseases. *Parasitol Res*; 104(2):217-221.

Risher, J.F., Mink, F.L., Stara, J.F., (1987). The toxicologic effects of the carbamate insecticide aldicarb in mammals: a review. *Environ. Health Perspect*. 72, 267–281.

Roos, M.H., (1997). The role of drugs in the control of parasitic nematode infections: must we do without? *Parasitology* 114 (Suppl.), S137–S144.

Sali, A and Blundell, T.L. (1993). Comparative Protein Modelling by satisfaction of spartial restrains. *J. Mol. Biol*. 234, 779-815,.

Sangster, N. C. (2001). Managing parasiticide resistance. *Veterinary Parasitology* 98, 89–109.

Schneider, S.M., Rosskopf, E.N., Leesch, J.G., Chellemi, D.O., Bull, C.T., Mazzola, M., (2003). United States Department of Agriculture– Agricultural Research Service research on alternatives to methyl bromide: pre-plant and post-harvest. *Pest Manag. Sci*. 59, 814–826.

Schneidman-Duhovny, D.; Nussinov, R.; Wolfson, H.J. (2004). Predicting molecular interactions in silico: II. Protein-protein and protein-drug docking. *Curr. Med. Chem*, *11*, 91-107.

Schoneberg, T., Schulz, A. and Gudermann, T. (2002). The structural basis of G protein-coupled receptor function and dysfunction in human diseases. *Rev. Phys.Biochem. Pharm.* 144, 145-227.

Schöneberg, T., Schulz, A., and Gudermann, T. (2002). The Structural Basis of G Protein-Coupled Receptor Function and Dysfunction in Human Diseases. *Reviews of Physiology, Biochemistry and Pharmacology 144*:143-227.

Shimamura, T.; Hiraki, K.; Takahashi, N.; Hori, T.; Ago, H.; Masuda, K.; Takao, K.; Ishiguro,M.; and Miyano, M. (2008). Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *J. Biol. Chem*., 283, 17753-17756.

Sousa S.F., Fernandes P.A. and Ramos M.J. (2006) Protein-ligand docking: current status and future challenges. *Proteins* 65, 15-26.

Spratt, D.M., (1997). Endoparasite control strategies: implications for biodiversity of native fauna. Int. *J. Parasitol*. 27, 173–180.

Stahura, F.L.; Bajorath, J.(2005). Novel methodologies for virtual screening. *Curr. Pharm Design., 11*, 1189-1202.

Stepek, G., Buttle, D.J., Duce, I.R., Behnke, J.M., (2006). Human gastrointestinal nematode infections: are new control methods required? *Int. J. Exp. Pathol*. 87, 325–341.

Strader, C.D., Fong, T. M., Tota, M.R., Underwood, D. and Dixon, R.A.F., (1994). Structure and Function of G-Protein Coupled Receptors. *Annu. Rev.Biochem.* 63, 101-132.

Strader, C.D., Sigal, IS., Register, R.B., Candelore, M.R., Rands, E. and Dixon, R.A.F. (1987). Identification of residues required for ligand binding to the P-adrenergic receptor. *Proc. Natl. Acad. Sci.* USA 84, 4384-4388.

Svoboda, P.; Teisinger, Crompton, D.W.T. (1999). How much human helminthiasis is there in the world? *Journal of parasitology*, 85, 397-403.

Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K. & Stenkamp, R. E. (2001). *Biochemistry* 40, 7761–7772.

Teller, D., Okada, T., Behnke, C., Palczewski, K. and Stenkamp, R. (2001). Advances in Determination of a High-Resolution Three-Dimensional Structure of Rhodopsin, A Model of G-Protein-Coupled Receptors (GPCRs). *Biochemistry* 40, 7761-7772.

Thanki, N.; Zeelen, J.P.; Mathieu, M.; Laenicke, R.; Abagyan, R.A.; Wierenga, R.K.; Schliebs, W. (1997). Protein engineering with monomeric triosephosphate isomerase (monoTIM): The modelling and structure verification of a seven residue loop. *Protein Eng*., 10, 159–167.

Trumpp-Kallmeyer, S., Hoflack, J., Bruinvels, A. and Hibert, M. (1992).Modelling of G-protein-coupled receptors: Application to dopamine, adrenaline, serotonin, acetylcholine and mammalian opsin receptors. *J. Med. Chem*. 35, 3448-3462.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J. (2005). "GROMACS: fast, flexible, and free". *J Comput Chem*26 (16): 1701–18.

Van der Spoel, D., Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Tieleman, D. P., Sijbers, A. L. T. M., Feenstra, K. A., vanDrunen, R., Berendsen, H. J. C. (2005). *Gromacs User Manual, Versions 3.3*. www.gromacs.org.

Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S.(2003). Molecular modelling of the three-dimensional structure of dopamine 3 (D3) subtype receptor: discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.*, *46*, 4377-4392.

Waghorn, T.S., Leathwick, D.M., Chen, L.Y., Skipp, R.A., (2003). Efficacy of the nematode-trapping fungus Duddingtonia flagrans against three species of gastro-intestinal nematodes in laboratory faecal cultures from sheep and goats. *Vet. Parasitol*. 118, 227–234.

Walker, R.J., Holden-Dye, L., and Franks, C.J. (1993). Physiological and pharmacological studies on annelid and nematode body wall muscle. *Comp Biochem Physiol*; 106C:49–58.

Wamae, C. N. (1994). Advances in the diagnosis of human lymphatic filariases: a review. *East Afr Med J* 71(3): 171-82.

Warne, A.; Serrano-Vega, M.J.; Baker, J.G.; Moukhametzianov, R.; Edwards, P.C.; Henderson,R.; Leslie, A.G.W.; Tate, C.G.; Schertler, G.F.X. (2008). Structure of the β1-adrenergic G protein-coupled receptor. *Nature*, *454*, 486-491.

Wieman H, Tøndel K, Anderssen E, Drabløs F (2004). Homology-based modelling of targets for rational drug design. *Mini Rev Med Chem* 4:793–804.

Williams, S. A., Lizotte-Waniewski, M. R. (2000). The filarial genome project analysis of the nuclear, mitochondrial and endosymbiont genomes of Brugia malayi. *Int J Parasitol* 30(4): 411-9.

Williamson A.R. (2000). Creating a structural genomics consortium. *Nature Struct. Biol.* Suppl. 953.

Wordemann, M., Diaz, R.J., Heredia, L.M., Collado Madurga, A.M., Ruiz Espinosa, A., and Prado, R.C., (2008). Association of atopy, asthma, allergic rhino conjunctivitis, atopic dermatitis and intestinal helminth infections in Cuban children. *Trop Med Int Health*; 13(2):180-186.

Wu, B.; Chien, E.Y.; Mol, C.D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.;Wells, P.; Bi, F.C.; (2010). Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*, *330*, 1066-107.