Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Computational Study

of the α_{1A}-adrenoceptor and its ligands

By: Gemma Kinsella

A thesis presented to Trinity College Dublin for the degree of Doctor of Philosophy

October 2005

Declaration

I hereby declare that:

This thesis has not been submitted as an exercise for a degree at this or any other University.

The work contained herein is entirely my own, except where otherwise cited, referenced, acknowledged or accredited.

I agree that the Library of the University of Dublin, may at their discretion lend or copy the thesis upon request.

.....G. Kinsella.....

Gemma Kinsella

October 2005

Trinity College Dublin

Thesis Summary

Adrenoceptors belong to the 'super-family' of seven transmembrane domain receptors known as G protein coupled receptors (GPCRs), which produce their effects through coupling with G proteins. Nine different subtypes of adrenoceptors (ARs) have been identified based on the results of pharmacological and molecular cloning studies, of which the α_1-AR subtypes are associated with the condition Benign Prostatic Hyperplasia (BPH). BPH effects 50 % of men over the age of 50 and due to increased life expectancy the proportion of men suffering from BPH will increase significantly over the coming years. Prostatic smooth muscle contraction occurs predominantly, if not exclusively, via the α_{1A}-AR subtype and antagonists selective for this particular receptor might be most useful in alleviating the urological symptoms of BPH which include increased frequency in urination, nocturia, a poor urine stream and hesitancy or delay in starting flow.

Rational structure-based drug design (SBDD) requires prior knowledge of the structure of the target receptor and the principal ligand binding site. Currently, there are few experimentally solved structures of membrane proteins due to the difficulty in obtaining sufficient quantities of the protein for crystallization trials. In this work, homology models of the α_{1A}-AR were developed based on an available bovine Rhodopsin crystal structure (*1l9h*). The prediction process consisted of fold assignment, target-template alignment, model building and evaluation. Structural refinement was achieved through molecular dynamics simulations (MD) of the receptor in a variety of solvent environments, including a water box and two biphasic systems ($H_2O/CHCl_3/H_2O$ and $H_2O/CCl_4/H_2O$) before selecting $H_2O/CHCl_3/H_2O$ as the membrane mimic for further studies. Structural analyses of the refined α_{1A}-AR models were pursued to examine the inactive conformation of this receptor while a control simulation was also performed on the template rhodopsin crystal structure.

Adverse side effects limit the use of the α_1-AR antagonists available in clinical practice. In a search for novel antagonists, our group had previously prepared a series of bis-imino-imidazolidinium and bis-guanidinium diphenyl derivatives as possible selective α_{1A}-AR antagonists. These compounds along with others found in the literature have been computationally investigated through proton affinity calculations and detailed conformational analyses. We have also examined the use of α_1-AR

pharmacophores in a form of *in silico* screening of a series of α_1 and α_{1A} adrenoceptor antagonists spanning different structural classes.

Furthermore, by means of different molecular modelling techniques, such as docking and molecular dynamics simulations, we have developed a computational protocol to study the interactions of the agonists, adrenaline and noradrenaline, with the developed α_{1A}-AR homology models. MD simulations were pursued to optimise the agonist/receptor complexes and the produced 'agonist-bound' receptor forms were used for subsequent mutating, redocking and crossdocking of the agonists. Such three dimensional (3D) models of the receptor, allowed us to study the α_{1A}-AR in great detail, improving our understanding of conformational changes that take place upon agonist binding to the receptor. An accumulation of evidence suggests that the activation of GPCRs is connected to the movement of TM-VI with respect to TM-III and structural changes to the helices were examined. The role of conserved residues, such as the DRY and NPxxY motifs, in the activation process have also been examined.

Following a similar computational protocol involving MD simulations of antagonist/α_{1A}-AR complexes, three initial 'antagonist-induced' receptor conformations were optimised. One complex with the non-selective antagonist, doxazosin, one with the α_{1A}-AR selective antagonist, tamsulosin and one with a guanidinium derivative, compound **6**. The results illustrate how a detailed map of the binding site pocket of α_{1A}-AR obtained through combined experimental mutagenesis and dynamic receptor modelling can lead to a fuller understanding of the molecular basis for the effects of α_{1A}-AR antagonists on the receptor. We examined the use of the 'antagonist-induced' receptor forms in an *in silico* screening of antagonists of different structural classes. Our study highlights the need to develop an 'antagonist-bound' receptor form for each structural class of antagonist to be studied.

Finally, comparative MD simulations were performed on the uncomplexed inactive α_{1A}-AR structure, the 'antagonist-bound' form and the 'agonist-bound' receptor forms developed in this work to examine the effect of various ligands on the receptor conformation. Modelling approaches such as those undertaken in this study bridge experimental gaps and yield better models and theories that will ultimately aid in the design of new drugs.

Acknowledgements

I would like to take this opportunity to acknowledge and thank a number of people who have helped me in so many ways over the past few years.

I would firstly like to thank my supervisors, Graeme and Isabel, for all their help and support during the last number of years. Additionally, I would like to thank Professor Corish of the Trinity Chemistry Department for allowing me to complete this Ph.D. and letting me use the facilities of the Department.

My thanks to the TCHPC staff for all their help, assistance and computer time on aegir and moloch and a special thanks to Audrey for all her organisational skills and assistance with getting to conferences and acquiring software. I would also like to thank Dr. Peter Oliver at RAL for access and assistance with Hrothgar, a 16 node Beowulf cluster and the Centre for Synthesis and Chemical Biology (CSCB) for access to their cluster.

To everyone in the Comp. Lab., especially Aron, Berry, Gary, James and Paul and the other chemistry postgrads for making it such an interesting place to work. My thanks to Hugh for putting up with me through the good and bad times and for all the Italian coffee shop trips and Cadburys chocolate. A special thanks to Clare, Elaina, Gary and Iain for all the distractions over the years and to Joe for getting me over the last few hurdles.

Finally, but not least, I would like to express my gratitude to my family, for putting up with me. This thesis could never have been written without their help and support.

Table of Contents:

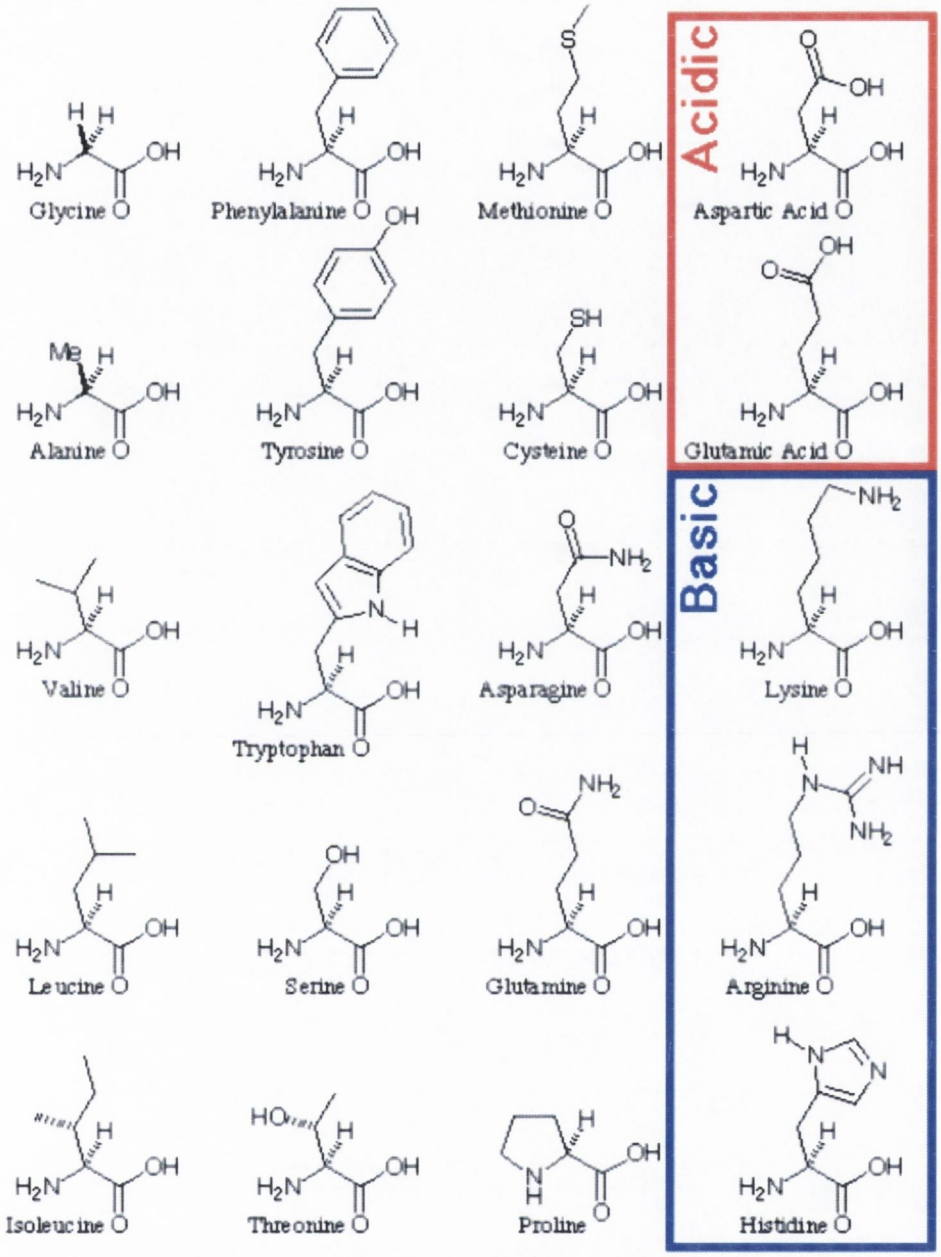
Abbreviations

3D	Three-Dimensional
5AR	5α-reductase
α_{1A}-AR	α_{1A} adrenoceptor
α_{1B}-AR	α_{1B} adrenoceptor
α_{1D}-AR	α_{1D} adrenoceptor
α_{2A}-AR	α_{2A} adrenoceptor
α_{2B}-AR	α_{2B} adrenoceptor
α_{2C}-AR	α_{2C} adrenoceptor
AD	Adrenaline
AM1	Austin Model 1
AMBER	Assisted Model Building with Energy Refinement
AO	Atomic orbital
β_1-AR	β_1 adrenoceptor
β_2-AR	β_2 adrenoceptor
β_3-AR	β_3 adrenoceptor
BA	Binding affinity
BE	Binding energy
BPH	Benign Prostatic Hyperplasia
CASP	Critical Assessment of Methods of Protein Structure Prediction
CG	Conjugate Gradient
CHARMM	Chemistry at HARvard using Molecular Mechanics
CHELP	Charges from Electrostatic Potentials
CHELPG	Charges from Electrostatic Potentials using a grid
DFT	Density Functional Theory
DHT	Dihydrotestosterone
DMPC	Dimyristoylphosphatidylcholine
DPPC	Dipalmitoylphosphatidylcholine
DSSP	Definition of Secondary Structure of Proteins
DZ	Double zeta
EC-I	1st extracellular loop
EC-II	2nd extracellular loop
EC-III	3rd extracellular loop
E/DRY	Glu/Asp-Arg-Tyr

ESP	Electrostatic Potential
FF	Force field
FFT	Fast Fourier Transform
GAFF	General AMBER Force Field
G-proteins	Guanine nucleotide binding proteins
GDP	Guanosine diphosphate
GPCR	G protein coupled receptors
GROMOS	GROningen Molecular Simulation Software
GTO	Gaussian Type Orbital
GTP	Guanosine triphosphate
HB	Hydrogen Bond
HF	Hartree Fock
IC-I	1st intracellular loop
IC-II	2nd intracellular loop
IC-III	3rd intracellular loop
IE	Interaction Energy
IR	Infrared
IP_3	Inositol triphosphate
LCAO	Linear Combination Atomic Orbital
LDA	Local Density Approximation
LUT	Lower Urinary Tract
MC	Monte Carlo
MD	Molecular Dynamics
MM	Molecular Mechanics
ND	Noradrenaline
NMR	Nuclear Magnetic Resonance
NPxxY	Asn-Pro-x-x-Tyr
PA	Proton Affinity
PASS	Putative Active Sites with Spheres
PBC	Periodic Boundary Conditions
PCM	Polarisable Continuum Model
POPC	Palmitoyoleoylphosphatidylcholine
PDB	Protein databank
PME	Particle Mesh Ewald
PKC	Protein kinase C

QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanical
QOL	Quality of life
QSAR	Quantitative Structural Activity Relationship
RMSD	Root Mean Square Deviation
SBDD	Structure based drug design
SCF	Self Consistent Field
SCR	Structurally Conserved Regions
SGI	Silicon graphics Inc.
SPC	Single point charge
STO	Slater type orbital
TM	Transmembrane
TIP3P	three point transferable intermolecular potential
TURP	Transurethral resection of the prostate
vdw	van der Waals
vHTS	virtual High Throughput Screening
VV	Velocity Verlet
ZDO	Zero Differential Overlap

Amino Acids


Glycine
Phenylalanine
Methionine
Acidic
Aspartic Acid
Alanine
Tyrosine
Cysteine
Glutamic Acid
Basic
Valine
Tryptophan
Asparagine
Lysine
Leucine
Serine
Glutamine
Arginine
Isoleucine
Threonine
Proline
Histidine

Chapter I: Introduction

"Art is the lie that helps tell the truth"

Pablo Picasso

1.0 Introduction

Commenting on the 1998 Nobel Prize awardees in quantum chemistry – Walter Kohn and John Pople – 'The Economist' wrote "In the real world, this could eventually mean that most chemical experiments are conducted inside the silicon of chips instead of in the glassware of laboratories."[1] An area in which such "*in silico*" experiments have been successfully applied is in the field of medicinal chemistry. Until recently, drugs that exert a therapeutic effect on the human body were discovered through random screening processes with an average design cycle for a drug of 6-12 years.[2] A more rational approach to drug design is necessary to meet market demand and modelling of ligand/protein interactions is now widely considered as a viable approach to accomplish this. Studying molecular structure and function through model building and computation has been aided in recent years by improvements in instrumental and experimental techniques, new models and algorithms for molecular simulations and increasing speeds of supercomputers.[2]

In the language of proteins, primary structure refers to the linear sequence of amino acids and as a consequence of the concerted effort in genomic research such as the Human Genome Project, a large number of human protein sequences have been determined. Protein secondary structure refers to the local ordered structure brought about via hydrogen bonding (HB) mainly within the peptide backbone structure (*e.g.* α helix or β strand). Tertiary structure refers to the global folding of a single polypeptide chain and finally quaternary structure involves the clustering of several individual peptide or protein chains into a final specific shape. The first reported experimental 3D structure of a protein was that of myoglobin, described in 1958 by J.C. Kendrew, (Nobel Prize 1962)[3] and currently the Protein Data Bank (PDB)[4] holds 32,727, experimentally, determined structures (20th September 2005).

Unfortunately, protein structure determination using experimental methods such as X-ray crystallography or Nuclear Magnetic Resonance (NMR) spectroscopy is time consuming and not possible for many proteins, particularly those embedded in a cell membrane. The ongoing challenge in protein crystallisation was highlighted by the awarding of the 2003 Nobel Prize to Peter Agre for discovering and characterizing the first water channel protein, aquaporin, and to Roderick MacKinnon who elucidated the structural and mechanistic basis for ion channel function.[5] As a consequence the number

of known protein sequences grows much faster than the number of 3D structures solved thereby increasing the relevance of meaningful structural information that non-experimentalists can provide.[6]

Protein receptors form the sensing elements in the system of chemical communications that coordinate the function of all cells in the body.[7] The broad spectrum of signals sensed by G protein coupled receptors (GPCRs) makes them one of the most intriguing targets for pharmacological interventions. Especially for the pharmaceutical industry as approximately 52 % of all existing medicines act on a GPCR.[8] Modelling large biological polymers, such as GPCRs, is a truly multidisciplinary enterprise; the elucidation of the relationship between structure and function in GPCRs being made possible by the combination of computational modelling and molecular pharmacological experiments.

The aim of this thesis is to apply computational medicinal chemistry techniques to the modelling of a GPCR, the α_{1A} adrenoceptor (α_{1A}-AR) and the study of the agonist and antagonist interactions with this receptor towards improved treatment of the Benign Prostatic Hyperplasia (BPH) condition.

1.1 Benign Prostatic Hyperplasia

The prostate is a small, walnut sized and shaped gland, comprised mostly of muscle and soft tissues that wrap around the upper part of the urethra, which carries urine from the bladder out through the penis (Figure 1.1). The prostates key role is in producing a milky fluid that helps transport and protect sperm during ejaculation.

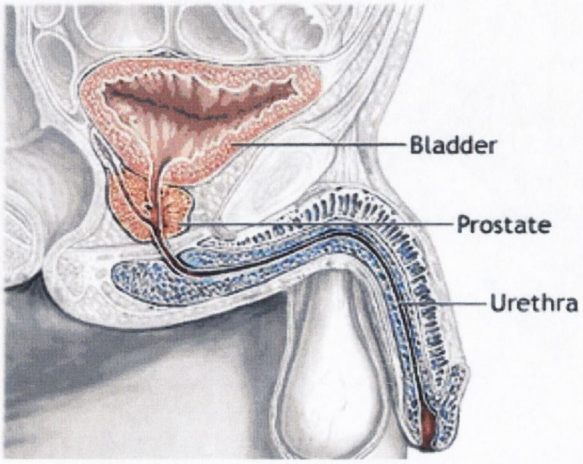


Figure 1.1: Image of the human bladder, prostrate and urethra.[9]

Benign prostate conditions, where the prostate has grown larger in size than expected normally, are prevalent in the male population. They include hypertrophy (growth by virtue of an increase in the size of cells) and hyperplasia (growth by virtue of an increase in the total number of cells resulting from mitosis). BPH is a condition whose prevalence in mature men occurs approximately linearly with age *e.g.* 50 % of men over the age of 50 and 90 % over 80 years.[10,11] Thus, the number of patients effected is rising worldwide as a result of the aging population.[12]

As the prostate grows, normal prostatic tissue becomes increasingly compressed by hyperplastic tissue and the urethra is impinged. This makes urinating difficult and leads to a variety of urological symptoms, that may vary in severity and impact to different degrees on a sufferers quality of life.[13,14,15] The symptoms are generally classified as *storage / irritative* (including urinary frequency, nocturia, burning, urgency or urge incontinence) or *voiding / obstructive* symptoms (including hesitancy, weak stream, slow termination/dribbling, sensation of incomplete voiding and urinary retention).[16,17] Although voiding symptoms occur with greater frequency, storage symptoms are considered to be more bothersome *i.e.* they interfere to a greater extent

with daily life activities and have a greater impact on the patient and his partner's quality of life.[18]

There are currently three modes of treatment for BPH, which are, watchful waiting, surgery or medication,[15] all of which have varying degrees of success and different side effects. In the first instance, watchful waiting, no active treatment is pursued but the situation is monitored by regular check-ups with a general practitioner. In the past, surgical procedures such as transurethral resection of the prostate (TURP), where the obstruction is surgically removed were considered the 'gold standard' in BPH treatment.[19] Open prostatectomy is a more complicated operation and is only usually recommended in the case of a much enlarged prostate.[11,12] The third option, medical therapy, became increasingly the preferred option over the early 1990s.[19] Souverin *et al.*,[19] reflected on this trend by examining the prevalence of BPH drug use and prostatic surgery in the period 1991-2000 among men aged 50 years or older. The prevalence of BPH drug use increased, while there was a statistically significant decrease in prostatic surgery in the same period.[11]

Clinical trials have shown that two main medical approaches are effective in reducing Lower Urinary Tract Symptoms (LUTS) suggestive of BPH and of improving urinary flow.[13,19] These are:

1) Inhibitors of the enzyme steroid 5α-reductase (5AR).
2) Treatment by high affinity α_1 adrenoceptor (α_1-AR) antagonists.

The growth and development of the prostate is under the influence of the male hormone testosterone and its more active metabolite dihydrotestosterone (DHT).[13] 5AR inhibitors block the conversion of DHT to testosterone and as a result shrink the prostate often up to 20-30 %.[20] However, BPH symptoms are found to correlate more with the extent of increase in smooth muscle tone rather than increase in gland size. α_1-AR antagonists act by reducing the smooth muscle tone in the bladder neck and the prostate.[19] Also, the onset of the α_1–AR antagonistic action is developed more rapidly (in the order of weeks) than the steroid 5AR inhibitors (in the order of months).[19] As approximately, 80 % of patients receiving medical therapy for BPH are prescribed an α-blocker by their general practitioner,[21] this approach shall be our main focus.

1.2 G Protein Coupled Receptors - Adrenoceptors

Adrenoceptors belong to Class A of the super family of GPCRs that receive a ligand stimulus at the extracellular side, transduce signals across a cell membrane and thus initiate a variety of intracellular biochemical events. In 1948, the adrenoceptors (AR) were initially divided into α- and β- AR[22] forms and subsequently subdivided according to sequence homology, drug specificity and mechanism of signal transduction to a final nine subtypes.[13,26]

α_1 (α_1-AR)	-	α_{1A} (α_{1A}-AR), α_{1B} (α_{1B}-AR), α_{1D} (α_{1D}-AR)
α_2 (α_2-AR)	-	α_{2A} (α_{2A}-AR), α_{2B} (α_{2B}-AR), α_{2C} (α_{2C}-AR)
β (β-AR)	-	β_1 (β_1-AR), β_2 (β_2-AR), β_3 (β_3-AR)

The adrenoceptors are activated upon binding of the catecholamines (adrenaline (AD) or noradrenaline (ND)) to the receptor (Figure 1.2). Both catecholamines contain an amino group protonated at physiological pH, separated from the aromatic catechol ring by a *beta* hydroxylethyl chain.

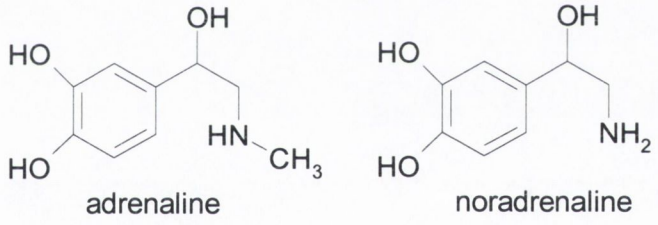


Figure 1.2: Adrenaline and noradrenaline - the catecholamine agonists for adrenoceptors.

Adrenaline is released from the adrenal medulla while noradrenaline is released from vesicles contained within the sympathetic nerve terminal. Noradrenaline diffuses across the synaptic gap to bind to numerous α_1-ARs located on the membrane of prostatic smooth muscle cells, inducing a conformational change in the adrenoceptor causing it to be activated. The signals are amplified by a family of proteins with homologous amino acid sequences, guanine nucleotide binding proteins, termed G proteins, which ultimately pass the orders to the final executors, enzymes or ion channels (Figure 1.3).[23]

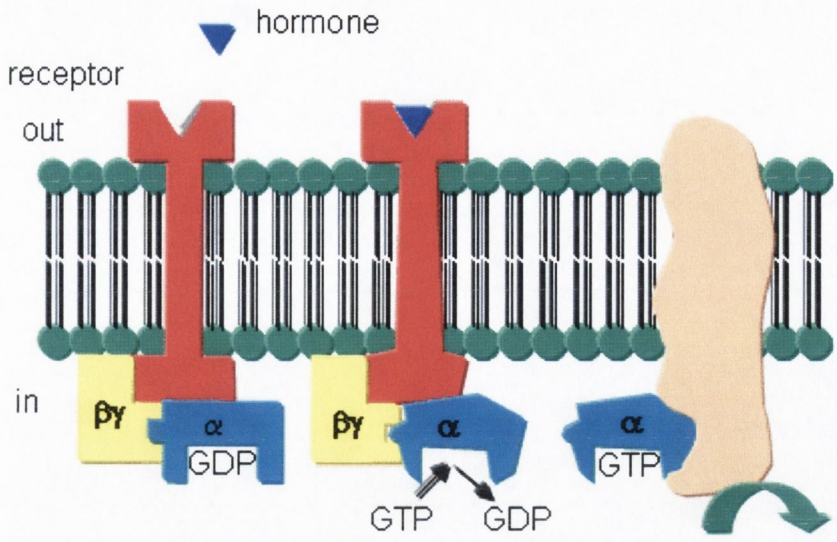


Figure 1.3: Generic scheme for activation of a G protein coupled receptor the G protein and the subsequent enzyme or ion channel.[24]

G-proteins consist of three subunits α, β and γ units, and activation causes the G protein to exchange nucleotide guanosine diphosphate (GDP) for nucleotide guanosine triphosphate (GTP). The G protein then dissociates to form an α and a $\beta\gamma$ unit, after which the GTP-bound α subunit diffuses along the membrane and binds to an effector, activating it. Subsequently, the α subunit converts GTP back to GDP, thereby inactivating itself before reassociating with a $\beta\gamma$ complex.[25] α_1-ARs couple predominantly through G-protein q (Gq), resulting in hydrolysis of membrane phospholipids to yield the second messengers, inositol triphosphate (IP_3) and diacylglycerol, leading to muscle contraction through mobilization of intracellular Ca^{2+} and activation of protein kinase C (PKC).[26] The resultant influx of calcium increases the prostatic smooth muscle tone.

1.2.1 Adrenoceptors in the prostate

The human prostate (Figure 1.4) contains mainly the α_{1A}-AR subtype (70 % of all α_1-AR present in the prostate).[27,28] Noradrenaline is equally potent with adrenaline at most of the adrenoceptor subtypes, although at some α_{1A}-AR sites it is considered to be more potent.[29] The catecholamine levels increase with age[23] inducing augmented contraction of human prostate tissue, resulting in an increased closure pressure on the urethra.[30] Antagonists improve urinary flow and voiding symptoms by blocking the action of α_{1A}-ARs, reducing contraction, which results in relaxation of smooth muscle of the bladder neck and prostate capsule. Antagonist inhibition at the α_{1A}-AR subtype is

thought to be most useful in alleviating obstructive symptoms in cases of BPH, while the blockade of α_{1D}-AR may also be important for patients with irritative symptoms by reducing bladder overactivity.[17]

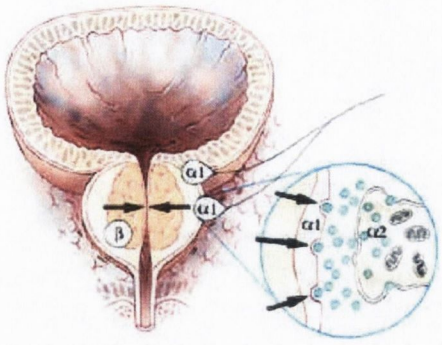


Figure 1.4: Illustration of the prostate with the locations of the α_1-AR indicated.[31]

Each adrenoceptor differs in structural details at the atomic level, so that the more perfectly a ligand fits one target the less well it fits another. In an effort to minimise side effects of the drug we are interested in selective binding to the α_{1A}-AR, the most abundant receptor in the prostate. We firstly need a structure for this α_{1A}-AR at the atomistic level, which unfortunately has eluded X-ray crystallography to date. Analysis of genome sequences has revealed that membrane proteins such as the adrenoceptors are abundant, comprising between 20-30 % of known proteins.[32,33,34] However, knowledge of the 3D structures of membrane proteins is limited, due to difficulties linked to their non-degenerative purification and crystallisation.[6,25]

1.2.2 G Protein Coupled Rreceptor Structure

On the basis of sequence comparisons, GPCRs have been classified into three different families, Class A 'rhodopsin-like', Class B 'secretin-like' and Class C 'glutamate-metabotropic-like'. Apart from their conserved secondary domain structure, the GPCR families do not share significant sequence homologies. Family A is by far the largest, displaying short amino-terminal tails and having highly conserved amino acid residues within each transmembrane helix.[35] At present, the most thoroughly documented example of a Class A GPCR is Rhodopsin, which is naturally expressed in large amounts in the disk membranes of the rod and cone cells in the retina.[36] Significant progress in understanding the structure and function of bovine Rhodopsin has been made in recent years[45] and the presence of a bundle of seven TM α helices

(TM I-VII) was established by high-resolution electron diffraction and by a crystal structure at a 2.8 Å resolution in 2000.[37] The α helical bundle is connected by intracellular (IC) and extracellular (EC) loops as illustrated in the generic GPCR model in Figure 1.5. An additional short eighth helix (TM-VIII) was determined almost parallel to the plane of the membrane for Rhodopsin.

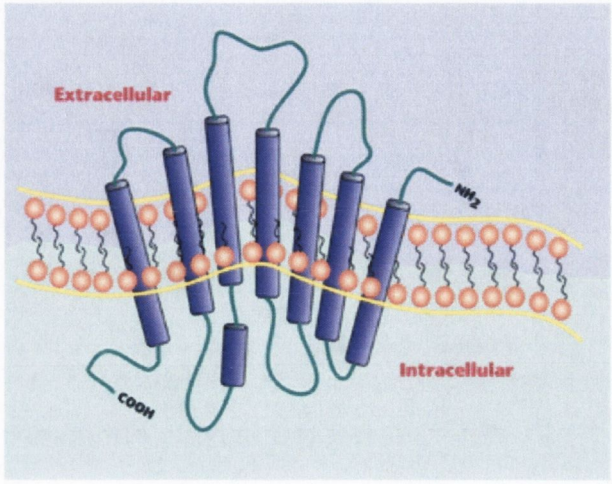


Figure 1.5: Model of the seven transmembrane region of a GPCR embedded in a membrane.[38]

GPCRs share a common membrane topology they are remarkably diverse in sequence and vary especially in the size of the extracellular amino-terminal tails, cytoplasmic loops and carboxy-terminal tails.[39] Although bovine Rhodopsin has only a low sequence similarity to other GPCRs, the specific arrangement of the seven TM helices, stabilized by a series of intramolecular interactions mediated by several backbone and side-chain atoms, appears to be conserved among the Class A receptors.

1.2.3 Development of an Adrenoceptor Model

GPCRs make an interesting and relevant challenge for the realm of protein structure prediction. Theoretical protein structure prediction hinges on the Levinthal paradox - the observation that there is insufficient time to randomly search the entire conformational space available to a polypeptide chain. How then does one move from the Levinthal Paradox to protein structure prediction?[40] Due to the enormously complex energy landscape of proteins, the number of local minima must be reduced by predicting the structure through an *ab initio* or homology/comparative modelling method in order to obtain a set of predictions in a reasonable timeframe. To validate the

available tools, a series of benchmarks were developed to measure the performance of fold recognition methods, known as Critical Assessment of Protein Structure Predictions (CASP). Experimentalists provide amino acid sequences of soon-to-be-determined protein structures and invite the protein prediction community to test their methods on the unknown targets. This format propels the improvement of methods and the progress in the field can also be seen by evaluations of the biannual prediction exercises.[41]

As protein structure is more conserved than protein sequences, homology modelling methods, which were first reported by Browne *et al.*,[42] attempt to predict the 3D structure of a protein sequence (target) by using information derived from an homologous protein of known structure (template).[43,44] Class A GPCRs for which the rhodopsin template is available, share sequence homology and also key structural features such as a disulfide bond between TM-III and the extracellular (EC) region and a tripeptide Glu/Asp-Arg-Tyr (E/DRY) motif located at the intracellular (IC) end of TM-III. There are several other highly conserved residues, such an Asn-Asp pair located in TM-I and TM-II, respectively, Pro residues in TM-V and TM-VI, aromatic residues in TM-IV and TM-VI and a common Asn-Pro-X-X-Tyr (NPxxY) motif in TM-VII.[45]

In the modelling procedure, structurally conserved regions (SCRs) such as the helical bundle are identified between the template and the target and their sequences are aligned.[46] A multiple alignment procedure such as that used in clustalW[47] produces biologically meaningful alignments of divergent sequences. Using such an alignment, GPCR models can be constructed on a template of Rhodopsin α helices, which provide guidelines for positioning, tilting and orientating the helices in the adrenoceptor TM bundle. The receptor model can be further improved and validated by moving from a static to a dynamic picture, through molecular dynamics simulations (MD), a technique frequently applied since the first simulation of bovine pancreatic trypsin inhibitor (BPTI) was reported in 1977.[48] MD simulations yield insights into the natural dynamics of biomolecules on different timescales, affording averages of molecular properties and exploring thermally accessible conformations of a receptor or a ligand/receptor complex.[49]

A major driving force in determining the tertiary structure of globular proteins is the hydrophobic effect.[50] A receptor, as it goes through a cell membrane is exposed to a heterogeneous environment, where the TM regions are embedded in a phospholipid bilayer and the extramembrane domains are surrounded by water. The water exposed polypeptides (loop regions) can adopt a diverse array of folds, whereas the physical and chemical constraints imposed by the lipid bilayer appear to restrict the structural diversity of the embedded protein. Hence, a MD simulation of the receptor in its natural environment requires the introduction of hydrophobic interactions to mimic the effect of the membrane on the receptor. This can be achieved through using explicit models of phospholipids bilayer or a less time consuming membrane-mimetic system.[51,52] The time scale of lipid motion is slow, with lateral motions within the bilayer requiring simulations to the order of nanoseconds. Starting from a homology model rather than a crystal structure, it is possible that the slow time scale of lipid motion effectively restricting receptor mobility could mask large-scale conformational drift that might provide evidence for an incorrect fold in the model. As an alternative, a water and carbon tetrachloride membrane mimic was used to model the interaction of the antagonist Hoe-140 with the bradykinin receptor.[53] A similar approach was applied by Pellegrini *et al.*[54] to study the threonine/bradykinin receptor interaction. Further biphasic systems studied include a water/chloroform system by Troxler *et al.*,[55] and a water/hexane system by Wymore *et al.*.[56] The latter concluded that the same qualitative features were achieved in simulations of the biphasic (hexane/H_2O) and micellar (SDS dodecylsulfate) system to that of the phospholipid dimyristoylphosphatidyl (DMPC).

1.3 Current antagonists of α_1 adrenoceptors

As previously mentioned, prostatic and bladder neck smooth muscle tone is mediated by α-ARs and many clinical trials have shown that α-AR antagonists can significantly improve symptoms of BPH.[21] The first α-AR antagonist observed to reduce BPH symptoms was the long acting nonselective α_1/α_2 AR blocker, phenyoxybenzamine.[16] The increase in urinary flow rates observed with phenyoxybenzamine provided 'proof of concept' that α_1-ARs antagonists could have a therapeutic effect on BPH. However, its use was restricted due to significant side effects in a large percentage of patients (10-30 %),[57] which included dizziness, weakness and palpitations, decreased blood pressure, nasal congestion and impotence. Many of the side effects are due to interactions with α_2-ARs in the cardiovascular system where they

cause venodilation.[16,58] As there is considerable overlap (39.2 %)[59] between the populations of men who suffer from high blood pressure and BPH, this may provide clinicians with the opportunity to treat two conditions with a single agent. In cases were this dual effect is undesirable, a selective α_1-AR over α_2-AR agent is required, necessitating further research into more selective α_1-AR agents.

The first selective α_1-AR antagonist determined was prazosin, which is as effective as phenyoxybenzamine but with fewer side effects, see Figure 1.6.[16] The primary advantage of prazosin is its cost as it is available generically (Minipress®, Hypovase®, Alza). However, its widespread use has been limited by its relatively short half-life, which necessitates frequent dosing. Alfuzosin (Xatral®, Sanofi-Synthelabo) is also selective for α_1-ARs over α_2-ARs and for the lower urinary tract but without subtype selectivity.[13] It is as effective as prazosin in the treatment of BPH with a lower incidence of adverse events. Further development in the quinazoline class led to the longer acting agents doxazosin (Carduara®, Pfizer) and terazosin (Hytrin®, Abbott Laboratories), which are selective, α_1-AR blockers and have once daily dosing.[13,21]

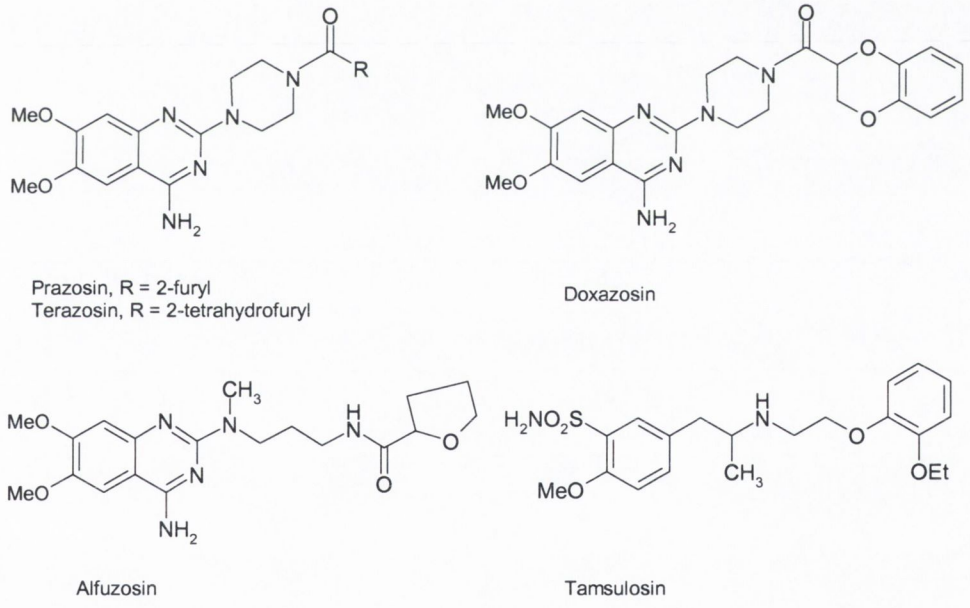


Figure 1.6: Current antagonists available in the treatment of BPH; prazosin, terazosin, doxazosin, alfuzosin and tamsulosin.

Selective α_1-AR over α_2-AR antagonists are typically effective for treating BPH symptoms but side effects of dizziness, asthenia and peripheral edema are still present.[16] However, α_{1A}-AR selective antagonists have the potential benefits of ease of dosing, better compliance, reduced cost, improved efficacy, better drug-drug interaction profile

and fewer side effects. Tamsulosin (Flomax®, Yamanouchi Pharmaceuticals Co. Ltd.), a phenylethylamine type α blocker, is the only agent currently used for the treatment of BPH, which shows substantial subtype selectivity for the α_{1A}-AR (Figure 1.6).[13]

1.3.1 Towards selective α_{1A}-AR Antagonists

The success of Tamsulosin has stimulated the quest for even more α_{1A}–AR selective blockers, but none have been approved to date. As a consequence of the limited availability of structural data on GPCRs the design of antagonists heavily relies on ligand based drug design techniques. Imidazoline derivatives have traditionally been considered as one of the major types of drugs that interact with α-ARs. Compounds such as clonidine (often prescribed for high blood pressure) or naphazoline containing a 2-iminoimidazolidine or an imidazoline ring respectively, show α_1 and α_2 AR activities and more specifically, phentolamine containing an imidazoline ring, is a known α_1-AR antagonist (Figure 1.7).

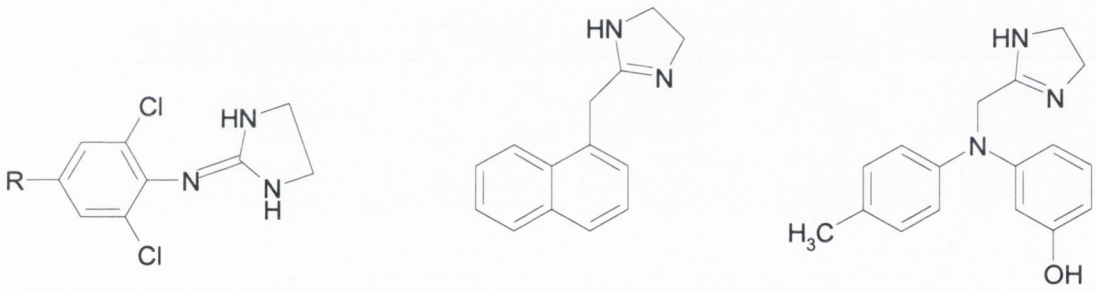
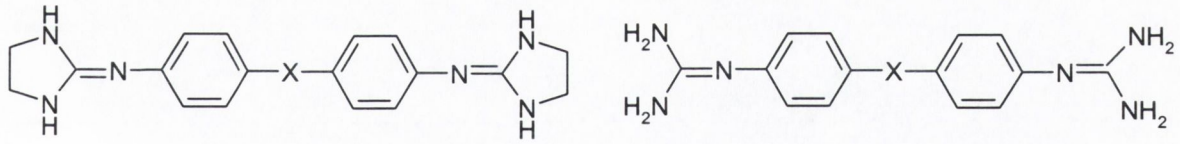


Figure 1.7: Imidazoline derivatives containing a 2-iminoimidazolidine or an imidazoline ring, which is known to interact with α-ARs including Clonidine (R=H) on left, Naphazoline (in middle) and Phentolamine on right.

Previous studies have led to the preparation of two new families of biphenyl compounds for both imidazolidinium and guanidinium compounds connected by a group (X) containing polar atoms, which were pursued as α_{1A}-AR antagonists (Figure 1.8).[60] The imidazolidinium compounds **1-4**, are named denoting the bridging group X as, **1** X=NH; **2** X=CO; **3** X=SO_2 and **4** X=CH_2 for a non polar comparison. Similarly for the guanidinium compounds, **5-8**, where **5** denotes X=NH; **6** X=CO; **7** X=SO_2 and **8** X=CH_2.

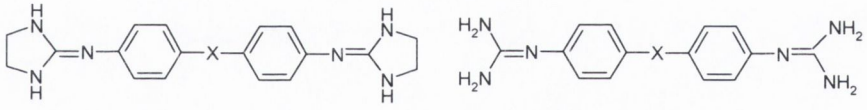


1-4 5-8

Figure 1.8: Imidazolidinium compounds **1-4** and guanidinium compounds **5-8** where for compounds **1** and **5** (X=NH); compounds **2** and **6** (X=CO); compounds **3** and **7** ($X=SO_2$) and compounds **4** and **8** ($X=CH_2$).

Pharmacological studies of compounds (**1-8**) were performed on slices of human prostatic tissues with BPH, introduced in organ baths with Krebs solution. Cumulative concentration response curves to ND (10^{-7}-10^{-4} M) were constructed and the drug to be tested was added and incubated for 10 minutes, before a second ND cumulative curve was produced. The capacity of the bis-imino-imidazolidinium derivatives **1-4** and bis-guanidinium derivatives **5-8** to inhibit the contraction induced by noradrenaline on human prostate tissue with BPH was tested (Table 1.1). The clinical antagonist doxazosin was used as a reference compound. For the guanidinium compounds **6**, **8** and **5** were most active, with compound **5** having a better dosage dependence. For the imidazolidinium compounds, major blockage was achieved for compounds **2** and for **4**.[61]

Table 1.1: Percentage inhibition of contraction of compounds **1-8** on human prostate tissue with BPH.



1-4 5-8

X	Imidazolidinium	Guanidinium	Doxazosin
Doxazosin			95
NH	78	90	
CO	82	**95**	
SO_2	29	-	
CH_2	80	92	

Considerable industrial and academic effort has been dedicated to identifying subtype selective α_{1A}-AR antagonists; examples industrially include Johnson & Johnson (Kuo *et al.*[62]) and at Abbott laboratories (Meyer *et al.*[63]), while academic interest,

amongst others includes Chern *et al.*,[64] Quaglia *et al.*,[65] Barbaro *et al.*,[66] Wong *et al.*,[67] Lopez *et al.*[68] and Corsano *et al.*.[69] A selection of α_1-AR antagonists from the literature in different structural classes shall be further examined.

1.3.2 Computational Study of α_1-AR antagonists

Although α_1-AR antagonists cover a range of structural types, all the compounds possess a central basic unit flanked on at least one side by aromatic systems. The precise profile in terms of subtype selectivity is heavily dependent on the nature of the basic centre, the substitution of the aromatic rings and the spatial orientation of the groups. The presence of a protonated nitrogen at physiological pH would appear to be a vital feature for α_1–AR antagonists. The protonated nitrogen can facilitate the formation of weak, attractive HBs with receptor residues. Hence, we are interested in quantifying the proton affinity of the nitrogens in the ligands and in estimating the likelihood of these nitrogens to protonate under physiological conditions. This can be estimated by following a thermodynamic cycle relating pKa to the gas-phase proton basicity (ΔG^{gp}) via the solvation energies (ΔG_s) of the products and the reactants. Similar pKa calculations have been successfully performed on a series of molecules.[70,71,72]

A commonly adopted approach to drug design relies on establishing the structure-activity relationships of a number of known ligands of a receptor in order to develop a pharmacophore model. The term pharmacophore is traced to Emil Fischer (1894) and Paul Ehrlich (1909) and refers to the molecular framework (phoros) that carries the essential features responsible for a drugs biological activity (pharmacon). Structural features necessary for the selective affinity of ligands for α_1-ARs were previously investigated.[73] Low energy conformations of a set of known antagonists were used as input for a training set and the models were validated by testing the fit of known antagonists not included in the training set to the pharmacophore model.[73] The pharmacophore model for the α_{1A}-AR required that the distance of the protonated nitrogen from a polar group (polar distance) should be in the range of 6-8 Å and its distance from an aromatic group (aromatic distance) should be in the range of 5.2-5.8 Å (Figure 1.9). For α_{1B}-AR, a polar distance of 5-6 Å and an aromatic distance of 6.2-7.8 Å are necessary. Later, Bremner *et al.*,[74] also published the first pharmacophore model for the α_{1D}-AR where a polar distance of 4.5 Å and an aromatic distance of 5.4 Å are

required. Recently, an α_{1A}-AR pharmacophore[75] was developed from a larger set of compounds, which also contains the main features of that of Bremner *et al.*.[74]

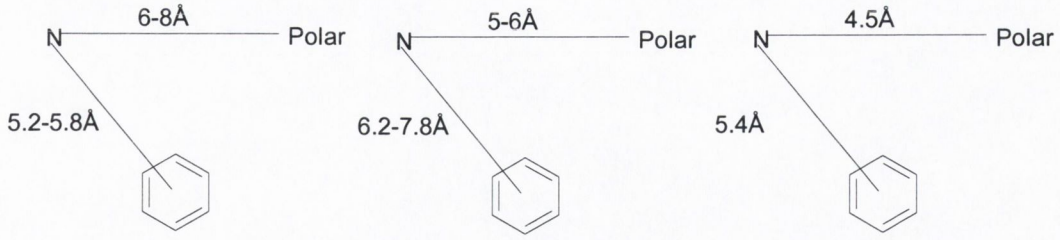


Figure 1.9: Pharmacophore details for antagonists of α_{1A}-AR, α_{1B}-AR and α_{1D}-AR indicating the relevant 'polar distances' and 'aromatic distances'.[73,74]

1.4 Ligand/α_{1A}-AR Complexes

Once the appropriate protonation states of the agonists and antagonists have been established and homology models for the receptor generated, it would be useful to be able to predict (a) in what configuration and how strongly the two will bind to each other and (b) the resultant conformational changes induced.

1.4.1 Experimental Agonist Binding Studies

Both agonists, adrenaline and noradrenaline, contain a protonated amino group separated from the aromatic catechol ring by a β-hydroxylethyl chain. Mutagenesis studies of the β_2–AR and α_{2A}–AR suggest that the amino group of the catecholamines makes an electrostatic interaction with the carboxylate side chain of an Asp residue of TM-III, which is highly conserved in all GPCRs binding with amine ligands. Mutation of this Asp residue in the β_2-AR (Asp-113),[76] and in α_{1B}-AR (Asp-125)[77] resulted in lower or nondetectable binding affinity of the receptor agonists. There are two Asp residues (Asp-106 and Asp-123) in TM-III of the α_{1A}–AR; Asp-106 is located at the extracellular side and is conserved only in the biogenic amine receptors, while Asp-123 of the conserved DRY motif is located at the intracellular side, is highly conserved in Class A GPCRs and is likely to be involved in interactions with the G protein.[78]

Another fundamental interaction in the binding of catecholamine agonists is hydrogen bond (HB) formation between the catecholic hydroxyl groups and two conserved Ser residues located in TM-V.[79] In the α_{1A}-AR it was determined that Ser-188 may interact with the *meta*-hydroxyl whereas Ser-192 may interact with the *para*-

hydroxyl of the catechol ring.[80] Ambrosio *et al.*,[81] also concluded that the equivalent Ser-204 and Ser-207 of the β_2 -AR not only provides a docking site for the agonist, but also controls the equilibrium of the receptor between the active (R*) and the inactive (R) forms. To date the role of the *beta* hydroxyl in the interactions of the catecholamines remains unclear.

1.4.2 Agonist Activation of the α_{1A}–AR

Obtaining structural information about agonist/α_{1A}-AR complexes has been hampered by the lack of a GPCR crystal structure in the activated state. Currently, there are a number of postulated activation mechanisms for Class A GPCRs, which essentially consist of three steps:

1 Admission of the agonist into the binding pocket.
2 Movement of the receptor from the inactive state to the active state.
3 Activation of the G protein.

The concept that agonist induced activation of a receptor, involves a conformational change, was originally developed in 1957 by del Castillo and Katz, who proposed that binding of acetyl-choline to the nicotinic acetylcholine receptor led to a conformational change resulting in channel opening.[82] Agonist binding to GPCRs is believed to promote a conformational change that leads to a switch from the inactive ground state (R) to the active (R*) conformation of the receptor.[83] A movement of the second extracellular loop (EC-II) in response to ligand binding may cause TM-III to translate in the cytoplasmic region near the G protein and thus initiate the signal transduction pathway.[84] Agonist induced conformational change of the β_2-AR is associated with just such a change in the relative orientation of TM-III and TM-VI (with a rotation of TM-III and a separation from TM-VI),[83] which could unmask G protein-binding sites. The rigid body motions of the two portions of the Pro kinked TM-VI were proposed as key dynamic components of the rearrangement of GPCR molecules upon activation by agonist binding.[85] Subsequent experimental probing of the structural details of molecular rearrangements following activation of Rhodopsin[86] or β_2-ARs[87] identified the same type of rigid body motions and the probable involvement of the conserved Pro in TM-VI.

All Class A GPCRs share a set of highly conserved residues such as the E/DRY motif at the G-protein coupling interface.[88] The high-resolution Rhodopsin structure indicates that Glu and Arg of the ERY motif in TM-III interact and thus a similar interaction is predicted amongst all Class A receptors. Whether the conformation and the interaction pattern of the Arg is peculiar to the available Rhodopsin structure, will await the structural resolution of other closely related GPCRs. An additional interaction between this Arg of TM-III and a conserved Glu in TM-VI is proposed for the β_2–AR[89] and the α_{1B}–AR[90] where it was suggested that an interaction between these residues would constrain the relative positions of TM-III and TM-VI and form an ionic switch that controls transition of the receptor from its inactive to active state.[91] A study of the 5-HT1A receptor, (another GPCR), strengthens the hypothesis that the receptor portions close to the E/DRY motif, with prominence to the cytosolic extensions of TM-III and TM-VI, are particularly susceptible to structural modification in response to agonist binding. Both agonists studied exerted a destabilization of the intrahelical and interhelical interactions found in the empty and antagonist bound receptor forms.[92] Alternatively, receptor activation for the α_{2A}-AR was envisioned as a rotation that exposes both TM-V Ser residues to the ligand binding cavity and movement of TM-V toward the ligand, TM-III and TM-VI.[93] Their results indicate that the catecholic hydroxyl groups are not critical for binding, but are very important for receptor activation.

In summary, an activation mechanism has only been postulated for a number of Class A receptors including the α_{2A}–AR and β_2-AR (for reviews[94,95]). Agonist binding may induce specific rearrangements in Pro kink regions or rigid body motion of TM-V and TM-VI.[96] Such studies support a movement of TM-VI relative to TM-III as a consequence of agonist induced receptor activation, coupled with a change in the interactions of residues in the conserved E/DRY motif.

1.4.3 Experimental Antagonist Binding Studies

Based on molecular modelling and mutagenesis studies performed on several GPCRs binding diverse types of ligands (biogenic amines, peptides, nucleotides and small proteins) it is suggested that all Class A GPCRs share a binding pocket in the 7-TM region. This pocket can serve as an interaction site, not only for the monoamines, but for all agonists or antagonists of this receptor family.[97,98] Ishiguro *et al.*,[99] postulated

that the principal interactions of the antagonists, prazosin and tamsulosin were with Asp-106 and Ser-188 of α_{1A}-AR. The other molecular moieties then fit into two binding pockets formed by TM-I, TM-II and TM-VII on one side and by TM-IV, TM-V and TM-VI on the other one, lying in an almost symmetrical topography with respect to Asp-106. Other studies suggest that in contrast to agonist binding, which is localized to the interior core of the receptor, antagonists interact with residues closer to the extracellular surface of ARs. Zhao *et al.*,[100] determined that residues, Gln-177, Ile-178 and Asn-179 of EC-II, are responsible for some selective antagonist binding in α_{1A}-AR over α_{1B}-AR. Furthermore, attractive interactions between π systems are one of the principal noncovalent forces governing molecular recognition. According to Hamaguchi *et al.*,[101] the Phe-86 residue in TM-II of the α_{1A}-AR, located in the final helical turn closest to the extracellular region, is the principal determinant responsible for the subtype selectivity of antagonists for α_{1A}-AR over α_{1D}-AR. Based on mutagenesis studies, the same group,[102] determined that two Phe (Phe-308, Phe-312) residues in TM-VII of the α_{1A}-AR are involved in nonselective binding for almost all α_1–AR antagonists.

1.4.4 Theoretical Ligand Binding Studies

Molecular docking such as that implemented in Dock 4.0,[103] in combination with scoring functions[104,105,106] are used to predict ligand orientations in binding sites and subsequent binding affinities (BAs).[107] All docking methods make limiting assumptions to be computationally reasonable, (*i.e.* generate a large number of conformations quickly) with each method having its own unique advantages and disadvantages. Many docking programs often consider the receptor to be rigid and only allow flexibility to the ligand. However, as protein flexibility and the dynamics of inter molecular interfaces can regulate binding affinity and specificity in molecular recognition, the Flexidock[108] routine attempts to address the issue of flexibility of receptor residues. Subsequent MD simulations on the ligand/receptor complexes can provide further information on binding modes and complex stability.

1.5 Thesis Layout

The Economist[1] in 1998 stated, "Turn off the Bunsen burner; it will not be wanted these ten years," and indeed computational approaches including structure based drug design methods have had a significant impact on the creation of high-value compounds entering the market as drugs and clinical trials. Seven such compounds are now approved and marketed drugs,[109] including Captopril whose template structure used for its design was a homology model (August 2003). The primary objective of this thesis is to computationally investigate the interactions of agonists and antagonists with the α_{1A}-AR, towards an improved understanding of the activation mechanism and antagonist blockade in the treatment of BPH. An introduction to the topic and a brief overview of the issues at hand has been presented in Chapter I.

In Chapter II, the computational techniques employed are discussed. These include quantum mechanical, semi-empirical and molecular mechanical techniques. Computational methods to evaluate atomic charges and solvation energies are also examined.

In Chapter III, a computational pipeline for the work pursued is presented. The computational procedures and methods are divided into receptor based (homology based protein models), ligand based (quantum mechanical calculations, pharmacophore models) and ligand-receptor binding (docking, molecular dynamics) categories. In each of the four experimental chapters (IV-VII) we present the corresponding results and discussions.

In Chapter IV, homology models of the α_{1A}-AR are developed using the bovine Rhodopsin crystal structure as a template. Receptor refinement is pursued through MD simulations in a variety of phospholipid bilayer mimics to the natural environment of a membrane receptor. A detailed analysis is performed on the resultant trajectories and the conformation of the receptor in the inactive state are examined.

In Chapter V, a computational study of the agonists and antagonists is pursued. A proton affinity study is performed followed by conformational analyses and structural optimisations of the ligands. Estimates at ligand pKa values are performed and limitations in the applied theoretical approach are discussed. Conformations of the

ligands fitting the α_{1A}-AR pharmacophore model are determined in a form of *in silico* ligand screening.

In Chapter VI, we develop a computational strategy, which involves initial docking of the agonists into homology models of the α_{1A}–AR followed by a cycle of MD simulations of the agonist/α_{1A}–AR complexes in a $H_2O/CHCl_3/H_2O$ membrane mimic to optimise them. Insights into the activation process are also gained through examination of the conformational changes occurring over the complexed MD simulations and changes in the interactions of ligands and of the conserved receptor residues.

In Chapter VII, we perform combined docking and MD simulations of antagonist/α_{1A}–AR complexes, to produce a series of 'antagonist-bound' receptor forms. Furthermore, we perform MD simulations with an adapted force field of the antagonist/α_{1A}–AR complexes to yield an analysis of binding interactions and conformational changes occurring upon binding. Screening of a further series of antagonists, **9-18**, is performed utilising the ligand induced receptor structures. Finally, a thorough comparison is made of the uncomplexed, 'agonist-bound' and 'antagonist-bound' receptor conformations produced in the study to examine the effect of different ligands on the receptor structure.

The final Chapter VIII presents the conclusions and insights gained throughout the course of the work. Attention is drawn to a number of key directions for future development.

Chapter II: Computational Techniques

"Anyone who is not shocked by quantum theory has not understood it."

Neils Bohr

2.0 Introduction

In this chapter, we introduce quantum mechanical (QM) and molecular mechanical (MM) techniques. The QM techniques shall be discussed in terms of Hartree Fock (HF) Theory and Density Functional Theory (DFT). Further simplifications to the model shall then be described including semi-empirical techniques and MM potentials. The utilised basis sets, charge determination and solvation techniques shall also be explored. The application of these techniques to the current study will be developed in the subsequent chapter.

2.1 Quantum Mechanics

The term *ab initio* ("from the beginning") is used to describe rigorous QM techniques, which are formulated without experimental data. QM methods attempt to solve the Schrödinger equation (1) and are necessary to describe the electron distribution of a molecule in detail.[110] To solve the Schrödinger equation it is necessary to find values of the wavefunction (Ψ) which, when operated upon by the Hamiltonian (H), returns the wavefunction multiplied by the particle energy (E).

$$H\Psi = E\Psi \qquad \textbf{(1)}$$

Following the Born Oppenheimer approximation the motion of electrons can be decoupled from the motion of nuclei by assuming the nuclear centres of mass are fixed for a given molecular conformation. This is a reasonable approximation as the nuclei, which are much heavier than the electrons, are typically fixed on the timescale of electronic vibration. Once assumed, the problem is reduced to solving the electronic Schrödinger equation (2) for a set of nuclear geometries.

$$H_e\Psi_e=E_e\Psi_e \qquad \textbf{(2)}$$

In an atom that contains a single electron the Hamiltonian can be written as the kinetic and potential energies of the nuclei and the electron (3). The kinetic energy depends upon the mass m, Plancks constant $\hbar$ and the Laplacian, $\nabla^2 = \frac{\delta}{\delta x^2} + \frac{\delta}{\delta y^2} + \frac{\delta}{\delta z^2}$.

The potential energy depends upon the distance r, between the electron and the nucleus

as given by the Coulomb equation, were Z is the nuclear charge, e is the electronic charge and ϵ_0 is the dielectric constant in vacuum.

$$H_e = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\varepsilon_0 r} \tag{3}$$

The Schrödinger equation can be solved exactly for a one-electron system, as there is no electron-electron interaction term. However, no analytical solutions have been determined for systems that involve three (or more) interacting particles. Hence, a series of approximations are generally introduced to make the calculations more tractable.

2.1.1 Hartree-Fock Approach

Within the framework of the approximations used in the HF approach, a molecular wavefunction is calculated and various molecular properties are determined. The Hartree product (HP) formulates the total electronic wavefunction (Ψ^{HP}) by describing the motion of the electrons as the product of the individual one electron wavefunctions $\chi(x_N)$, (4).

$$\Psi^{HP}(x_1 \ldots x_N) = \chi_1(x_1)\chi_2(x_2)\chi_3(x_3)\ldots\ldots\chi_n(x_N) \tag{4}$$

In 1930, Fock[111] observed that the Hartree wavefunction is invalid as it does not satisfy the Pauli Exclusion Principle[112] - that the wavefunction must be antisymmetric with respect to electron interchange. Fock also demonstrated that a Hartree product could be made antisymmetric by appropriately adding and subtracting all possible permutations. Later, Slater established that the resulting wavefunction is simply the determinant of the system of N electrons and N spin orbitals, termed a Slater determinant (5), which yields the antisymmetric wavefunction for N indistinguishable particles.[113,114]

$$\Psi = \frac{1}{\sqrt{N}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_N) & \chi_1(x_1) & \cdots & \chi_N(x_N) \end{vmatrix} \quad \textbf{(5)}$$

Describing the electrons by an antisymmetrized product, the Slater determinant, is equivalent to assuming that each electron moves independently of all the others except that it feels the Coulomb repulsion due to the average positions of all electrons and an exchange interaction due to antisymmetrization. It is more convenient to write the energy expression in a concise form that recognises the three types of interaction that contribute to the total electronic energy of the system. The Fock operator, f_i, (6), is an effective one-electron Hamiltonian, comprising of a core Hamiltonian term $H^{core}(1)$, a Coulomb operator $J_j(1)$ and an exchange operator $K_j(1)$ (7).

$$f_i \chi_i(1) = \sum_{j=1}^{N} \varepsilon_{ij} \chi_j(1) \quad \textbf{(6)}$$

$$\left[H^{core}(1) + \sum_{j=1}^{N} \{ J_j(1) - K_j(1) \} \right] \chi_i(1) = \sum_{j=1}^{N} \varepsilon_{ij} \chi_j(1) \quad \textbf{(7)}$$

The core Hamiltonian (8) is the average kinetic energy and the potential energy for the electrostatic attraction between the nuclei and the electron. The Coulomb integral (9) is the potential energy for the electrostatic repulsion between two charge distributions. The exchange integral (10) arises from the requirement that Ψ be anti-symmetric with respect to the permutation of the coordinates of any two electrons and is the additional term acquired by using the Slater determinant.[115]

$$H^{core}(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}} \quad \textbf{Core Hamiltonian operator} \quad \textbf{(8)}$$

$$J_j(1) = \int dr_2 \chi_j(2) \frac{1}{r_{12}} \chi_j(2) \quad \textbf{Coulomb operator} \quad \textbf{(9)}$$

$$K_j(1) = \int dr_2 \chi_j(2) \frac{1}{r_{12}} \chi_i(2) \quad \textbf{Exchange operator} \quad \textbf{(10)}$$

The *variation theorem* states that the best approximation to the true wavefunction is the one closest in energy to it. The HF equations are obtained by imposing this condition on the energy expression subject to the constraint that the molecular orbitals (MO) remain orthonormal. The orthonormal condition is written in terms of the overlap integral S_{ij} (11).

$$S_{ij} = \int \chi_i \chi_j d\tau = \delta_{ij} \qquad (\delta_{ij} = \text{Kronecker delta}) \qquad \textbf{(11)}$$

So far we have been discussing abstract molecular orbitals. To implement a self consistent field (SCF) method we now introduce a systematic way of varying the MOs and hence minimizing the energy. This is achieved by representing the MO's as a linear combination of basis functions. In this study, our basis functions are similar to atomic orbitals and the method is referred to as a linear combination of atomic orbitals (LCAO). Thus each molecular orbital can be written as a summation, where ψ_i is a (spatial) molecular orbital, ϕ_μ is one of K basis functions and $c_{\mu i}$ is a coefficient (12).

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i} \phi_\mu \qquad \textbf{(12)}$$

The HF equations form a set of pseudo-eigenvalue equations as the Fock operator depends on all the occupied MOs. The derivation of the HF equations for a closed shell system, was proposed by Roothan[116] and by Hall,[117] the resulting equations are known as the Roothan-Hall equations, which are in matrix form and can be conveniently written as a matrix equation (13). C is the matrix of coefficients $c_{\mu i}$, E is a matrix of energy eigenvalues, S is the overlap matrix between two atomic orbitals and F is the matrix representation of the Fock operator.

$$FC = SCE \qquad \textbf{(13)}$$

Elements of the Fock matrix which appear on the left hand side of equation 13 depend on the MO coefficients, $c_{\mu i}$, which appear on the right hand side of equation 13. Thus an iterative procedure is required to find a solution and the SCF procedure is implemented. First, a set of trial solutions for the HF eigenvalue equations are obtained.

These are used to calculate the Coulomb and Exchange operators. The HF equations are solved, giving a second set of solutions, which are used in the next iteration. The implementation of the SCF method gradually refines the individual electronic structures that correspond to lower total energies until the point is reached at which the results for all the electrons are unchanged and they are then said to be self-consistent.

The HF energy is not as low as the true energy of the system as the Fock operator treats each electron as though it were moving in a time averaged potential field due to other electrons. However, as the motion of electrons is correlated, they tend to 'avoid' each other more than the HF theory would suggest. In order to consider correlated properties, it is necessary to go beyond the HF method. One way of handling the electron correlation problem with only the computational expense of HF theory is through the density functional theory (DFT), which shall now be explored.

2.2 Density Functional Theory

The basis for density functional theory (DFT) is the proof by Hohenberg and Kohn[118] that the ground state electronic energy can be determined completely from the electronic density $\rho(r)$.[115] The energy E, is a unique function of $\rho(r)$ and the energy functional (a function of a function) can be written as:

$$E[\rho(r)] = \int V_{ext}(r)\rho(r)dr + F[\rho(r)] \quad \textbf{(14)}$$

The first term $V_{ext}(r)$, arises from the interaction of the electrons with an external potential (typically due to the Coulomb interaction with the nuclei). The second term $F[\rho(r)]$, is the sum of the kinetic energy of the electrons and the contribution from the inter-electron interactions. In 1965, Kohn and Sham,[119] suggested that $F[\rho(r)]$ could be approximated as the sum of $E_{KE}[\rho(r)]$ the kinetic energy; $E_H[\rho(r)]$, the Hartree electrostatic energy, *i.e.* the electron-electron Columbic energy; and $E_{XC}[\rho(r)]$, which contains contributions from exchange and correlation (15).

$$F[\rho(r)] = E_{KE}[\rho(r)] + E_H[\rho(r)] + E_{XC}[\rho(r)] \quad \textbf{(15)}$$

The density $\rho(r)$ of the system is equated as the sum of the square of a set of one-electron orthonormal orbitals (16).

$$\rho(r) = \sum_{i=1}^{N} |\psi_i(r)|^2 \tag{16}$$

The kinetic energy of a system of non-interacting electrons (N) with the same density $\rho(r)$ as the real system is then given by equation 17.

$$E_{KE}[\rho(r)] = \sum_{i=1}^{N} \int \psi_i(r) \left(-\frac{\nabla^2}{2} \right) \psi_i(r) dr \tag{17}$$

In the Hartree approach, the electrostatic energy, $E_H[\rho(r)]$, arises from the classical interaction between two charge densities, $\rho(r_1)$ or $\rho(r_2)$, which is summed over all possible pairwise interactions (18).

$$E_H[\rho(r)] = \frac{1}{2} \sum_{i \neq j}^{N} \iint \frac{\psi_i(r_1)^2 \psi_j(r_2)^2}{|r_1 - r_2|} dr_1 dr_2 \tag{18}$$

Combining these two terms and adding the electron-nuclear interaction leads to the full expression for the energy of a N-electron system within the Kohn-Sham scheme (19). This equation defines the exchange-correlation energy functional $E_{XC}[\rho(r)]$, which thus contains not only contributions due to exchange and correlation but also a contribution due to the difference between the true kinetic energy of the system and $E_{KE}[\rho(r)]$. Hence, in practical terms with DFT, the exact exchange (HF) for a single determinant is replaced by a more general expression, the exchange-correlation functional, which can include terms accounting for both the exchange energy and electron correlation omitted from HF. This term shall be discussed in more detail in the following section.

$$E[\rho(r)] = \sum_{i=1}^{N} \int \psi_i(r) \left(-\frac{\nabla^2}{2} \right) \psi_i(r) dr + \frac{1}{2} \sum_{i \neq j}^{N} \iint \frac{\psi_i(r_1)^2 \psi_j(r_2)^2}{|r_1 - r_2|} dr_1 dr_2 + E_{XC}[\rho(r)] - \sum_{A=1}^{M} \int \frac{Z_A}{|r - R_A|} \rho(r) dr$$

(19)

To solve the Kohn-Sham equations a self-consistent approach is taken. A set of orbitals is derived from an initial guess of the density, leading to an improved value for the density, which is then used in the second iteration and so on until convergence is achieved.

2.2.1 Exchange Correlation Functional

The exchange-correlation functional $E_{XC}[\rho(r)]$, is key to the success of the DFT approach. The simplest way to obtain this contribution is through the local density approximation (LDA). This is based upon the uniform gas model in which the electron density is considered constant throughout all space. The total exchange-correlation energy for the system, E_{XC}, is then obtained by integrating over all space. However, the LDA approximation has been shown to be inadequate in some cases and for this reason extensions have been developed.

In the early 1980s, it was recognised that not only the density but also the gradient of the density should be included in the functional expressions. If the function f(r) thus depends in some way on the gradients (or higher derivatives) of $\rho(r)$ then the functional is referred to as being 'gradient-corrected'. These gradient corrections are typically divided into separate exchange and correlation terms *i.e.* $E_X[\rho(r)]$ the exchange functional and $E_C[\rho(r)]$ the correlation functional.[120] The gradient correction to the exchange functional proposed by Becke is popular (B88),[121] while the correlation functional of Lee, Yang and Parr (LYP)[122] is widely used.

One potentially attractive option is to add a correlation energy derived from DFT (e.g. the LDA approximation) to the HF energy. In 1993, with Becke's functionals a breakthrough in molecular calculations occurred. Functionals for non-interacting electrons were combined with those in which the electron-electron interaction is fully switched on.[123,124] The most popular form was launched under the acronym "B3LYP"[125] and is an empirical mixture of exact exchange, Becke's gradient correction for exchange B88, the Lee-Yang-Parr (LYP)[126] correlation functional (with the gradient term) and the standard local correlation functional due to Vosko, Wilk and Nusair (VWN).[127] The constants A, B and C are those determined by Becke who computed the values of A=0.20, B=0.72 and C=0.81.[110]

$$E_{XC} = A * E_X^{HF} + (1-A)E_X^{LDA} + B * \Delta E_X^{Becke88} + CE_C^{LYP} + (1-C)E_C^{VWN} \quad \textbf{(19)}$$

2.3 Basis sets

An approximation inherent in all quantum mechanical methods is the choice of basis set. Two types of basis functions are normally used for molecular calculations – the Slater type orbital (STO) or the Gaussian type orbital (GTO). Although Slater type orbitals are physically more realistic, some of the integrals can not be solved analytically. Hence, Slater orbitals are commonly replaced by functions based upon Gaussians. The Gaussian function has the form $\exp(-\alpha r^2)$ where α is the radial extent or spread of the Gaussian. QM calculations use basis functions comprising integral powers of x, y and z multiplied by the exponent (20).

$$x^a y^b z^c \exp(-\alpha r^2) \quad \textbf{(20)}$$

In practice, a Gaussian function has the wrong behaviour at the nucleus and at the large *r* tail of the function and so many more Gaussian functions must be used. To efficiently improve this, a fixed linear combination of primitive Gaussian functions is used to create a single *contracted function* for which a single coefficient is varied.

A *minimal basis set* is a representation that contains just the number of functions (or contracted functions) that are required to describe the free atom. A basis set which doubles the number of functions in the minimal basis set is described as a *double zeta basis* (DZ). The core orbitals, unlike the valence orbitals, do not effect chemical properties significantly and vary only slightly from one molecule to the next. A variation of the DZ type basis only doubles the number of valence orbitals, producing a *split valence basis*. 'Pople style' *split valence* (SV) basis sets are frequently used, such as the 3-21G basis set were a contracted function of three Gaussian functions describes the core orbitals, while the valence electrons are represented by two contracted functions and one diffuse function. Similarly, the basis set, 6-31G, has six Gaussian functions describing the core and four describing the valence electrons, three of which are contracted functions and one diffuse function.

As atoms are brought closer together their charge distribution causes a polarization effect, which distorts the shape of the atomic orbitals. Introducing orbitals that have more flexible shapes is accomplished by adding in basis functions of higher angular momentum than the valence orbitals. For example, one can distort the spherical *1s* orbital of hydrogen by mixing in an orbital with *p* symmetry. Similarly, we can polarise *p* orbitals if we mix in an orbital of *d* symmetry. Such polarisation functions are introduced to basis sets and are denoted by an asterisk '*'. Thus 6-31G*, refers to a 6-31G basis set with polarisation functions on heavy atoms. 6-31G** indicates the additional use of polarisation functions on hydrogen and helium.

A deficiency of the basis set arises as Gaussian basis functions fall off too rapidly with r and are rather low far from the nuclei. This becomes particularly important when studying an anionic or excited state, as the loosely bound electrons are responsible for the energy in the tail of wave function. To compensate, highly diffuse functions can be added to the basis set, indicated by (+). A single + indicates an additional single set of diffuse *s*- and *p*- type Gaussian functions, while ++ indicates that the diffuse functions are included for hydrogen as well as for heavy atoms.

2.4 Approximate Molecular Orbital Theories - Semi Empirical Techniques

Addition of further approximations to the HF model can invoke semi-empirical methods. In such calculations, valence electrons are only included while the core electrons are subsumed into the nuclear core. Furthermore, the computation of a large number of the integrals in the HF-SCF calculation are simplified by either neglecting or replacing them with parameterised formulae that can be calculated more easily. The *zero differential overlap* approximation (ZDOA), neglects all products of basis functions depending on the same electron coordinates when located on different atoms. If the ZDOA is applied to all orbital pairs then the Roothan-Hall equations (13) for a closed shell molecule simplify considerably. However, the electron-core interactions between pairs of orbitals and the nuclear cores are not subjected to the ZDOA, as chemical bonding is reliant on this overlap character. The core integrals are parameterised to experimental results, compensating for the ZDOA to some extent.[115]

The AM1 (Austin Model 1) and PM3 (Parametric Method Number 3) methods were developed to improve deficiencies of earlier semi-empirical methods, which

included a tendency to overestimate repulsions between atoms separated by distances approximately equal to the sum of their van der Waals (vdw) radii. For AM1, the core-core term was modified using Gaussian functions. PM3 has essentially the same features as AM1, but the parameters were derived using an automated parameterisation procedure whereas chemical knowledge and intuition were used for AM1. [115]

2.5 Molecular Mechanical Potentials

Although quantum mechanical calculations of molecular electronic structure can be highly accurate, they are also costly in computational time. Hence, simulations of large molecules such as biological macromolecules are generally performed using classical molecular mechanics. Molecular mechanical (MM) calculations arose naturally from the concept of molecular bonding and vdw forces and are often referred to as 'force-field' (FF) or 'potential energy' methods. Many force fields (designated Class I) have been derived for DNA and proteins, such as the Assisted Model Building with Energy Refinement (AMBER)[128] and the Chemistry at HARvard using Molecular Mechanics (CHARMM).[129] Furthermore, FFs can be simplified by applying united atom parameters, where hydrogens are not explicitly considered, such as the Gromos FF (GROningen Molecular Simulation).[130] Class II force fields concentrate on reproducing small to medium size molecules to a high degree of accuracy in terms of geometry and vibrational frequencies *e.g.* MM2[131] and MM3.[132]

Calculating an electronic energy for a given nuclear configuration is bypassed by writing the total energy (E_{tot}) as a function of the nuclear coordinates. Electrons are not explicitly considered and the field they generate is not actually calculated, but rather represented by an "effective" potential treated by classical mechanics. We shall examine each term of E_{tot} (21), which is the sum of bond stretching $v_r(r)$, angle bending $v_\theta(\theta)$, torsion potentials $v_\phi(\phi)$, improper torsions $v_\chi(\chi)$ and non-bonding interactions $v_{nb}(r_{ij})$:

$$E_{tot} = \sum_{i,j} v_r(r) + \sum_{i,j,k} v_\theta(\theta) + \sum_{i,j,k,l} v_\phi\phi + \sum_{\chi} v_\chi(\chi) + \sum_{i>j} v_{nb}(r_{ij}) \qquad \textbf{(21)}$$

2.5.1 Bonded Functions

Hookes law can be utilised to represent the potential energy curve for a typical bond. This potential has the form of an harmonic oscillator and is only adequate for

small deviations in the bond length r_{ij} from a reference value, r_0 (22). Stronger bonds tend to have larger force constants k_r, which is a measure of the bonds resistance to change from the reference or 'equilibrium' value and is related to the harmonic vibrational frequency.

$$v_r(r_{ij}) = \frac{k_r}{2}(r_{ij} - r_0)^2 \qquad \textbf{(22)}$$

2.5.2 Three-body Functions

The deviation of an angle (θ_{ijk}) between atoms i, j and k from its reference value (θ_0) is also frequently described using Hookes law (23). The force constants (k_θ), in the harmonic angle potential are proportionately smaller than for the bonded function as less energy is required to distort an angle away from equilibrium than to stretch or compress a bond.

$$v_\theta(\theta_{ijk}) = \frac{k_\theta}{2}(\theta_{ijk} - \theta_0)^2 \qquad \textbf{(23)}$$

2.5.3 Four-body Functions

The existence of barriers to rotation about chemical bonds is fundamental to understanding the structural properties of molecules and conformational analysis. Torsional potentials are usually expressed as a cosine series expansion and one functional form is given in equation 24.

$$v_\phi(\phi_{ijkl}) = \frac{V_n}{2}\left[1 + \cos(n\phi_{ijkl} - \gamma)\right] \qquad \textbf{(24)}$$

ϕ_{ijkl} is the torsion angle between atoms i,j,k,l; V_n is referred to as the 'barrier' height and gives a qualitative indication of the relative barriers to rotation; n is the multiplicity, its value giving the number of minimum points in the function as the bond is rotated through 360°. Finally, the phase factor γ, determines where the torsion angle passes through the minimum value.

Improper torsion terms, also known as 'out-of-plane bending' motions, are defined by four atoms *i*, *j*, *k* and *l* (χ_{ijkl}) for which the central *j* atom is bonded to atoms *i*, *k* and *l*. A torsional potential of the following form (25) is then used to maintain the improper torsion angle at 0° or 180°.

$$v_{\chi}(\chi_{ijkl}) = k_{\chi}(\chi_{ijkl} - \chi_0)^2 \quad \textbf{(25)}$$

2.5.4 Non bonded interactions

Molecules also interact through non-bonded forces, whose determination generally are the most time consuming of all those discussed. Such non-bonded interactions (v_{nb}) can be short ranged (v_{SR}) or long ranged, electrostatic ($v_{coulombic}$) interactions (26).

$$v_{nb} = v_{SR} + v_{coulombic} \quad \textbf{(26)}$$

In the Lennard-Jones form the nonbonded potential consists of a short range repulsive (r^{-12}) and an attractive (r^{-6}) component due to vdw forces. Together these terms mimic the tendency of atoms to repel one another when they are very close and attract one another as they approach an optimal internuclear distance. The pairwise potential energy, v_{ij}, between two non-bonded atoms can be expressed as a function of the internuclear separation, r_{ij}. The two parameters A_{ij} and B_{ij} are fitted from a set of atomic radii and hardness parameters (27).

$$v_{ij} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \quad \textbf{(27)}$$

In principle, the non-bonded interactions are calculated between every pair of atoms in the system. However, the Lennard-Jones potential falls off very rapidly with distance. Hence, the number of non-bonded atom pairs can be significantly reduced by using a cut-off distance beyond which the atoms are no longer considered to interact.

The electrostatic interaction ($v_{coulomb}$) is calculated as a sum of interactions between pairs of point charges (q), where q_i and q_j are the charges on atom *i* and *j*

respectively and ϵ_0 is the permittivity of free space. Using Coulombs Law the interaction energy between the two charges q_i and q_j becomes:

$$v_{ij}^{Coulomb} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{28}$$

Due to the long range nature of the $\frac{1}{r}$ term one can not simply apply a real space short range cut off. The use of periodic boundary conditions allows the use of the Ewald[133] sum method to treat the long range nature of the electrostatic interactions. In the Ewald method, the sum is split into real and reciprocal space summations (29). For the real space part, each point charge is surrounded by a Gaussian charge distribution of equal magnitude and opposite sign, which spreads out radially from the charge. The parameter α, determines the width of the distribution and r is the position relative to the centre of the distribution. This extra distribution acts as an ionic atmosphere to screen the interaction between neighbouring charges. The screened interactions are now short-ranged and the total screened potential is calculated by summing over all the molecules in the central cube and all their images in the real space lattice of image boxes. Erfc(x) is the complementary error function which falls to zero with increasing x. The cancelling distribution is summed in reciprocal space, where L is the length of a side of the box and k are reciprocal vectors. The interaction of the cancelling distribution with itself, the *self-term* must be subtracted from the total where N is the number of atoms.

$$U = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\begin{array}{l}\sum_{|n|=0}^{\infty}\frac{q_i q_j}{4\pi\varepsilon_0}\frac{erfc(\alpha\,|\,r_{ij}+n\,|)}{|\,r_{ij}+n\,|} \\ +\sum_{nk\neq 0}\frac{1}{\pi L^3}\frac{q_i q_j}{4\pi\varepsilon_0}\frac{4\pi^2}{k^2}\exp\left(-\frac{k^2}{4\alpha^2}\right)\cos(nk.r_{ij}) \\ -\frac{\alpha}{\sqrt{\pi}}\sum_{k=1}^{N}\frac{q_k^2}{4\pi\varepsilon_0}\end{array}\right] \tag{29}$$

Thus, with the Ewald method the 1/r functional form of the Coulomb energy is split into two functions that are rapidly convergent in real and reciprocal space, respectively. The calculation in reciprocal space is very expensive for large systems of

charged atoms, scaling at $N^{3/2}$. Alternatively, with the smooth particle mesh Ewald (SPME) method the cancelling distribution is calculated by means of an interpolation procedure involving B-splines, with the sum in reciprocol space being represented on a 3D rectangular grid. In this form, a Fast Fourier Transform (FFT) may be used to perform the primary mathematical operation. The SPME method requires as input parameters the order of the spline and the number of grid points in the x-, y- and z-directions. Its advantage relies on the NlogN dependence of the computational efficiency with the system size.[115]

2.6 Evaluation of Atomic Charges

Given the widespread use of the partial atomic charge model, (charges restricted to the nuclear centre) it is important to consider how these theoretical charges are obtained. The electrostatic properties of a molecule are a consequence of the distribution of the electrons and the nuclei and thus one can obtain a set of partial atomic charges using QM techniques. Unfortunately, the partial atomic charge can not be unambiguously determined from the wavefunction. A variety of techniques have been proposed to determine partial atomic charges including those of Mulliken, Gasteiger-Marsili and a variety of electrostatic potential (ESP) derived charges.

2.6.1 Mulliken Population Analysis

Mulliken population analysis[134] is an orbital based method in which the electrons are divided amongst atoms according to the degree to which the different atomic orbital basis functions contribute to the overall wavefunction. Electrons 'shared' between basis functions are divided evenly between the two atoms. This approach has been used extensively as it is conceptually simple and is included in most *ab initio* and semiempirical packages. However, it depends strongly on the basis set used for the calculation and problems arise due to its dependence upon the conformation of the molecule.[115]

2.6.2 Gasteiger-Marsili and Gasteiger-Huckel Charges

The Gasteiger-Marsili[135] technique calculates atomic charges quickly and solely from information about the atoms present in the molecule and their connectivity. It uses

the concept of the partial equalisation of orbital electronegativity, $\chi_{\mu i}$, which was defined by Pauling as 'the power of an atom to attract electrons to itself'. A polynomial relationship between the electronegativity ($\chi_{\mu i}$) of an orbital (ϕ_μ) and the charge q_i on the atom i is assumed (30). Values for the coefficients a_μ and $b_{\mu i}$ were derived for common elements in their usual valence states. Electrons flow from the less electronegative elements to the more electronegative ones. This flow of electrons results in a positive charge on the less electronegative atoms and a negative charge on the more electronegative atoms, and as such the flow acts to equalise the electronegativities. This effect is modelled in the Gasteiger and Marsili approach by an iterative procedure where less and less charge is transferred between bonded atoms at each step.

$$\chi_{\mu i} = a_\mu + b_{\mu i} q_i + \chi_{\mu i} q_i^2 \qquad \textbf{(30)}$$

A more recent development is the Gasteiger-Huckel technique which uses a combination of two other charge computation methods - the Gasteiger-Marsili method to calculate the σ component of the atomic charge and the Huckel method to calculate the π component of the atomic charge. The Huckel portion of the charges is calculated first and these charges are used as the basis for the Gasteiger-Marsili charge calculation. After having performed a Huckel calculation, the actual number of electrons associated with atom i, q_i^π, can be calculated by equation 31. The total charge is the sum of the charges calculated by the two methods.

$$q_i^\pi = \sum_j^{MO} n_j c_{ij}^2 \qquad \textbf{(31)}$$

2.6.3 Charges derived from Electrostatic Potentials

A considerable effort has been made to develop methods in which charges are derived from the electrostatic potential (ESP). The ESP is the force acting on a unit positive charge placed at a certain point. The nuclei give rise to a positive (*i.e.* repulsive) force, while the electrons give rise to a negative potential. The objective of ESP methods is to derive the set of partial charges that best reproduces the QM electrostatic potential at a series of points surrounding the molecule, subject to the constraint that the sum of the charges is equal to the net charge on the molecule. ESP

charges are frequently considered superior to Mulliken charges as they depend much less on the underlying theoretical method used to compute the wavefunction (and thus the ESP).

The points where the potential is fitted are chosen to heavily represent regions for which it is more important to model intermolecular interactions correctly, such as just beyond the vdw radii of the atoms involved. Chirlian and Francl[136] developed the CHELP (Charges from electrostatic potentials) method, which uses spherical shells centred on each atom with points symmetrically distributed on the surface. Any points within the vdw radius of any atom in the system are discarded and the shells extend to 3 Å from the vdw surface of the molecule. This technique uses a Lagrangian multiplier method for fitting the atomic charges to electrostatic potentials.

The CHELPG method (Charges from Electrostatic Potentials using a Grid),[137] is a modification of the CHELP method. A cubic grid of points (spaced 0.3 - 0.8 Å apart) is used and all grid points that lie within the vdw radius of any atom are discarded, together with all points that lie further than 2.8 Å away from any atom. After evaluating the ESP at all valid grid points, atomic charges are derived that reproduce the ESP in the optimum way.

Alternatively, the Merz-Singh-Kollman (MK)[138,139] method produces charges fitted to the ESP at sets of concentric spheres about each atom. The layers are constructed as an overlay of vdw spheres around each atom. The best results are achieved by sampling points not too close to the vdw surface and the spheres radii are therefore modified through scaling factors. The smallest layer is obtained by scaling all radii with a factor of 1.4. The default MK scheme then adds three more layers constructed with scaling factors of 1.6, 1.8 and 2.0. After evaluating the ESP at all valid grid points located in all four layers, atomic charges are derived that reproduce the ESP as close as possible.

2.7 Solvation Methods

An important consideration in computational calculations of biological molecules is the nature of the solvent environment. The fact that biological processes take place in solution has posed significant challenges to the theoretical study of

proteins and nucleic acids.[140] Unfortunately, water is a deceptively difficult molecule to model accurately and much effort has gone into correctly modelling it.[141,142] Two general approaches are used namely classical ensemble treatments and quantum mechanical continuum models.

2.7.1 Explicit Water Models

In the simulation of macromolecules generally the solute is allowed to have flexible bonds and angles, whereas the water models use rigid geometries with fixed bond lengths and angles. Examples of explicit water models include, the three point transferable intermolecular potentials (TIP3P) of Jorgenson *et al.*[141] and the simple point charge (SPC) model of Berendsen *et al.*.[142] Both models use a total of three sites for the electrostatic interactions; the partial positive charges on the hydrogen atoms are exactly balanced by an appropriate negative charge located on the oxygen atom. The vdw interactions between two water molecules are computed using a Lennard-Jones function with just a single interaction point per molecule centred on the oxygen atom.[143] The TIP3P and SCP models differ slightly in the geometry of each water molecule, in the hydrogen charges and in the Lennard-Jones parameters utilised.

2.7.2 Implicit Water Models

The solvation free energy (ΔG_{sol}) is the free energy change to transfer a molecule from vacuum to a solvent and can be considered to have three components, ΔG_{elec}, the electrostatic component, ΔG_{vdw}, the vdw interaction and ΔG_{cav}, the free energy required to form the solute cavity within the solvent:

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{vdw} + \Delta G_{cav} \qquad \textbf{(32)}$$

ΔG_{elec} is the electrostatic component which is particularly important for polar and charged solutes due to the polarisation of the solvent. ΔG_{vdw} is the vdw interaction between the solute and solvent, which is in turn divided into a repulsive term ΔG_{rep} and an attractive dispersion term, $\Delta G_{disp.}$ ΔG_{cav} is the free energy required to form the solute cavity within the solvent. This comprises of the entropic penalty associated with the reorganisation of the solvent molecules around the solute together with the work done

against the solvent pressure in creating the cavity. The cavity and the vdw terms are often combined and represented using an equation of the form indicated in equation 33, where A is the total solvent accessible area and γ and b are constants that are usually taken from experimentally determined free energies for the transfer of alkanes from vacuum to water. [110]

$$\Delta G_{vdw} + \Delta G_{cav} = \gamma A + b \qquad \textbf{(33)}$$

The simplest model of solvation is to treat the solvent as a structureless, continuous medium. Generically, a solute molecule is firstly placed in solution. The solute displaces the solvent, thereby creating a solvent excluded volume. The boundary of this volume is the solute molecular surface, A. The region within the surface is assigned a dielectric constant ϵ of one. The rest of space is assigned the dielectric constant of the solution, 78.39 for aqueous applications. The solute charge density imposes an electric field on the surrounding continuum. This induces a solvent polarization that achieves equilibrium with the solute electric field. The subsequent solute-solvent interaction defines the electrostatic component of the solvation free energy, ΔG_{elec}.

2.7.3 Onsager Solvation Method

There is a number of approaches to solvation which differ in how the electrostatic term, ΔG_{elec}, is calculated. Important contributions to the study of solvation effects were made by Born[144] in 1920 and Onsager[145] in 1936. Born derived the electrostatic component of the free energy of solvation for placing a charge within a spherical cavity and Onsager extended this to a dipole in a spherical cavity. A fixed dipole moment of μ gives rise to an energy stabilization (34) where a is the cavity radius.[115]

$$\Delta G_{Onsager} = -\frac{\varepsilon - 1}{2\varepsilon + 1}\frac{\mu^2}{a^3} \qquad \textbf{(34)}$$

To develop the Self Consistent Reaction Field (SCRF) Onsager method, the back-polarisation of the medium is taken into account, as the dipole moment changes, depending on how polarisable the molecule is. Considering only the first-order effect

the stabilization is indicated in equation 35, where α is the molecular polarizability, the first-order change in the dipole moment with respect to an electric field.[115]

$$\Delta G_{Onsager} = -\frac{\varepsilon - 1}{2\varepsilon + 1}\frac{\mu^2}{a^3}\left[1 - \frac{\varepsilon - 1}{2\varepsilon + 1}\frac{2\alpha}{a^3}\right]^{-1} \quad \textbf{(35)}$$

The cavity size in the Onsager model strongly influences the calculated stabilization. The procedure requires that a somewhat arbitrary cavity radius, *a*, be assigned to solute molecules. A quantum mechanical recipe, proposed by Frisch and co-workers, is based on computing the 0.001 au electron density envelope of the gas phase geometry and applying a scaling factor of 1.33 to obtain an estimate of the molecular volume and finally adding a value of 0.5 Å to the calculated radius *a* in order to account for the nearest approach of solvent molecules.[146] This recipe has been shown to provide reasonable estimates of *a*, and has the advantage that no experimental information is required.

The Onsager model has the advantage of simplicity, rapid computation times and solute geometry optimisation. The main shortcomings of the Onsager model are:

- The spherical form of the cavity; molecules are rarely exactly spherical in shape.

- The description of the solute as a point dipole moment and its location at the centre of the sphere.

- The assumption of a fixed cavity radius, determined only by the nature of the solute and independent of the solvent.

- The lack of solvent structure in these models means it is not appropriate for solvents that have important specific interactions (such as HBs) with the solute.

2.7.4 Polarisable Continuum Models

The Polarisable Continuum Model (PCM) method, developed by Tomasi and coworkers,[147,148] forms a solute cavity from a union of spheres centred on each atom thus generating a more realistic cavity shape than utilised in the Onsager technique. PCM also includes a more exact treatment of the electrostatic interaction with the surrounding medium. The computational procedure divides this surface into small tesserae, on which the charges are evaluated. Typically, the spheres defining the cavity are taken to be 1.2 times the vdw radii.

2.8 Summary

When addressing problems in computational chemistry, the choice of computational scheme is dominated by the applicability of the method (*i.e.* the types of atoms and/or molecules that can be treated satisfactorily) and the size of the system to be investigated.[149] In this chapter, we have examined HF, DFT, semiempirical techniques and molecular mechanical potentials in terms of bond stretching $v_r(r_{ij})$, angle bending $v_\theta(\theta_{ijk})$, torsional $v_\phi(\phi_{ijkl})$, improper torsions $v_\chi(\chi_{ijkl})$ and non-bonding interactions $v_{nb}(r_{ij})$ terms. The cost of performing a HF calculation scales formally as the fourth power of the number of basis functions, M^4. Semi-empirical methods reduce the computational cost by reducing the number of integrals, from a formal order of M^4 to M^2 while the computational cost of DFT scales as M^2-M^3. The issue of cost versus accuracy must hence be born in mind when choosing a level of theory.

Finally, we examined a number of current methods to evaluate atomic charges based on ESP fits and include solvation effects in the calculations. One of the main concerns of using continuum models is the use of a dielectric constant (ϵ) at a microscopic level. The dielectric constant is a macroscopic property and is only meaningful for large groups of molecules. Also, the cost of simulations including correlation effects rapidly becomes prohibitively expensive. The application of continuum solvation modelling is, therefore, limited to small systems.

Chapter III: Computational Methodology

"Computers in the future may weigh no more than 1.5 tons."

Popular Mechanics (1939)

3.0 Introduction

Computational techniques, as explored in the previous chapter, are recognised as invaluable tools for revealing details of molecular conformation, dynamics and interactions between ligands and receptors. In this chapter, we will examine a number of computational applications, in the form of comparative modelling, optimisation techniques, conformational analysis, molecular dynamics simulations and docking techniques that are utilised in this study.

3.1 Protein Sequence Analysis and Structure Prediction

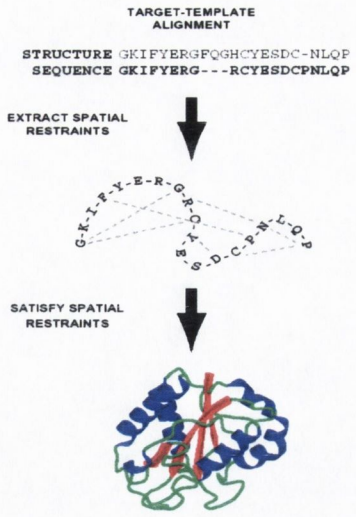
The experimental difficulties in obtaining protein crystal structures have lead to considerable interest in theoretical methods for protein structure prediction. The most ambitious approaches to theoretical *protein folding* attempt to solve it from first principles *i.e.* to explore the conformational space of the molecule in order to identify the most appropriate structure. A possible way to reduce the search in conformational space is to predict the 3D structure of a protein sequence by using information derived from a homologous protein of known structure. This methodology is known as homology or comparative modelling.

3.1.1 Comparative Modelling

Comparative modelling is based on the observation that the secondary structural units (such as α helices) of related proteins occupy the same relative orientations throughout a protein family. Thus, comparative modelling techniques use experimentally determined protein structures (templates) to predict the conformation of other proteins with similar amino acid sequences (targets). When determining if proteins are homologous or related, the target sequence is assigned a score relative to the target sequence. The simplest type of sequence score is the percentage sequence identity, which gives the percentage of amino acids that are identical in the two sequences. An alternative approach, the percentage sequence similarity, allows for 'conservative' substitutions, which are generally defined as amino acid replacements that preserve the structure and functional properties of proteins. For both approaches, a higher homology between the proteins results in a higher score.

The α_{1A}-AR modelling pursued in this work follows a three-step process consisting of sequence alignment, generation of spatial restraints and building of the receptor models, followed by subsequent evaluation of the models quality. Once each amino acid has been assigned to one of three secondary structure classes: α-helices, β-strand or loop region, the sequence alignment is used to determine equivalent residues between the target and the template. The objective of a sequence alignment algorithm is to match regions of amino acids that correspond to common structural or functional features. Global alignment algorithms attempt to match two sequences along their entire length *e.g.* Needleman and Wuncsh,[150] while local alignment algorithms align sections from the sequences *e.g.* Smith-Waterman.[151] Utilising multiple sequences of homologous proteins in the alignment procedure often improves the performance of secondary structure prediction as the algorithm is able to search for a consensus over the aligned sequences. Multiple sequence alignments can also suggest whether certain residues are conserved more frequently than others.

Once aligned, the template coordinates and a series of spatial restraints are used in conjunction with an optimisation procedure to derive a structure of the target protein (Scheme 3.1).[43,44] Typical spatial restraints include the distribution of distances between Cα atoms, residue solvent accessibilities or side-chain torsion angles. The restraints are expressed as probability density functions (pdf) and are combined to give a molecular function, which is then optimised. The procedure uses a combination of conjugate gradient optimisations, molecular dynamics simulations and simulated annealing, as shall be discussed in the following sections. In our work, one hundred models were generated for each alignment and the Modeller scores were examined.[197] The loop regions, restrained by the end positions of the helices were optimised independently using a molecular dynamics/simulated annealing procedure.



Scheme 3.1: Flowchart for comparative modelling by spatial restraints.[43]

The ability to then separate appropriate models of protein structures from incorrect models is of great importance for protein structure prediction methods. The conformation of an amino acid can be classified according to the torsion angles of its rotatable bonds (see Figure 3.1). There are three backbone torsion angles labelled ϕ (angle about the Cα-N bond), ψ (angle about the Cα-C bond) and ω (the amide bond). The amide bond has a relatively high energy barrier for rotation away from planarity and so ω rarely deviates from 0° or 180° with a significant preference for the trans (ω = 180°) conformation (except for Pro residues).

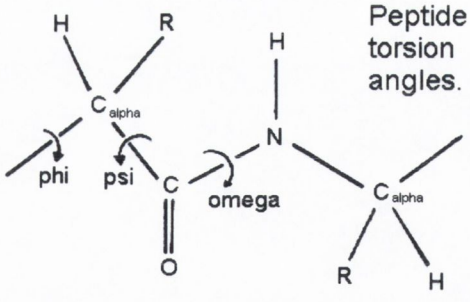


Figure 3.1: The three, backbone torsion angles labelled ϕ (angle about the Cα-N bond), ψ (angle about the Cα-C bond) and ω (the amide bond).[152]

A Ramachandran plot[153] is a contour map of energy as the backbone torsions ϕ and ψ are rotated. The sterically favourable combinations are the basis for preferred secondary structures, such as α helices or β sheets. Ramachrandran plots also indicate regions of steric conflict, which correspond to conformations where atoms in the polypeptide come closer than the sum of their vdw radii.

3.2 Structural Optimizations

Further refinement of the models can be pursued through structural optimizations. Such techniques gradually alter the coordinates of the atoms as they move the system closer to a minimum energy point. Alternatively expressed, given a function *f* which depends on one or more independent variables, x_1, $x_2 \ldots x_i$, one aims to find the values of those variables where *f* has a minimum value. The choice of minimisation technique is dictated by a number of factors including memory requirements, computational cost, speed and robustness of the method and a number of such methods shall now be discussed.

3.2.1 Simplex Optimization

A simplex is a geometrical figure with M + 1 interconnected vertices, where M is the dimensionality of the function. Following a simplex optimisation routine, three basic moves are possible on the potential energy surface, with the most common being a reflection (R) of the vertex with the highest value through the opposite face of the simplex, in an attempt to generate a new point that has a lower energy value, see Figure 3.2. If this new point is lower in energy than any other point in the simplex then a 'reflection and expansion' (E) move may be applied. If a 'valley floor' is reached, a reflection move will fail to produce a better point and the simplex contracts (C) along one dimension from the highest point. If this fails to reduce the energy another type of move is possible, in which the simplex contracts in all directions, pulling around the lowest point. Such a simplex optimization technique is implemented in the Dock 4.0 program, see section 3.7.2.

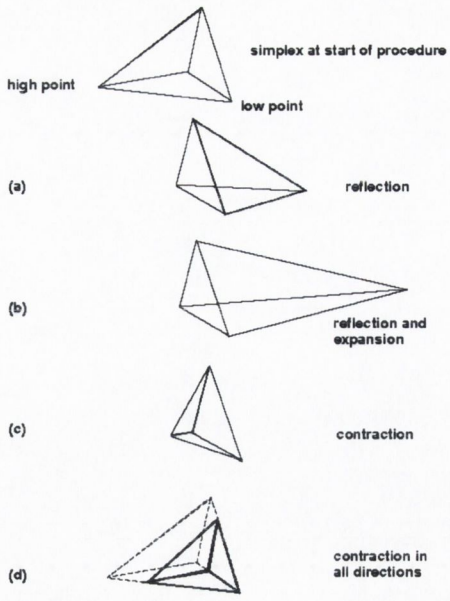


Figure 3.2: Simplex optimization; the position of the next point to be evaluated is indicated by either reflection (R), expansion (E), or contraction (C) operations.[154]

3.22 First Order Optimization Techniques

In general, first-order minimisation algorithms are preferred to the inefficient simplex method. The direction of the first derivative of the energy (the gradient, g) indicates the direction of maximum reduction in f and the magnitude of the gradient indicates the steepness of the local slope. The steepest descent method, over k iterations, moves in the direction parallel to the net force, *e.g.* in a geographical analogy this corresponds to walking downhill. For 3N cartesian coordinates this search direction is most conveniently represented by a 3N-dimensional unit vector, s_k.

$$s_k = -\frac{g_k}{|g_k|} \tag{36}$$

For the first iteration the starting point is the initial configuration of the system provided by the user, the vector x_1. The starting point for each subsequent iteration (k) is then the molecular configuration obtained from the previous step. The step size moved along the gradient unit vector s_k is given by γ_k. Hence, the new set of coordinates after step k would then be given by equation 37.

$$x_{k+1} = x_k + \gamma_k s_k \tag{37}$$

With the steepest descent technique two subsequent line searches are perpendicular to each other. If there were a gradient component along the previous

search direction the energy could be further lowered in this direction. There is therefore a tendency for each line search to partly spoil the function lowering obtained by the previous search. The steepest descent path oscillates about the minimum path and the rate of convergence is slow near the energy minimum.[115]

With the conjugate gradient technique, each line search is not along the current gradient but along a line constructed so that it is conjugate to the previous search direction. The gradients at each point are orthogonal but the directions are conjugate. A set of conjugate directions has the property that for a quadratic function of M variables, the minimum will be reached in M steps. The conjugate gradient method moves in a direction (v_k) from point x_k where v_k is computed from the gradient (g_k) at the point and the previous search direction (v_{k-1})

$$v_k = -g_k + \gamma_k v_{k-1} \tag{38}$$

γ_k is a scalar constant and there are several ways of choosing it such as that proposed by Polak-Ribiere (39).[110] Both optimization algorithms will alter the variables of the system until changes in the gradient and the structure on two successive iterations are smaller than prefixed values (convergence criteria).

$$\gamma_k = \frac{(g_k - g_{k-1}) * g_k}{g_{k-1} * g_{k-1}} \tag{39}$$

3.2.3 Second Order Optimization Technique

Further methods use both the first and second derivatives to locate an energy minimum as the second derivatives provide information about the curvature of the function. The Newton-Raphson method is the simplest example of a second order optimisation technique and requires knowledge of a Taylor series expansion about the point x_k:

$$V(x) = V(x_k) + (x - x_k)V'(x_k) + \frac{(x - x_k)^2 V''(x_k)}{2} + \ldots\ldots \tag{40}$$

The first derivative of $V(x)$ is:

$$V'(x) = xV'(x_k) + (x - x_k)V''(x_k) \quad \textbf{(41)}$$

If the function is purely quadratic, the second derivative is the same everywhere and so $V''(x) = V''(x_k)$. At the minimum ($x = x^*$) $V'(x^*) = 0$ and so

$$x^* = x_k - \frac{V'(x_k)}{V''(x_k)} \quad \textbf{(42)}$$

For a multidimensional function:

$$x^* = x_k - V'(x_k)V''^{-1}(x_k) \quad \textbf{(43)}$$

$V''^{-1}(x_k)$ is the inverse Hessian matrix of second derivatives. This algorithm uses the forces acting on the atoms of a given structure together with the Hessian matrix to predict energetically more favourable structures and thus optimise the molecular structure towards the next local minimum on the potential energy surface. The inversion of the Hessian matrix can be computationally demanding for systems of many atoms and can require significant amounts of storage. In practice, it is common to use a more robust method to get near to the minimum before applying the Newton-Raphson method.

The Berny algorithm[155] is the implementation of the Newton-Raphson method in the Gaussian 98 and Gaussian 03 programs. As explicit calculation of the second derivative matrix is quite costly, the Berny algorithm constructs an approximate Hessian at the beginning of the optimisation procedure through application of a simple valence force field and then uses the energies and first derivatives calculated along the optimisation pathway to update this approximate Hessian matrix. The success of the optimisation procedure therefore depends to some degree on how well the approximate Hessian represents the true situation at a given point. The algorithm will vary the

coordinates of the system until changes in the gradient and the structure on two successive iterations are smaller than prefixed values (convergence criteria).

3.3 Conformational Searches

The physical, chemical and biological properties of a molecule often critically depend upon the 3D structures, or conformations, that it can adopt. Hence, another important aspect is the conformational flexibility of these compounds and its influence on their properties. The energetic surface of a molecule can be rugged, containing multiple local minima along with the global minimum. The objective of a conformational search is to sample the energy surface and identify the energetically 'preferred' conformations of a molecule. Such methods can be conveniently divided into two categories: systematic and random search algorithms.

3.3.1 Systematic Conformational Search

In a systematic search, all rotatable bonds are identified and each is systematically rotated through 360° using a fixed increment, while the bond lengths and angles remain fixed throughout. For each conformation, every atom pair in the molecule is examined to determine if they are in steric conflict.[156] Conformations are also subjected to a structural optimization and higher energy conformations are eliminated. A conformational search finishes when all possible combinations of torsion angles have been generated and minimised. However, the number of conformations required to adequately sample the entire conformational space of a molecule becomes extremely large. If A is the torsion angle increment and T is the number of rotatable bonds in the molecule, then the total number of possible conformers to be examined is given by:

$$V = \left(\frac{360}{A}\right)^{T} \qquad \textbf{(44)}$$

e.g. a molecule with six rotatable bonds at 30° increments generates 2,985,984 conformations, a phenomenon known as *combinatorial explosion*. The number of conformations sampled can be reduced by, limiting the range of rotation for a

symmetric substituent (*e.g.* 0 to 180° for a phenyl group) and/or increasing the rotation step size.

3.3.2 Random Conformational Search

With a random conformational search, a random change is applied to the current conformation and the system moves from one region of the potential energy surface to another in a single step. A random search can explore conformational space by changing either the atomic cartesian coordinates or the torsion angles of rotatable bonds. Each resulting conformation is optimised using a conjugate gradient method.[157] If the particular conformation has not been found previously, it is stored. The procedure continues until a given number of iterations have been performed or until it is decided that no new conformations can be found.

3.4 Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations enable the time-dependent behaviour of atomic and molecular systems to be examined.[158] The foundation for MD simulations is Newton's second law of motion, which states that a body's' acceleration equals the net force divided by its mass. From the force acting on the atoms, one can determine the accelerations of the particles. A finite difference method such as the velocity verlet (VV)[159] method is utilised to integrate Newtons' laws of motion, leading to trajectories in space and time. The initial velocity vector is set pseudorandomly so that the total kinetic energy of the system corresponds to the expected value at the target temperature. Given the positions $r(t)$, velocities $v(t)$ and forces $F(t)$ at time t, the propagation of the trajectory in the VV scheme at time $(t+h)$ is obtained via equation 46

$$r(t+h) = r(t) + h\,v(t) + \frac{h^2 F(t)}{2m} \qquad \textbf{(45)}$$

Calculation of forces $F(t+h)$

$$v(t+h) = v(t) + \frac{h[F(t) + F(t+h)]}{2m} \qquad \textbf{(46)}$$

The VV scheme allows for the generation of positions and velocities of the integrated system at the same time. The choice of time step, h, is important as if it is set too small it would only cover a limited proportion of the phase space. Conversely, if h is too large the system would not sample all vibrations and may miss bond stretching and bending which occur on a femtosecond (fs) to picosecond (ps) time scale.

3.4.1 Molecular Ensembles

During a MD simulation an average of conformations (an ensemble) is generated. Traditionally a constant NVE (constant Number of particles, constant Volume and constant Energy) microcanonical ensemble is utilised. Simulations in a canonical NVT (constant Number of particles, constant Volume and constant Temperature) or an isobaric-isothermal NPT (constant Number of particles, constant Pressure and constant Temperature)[160] ensemble are needed when controlling temperature and pressure of realistic systems. In such situations equations of motion that produce the statistical behaviour in agreement with the expected distribution function are used.

The most common approach for constant temperature MD simulation is using an extended system in which the whole simulation is placed in contact with a heat bath. Such a method was originally proposed by Nosé, in which the equations of motion were modified by having a variable length time step. However, this technique was difficult to implement and the approach was reformulated by Hoover, who removed the variable timestep and introduced a friction factor. This heat bath has its own potential and kinetic energy terms and an associated degree of freedom, which is introduced into the systems Hamiltonian for which the equation of motion can be derived.

Many of the methods used for pressure control in the constant temperature and isobaric ensembles (NPT) are analogous to those used for temperature control. The pressure can be maintained at a constant value by simply scaling the volume. In the extended pressure-coupling system methods, an extra degree of freedom, corresponding to the volume of the box is added to the system.

3.4.2 Molecular Dynamics Protocol

In this study, MD simulations with a time step of 1 fs were performed using the standard parameters of the Amber force field[161] for the receptor and parameters from the General Amber Force Field (GAFF)[162] for the ligands. Periodic boundary conditions (PBC) were applied in all three dimensions, which ensure that the system does not have an abrupt border with vacuum by surrounding each box by images of itself. The "nearest neighbour" convention is employed with PBC, in which an atom is assumed to interact only with the nearest images of any other atom. The direct evaluation of the non-bonded interactions involving all atom pairs has a complexity of N^2, where N is the number of atoms. To ease the computational expense the number of short range atom pairs can be significantly reduced by using a cut-off distance of 9 Å beyond which, the atoms are no longer considered to interact. Non-bonded neighbour[163] lists are utilised to store all atoms within the cutoff distance and the interaction only calculated between these atoms. The array is updated at regular intervals; if the update frequency is too high the procedure is inefficient; if too low the energies and forces may be calculated incorrectly due to atoms moving within the non-bonded cutoff. A distance of 10 Å was added from the edge of the box to the closest solute molecule to ensure a minimum distance of 20 Å between adjacent proteins in the lattice, with non-bonded interactions calculated for 1-4 interactions and higher. The Particle Mesh Ewald (PME) method is used to treat the long-range electrostatic interactions.

Each simulation cell, prior to MD, was optimised to remove bad contacts by performing 250 steps of steepest descent followed by 750 steps of conjugate gradient energy minimisation. The simulation cell was heated gradually to 300 K over 5 ps. It is necessary to run the system holding the protein fixed with positional restraints so that the solvent can come to equilibrium. A force is applied to fix heavy atom coordinates on predefined positions. Equilibration is performed using backbone restraints for 5 ps at each of 15, 10 and 5 kcal mol^{-1} followed by 65 ps without restraints before commencing a production phase. Data was collected at one ps intervals over the production runs.

3.4.3 Molecular Dynamics Trajectory Analysis Tools

MD simulations generate configurations of the system that are connected in time and so can be used to calculate time dependent properties. The most common measure of the fit between two structures is the root mean square deviation (RMSD) between

pairs of atoms. RMSDs are frequently calculated for the backbone heavy atoms after fitting to the initial reference structure. The RMSD is given by equation 47, where $r_i(t)$ is the position of atom *i* at time *t* and N is the number of atoms.

$$RMSD(t_1 t_2) = \left[\frac{1}{N} \sum_{i=1}^{N} m_i \mid r_i(t_1) - r_i(t_2) \mid^2 \right]^{\frac{1}{2}} \quad \textbf{(47)}$$

The radius of gyration (R_g) yields a measure of the compactness of the receptor structure. In Equation 48, m_i is the mass of atom *i* and r_i is the position of atom *i* with respect to the centre of mass of the molecule. R_g tends to increase as the protein unfolds, *i.e.* as the protein spreads out from its centre.

$$R_g = \left(\frac{\sum \mid r_i \mid^2 m_i}{\sum m_i} \right)^{\frac{1}{2}} \quad \textbf{(48)}$$

Finally, hydrogen bonds (HBs) were defined geometrically; the angle donor-acceptor-hydrogen distance has to be less than 60° and the donor-acceptor distance (d [donor...acceptor]) must be less than 4.0 Å.

3.5 Thermodynamic Corrections

The QM or MM energy at a minimum corresponds to a hypothetical, motionless state at 0 K. However, experimental measurements are made on molecules at a finite temperature where the molecules undergo translational, rotational and vibrational motion. To compare theoretical and experimental results it is necessary to make corrections to allow for these motions, which are calculated using standard statistical mechanical formulae.[164] The internal energy U(T) at a temperature T is given by:

$$U(T) = U_{trans}(T) + U_{rot}(T) + U_{vib}(T) + U_{vib}(0) \quad \textbf{(49)}$$

where U_{trans} is the translational energy, U_{rot} is the rotational energy and U_{vib} is the vibrational energy, which is broken into two components, $U_{vib}(0)$ at 0K and $U_{vib}(T)$ at a temperature T. If all translational and rotational modes are fully accessible in accordance with the equipartition theorem, then $U_{trans}(T)$ and $U_{rot}(T)$ are both equal to

$\frac{3}{2}k_BT$ per molecule, where k_B is Boltzmanns constant. The vibrational contribution equals the difference in the vibrational enthalpy at the temperature T and at 0 K, and it is summed over all vibrational frequencies ν_i, equation 50.

$$U_{vib} = \sum_i \left[\frac{1}{2} + \frac{1}{e^{\frac{\hbar\nu_i}{kT}} - 1} \right] \hbar\nu_i \qquad \textbf{(50)}$$

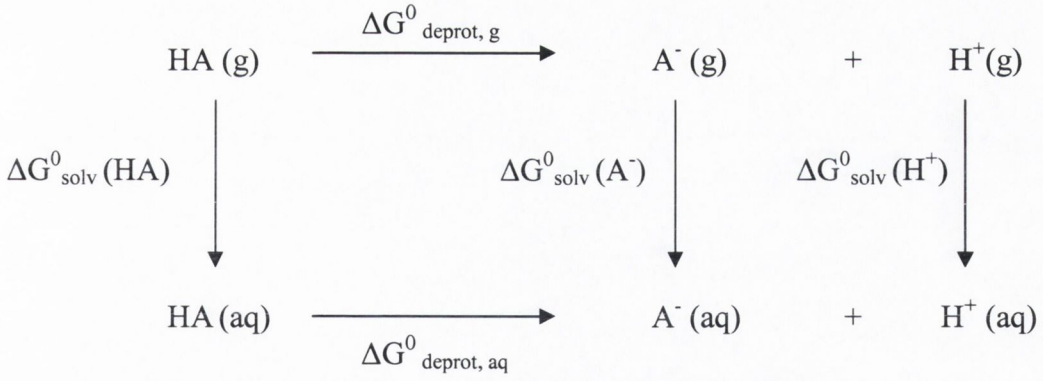
The vibrational free energy correction is given by equation 51, where the first term is the zero-point correction; and the second term corrects for the average thermal population of vibrational levels at the temperature *T*.

$$A_{vib} = \sum_i \left[\frac{\hbar\nu_i}{2} + kT \ln\left(1 - e^{\frac{-\hbar\nu_i}{kT}} \right) \right] \qquad \textbf{(51)}$$

Theoretical frequencies generally overestimate the observed experimental fundamentals due to the incomplete treatment of electron correlation, the neglect of mechanical anharmonicity and basis set truncation effects. The overestimation is generally uniform and to improve the agreement between the predicted and observed frequencies the computed harmonic frequencies are often scaled for comparison.[165]

3.6 Theoretical pKa Calculations

A computational methodology for pKa predictions of small molecules can be based on an *ab initio* description of a generalised thermodynamic cycle (see Scheme 3.2), relating pKa to the gas-phase proton basicity ($\Delta G^0_{deprot,g}$) via the solvation energies ($\Delta G^0_{deprot,aq}$) of the products and the reactants. Calculation of the pKa of an acid, HA, involves QM calculations to characterise the gas-phase system ($\Delta G^0_{deprot,g}$) for both the associated acid, HA(g) and the dissociated ions, H^+(g) and A^-(g) and to characterise the solvated system ($\Delta G^0_{deprot,aq}$) for the associated HA(aq) and dissociated H^+(aq) and A^-(aq) by coupling to continuum solvent using a self-consistent reaction field cycle.



Scheme 3.2: Thermodynamic cycle used in the calculation of theoretical pKas.

The pKa of HA(aq) is given by

$$pK_a = \frac{1}{2.303RT}\Delta G^0_{deprot,aq} \quad \textbf{(52)}$$

where

$$\Delta G^0{}_{deprot,\,aq} = \Delta G^0(A^-(aq)) + \Delta G^0(H^+(aq)) - \Delta G^0(HA(aq)) \quad \textbf{(53)}$$

$$\Delta G^0{}_{deprot,\,aq} = \{\Delta G^0(A^-(g)) + \Delta G^0{}_{solv}(A^-)\} + \{\Delta G^0(H^+(g)) + \Delta G^0{}_{solv}(H^+)\} - \{\Delta G^0(HA(g)) + \Delta G^0{}_{solv}(HA)\} \quad \textbf{(54)}$$

$$\Delta G^0{}_{deprot,\,aq} = \Delta G^0{}_{deprot,\,g} + \{\Delta G^0{}_{solv}(A^-) + \Delta G^0{}_{solv}(H^+) - \Delta G^0{}_{solv}(HA)\} \quad \textbf{(55)}$$

In terms of the theoretical calculations, two problematic issues remain, namely determining the free energy of a proton in the gaseous phase, $\Delta G^0_g(H^+)$ and in a solvated phase, $\Delta G^0_{solv}(H^+)$. $\Delta G^0_g(H^+)$ is given by,

$$\Delta G^0_g(H^+) = 2.5RT - T\Delta S = -6.28 \text{ kcal mol}^{-1} \quad \textbf{(56)}$$

Alternatively, there is no accurate determination experimental or otherwise, for $\Delta G^0_{solv}(H^+)$ in water. Experimental values range from -254 to -262.5 kcal mol^{-1}.[166] Theoretical studies also report a considerable range of values between -244.9 and -264.0 kcal mol^{-1}.[167] This limits absolute pKa determinations theoretically while the value of $\Delta G^0_{solv}(H^+)$ is not an issue when considering the relative pKa values.

3.7 Receptor Binding Sites

More than 100 years ago, Emil Fischer wrote that "enzyme and glycoside must fit together like a key and a lock in order to initiate a chemical action upon each other."[168] Subsequently, the target protein is often denoted as a *lock* with one or a few cavities in which a ligand or *key* may fit. Given the preponderance of cavities that may act as binding site locations a series of computational tools have been developed to detect potential binding regions. PASS (Putative Active Sites with Spheres)[169] is a simple tool that uses geometry to characterise regions of buried volume in proteins and to identify positions likely to represent binding sites based upon the size, shape and burial extent of these volumes. The PASS algorithm is designed to fill the cavities in a protein structure with a set of spheres and to identify a few of those spheres (called "active site points", ASPs) that most likely represent the centres of binding pockets. Once a receptor structure has been examined for potential binding sites, the interactions of ligands with the protein can be investigated.

3.7.1 Receptor-ligand Complex Study

In molecular docking, one attempts to predict the structure of the intermolecular complex formed between two molecules. The objective of a docking calculation is to score the interaction energies (IE) of many orientations of one molecule relative to another while searching for the orientations that result in low interaction energies *i.e.* the preferred binding mode of a ligand with a protein.[170] To this end, a number of methods have been developed which can be characterised according to the number of degrees of freedom that they ignore. The simplest algorithms treat the two molecules as rigid bodies and explore only the six degrees of rotational and translational freedom. Development of further search algorithms has sought to include limited treatment of molecular flexibility for the receptor[171] and the ligand.[172,173]

3.7.2 Docking Applications

A well-known example of a docking algorithm is that in the Dock 4.0 program of Kuntz and coworkers.[170] It has been successfully used to generate lead compounds for a number of important biological targets, including the human immunodeficiency virus (HIV)-1 protease,[174] dihydrofolate reductase[175] and hemagglutinin.[176] Dock 4.0

was designed to find molecules with a high degree of shape similarity to the binding site and thus it requires a topological 'map' of the active site to be created. A molecular surface of the binding region is obtained using the MSMS[177] (Michael Sanners Molecular Surface) software and a negative image produced. This negative image consists of a collection of overlapping spheres of varying radii, each of which touches the molecular surface at two points. A typical incremental construction algorithm then identifies one or more 'base fragments' within the ligand, often chosen as a rigid part of the molecule such as a ring system. Based on the location of rotatable bonds the molecule is divided into rigid and overlapping segments. Each docked orientation of the base fragment then represents the starting point for the conformational analysis of the rest of the ligand. Each conformation is optimised using the simplex technique (section 3.2.1). The ligand orientation is examined to ensure there are no unacceptable steric interactions between the ligand and the receptor. If the orientation is acceptable, an interaction energy (IE) is computed utilising the Amber forcefield to give the 'score' for the binding mode. The receptor-ligand intermolecular energy is often pre-computed using an underlying 3D grid, (with a grid spacing of 0.3 Å), constructed to enclose all site points plus an extra 5 Å margin along each axis. For each ligand atom the receptor-ligand IE is calculated from the nearest eight surrounding grid points using an interpolation method. The top-scoring orientations are retained for subsequent analysis.

Alternatively, the FlexiDock[108] software incorporates flexibility to a number of the receptor side chains as well as to the ligands. The FlexiDock routine utilises a genetic algorithm to determine the optimum ligand geometry through mimicing evolutionary behaviour. A solution to a problem is encoded in a chromosome and a fitness score is assigned to it based on the relative merit of the solution. Each chromosome consists of a number of 'genes' one for each torsional angle. Both the orientation and the internal conformation will vary as the populations evolve. A scoring function rates each chromosome and competition between chromosomes yields a set of results. A population of chromosomes then goes through a process of evolution in which only the fittest solutions survive. Flexidock takes into account not only the position and conformation of the ligand but also the hydrogen bonding network in the binding site.

In this study, charges for the ligands were taken from the electrostatic potential obtained during DFT structural optimisations using the B3LYP hybrid functional and

the 6-31G* basis set in Gaussian 98.[178] For the protein, Kollman charges were added. Initial docking was performed with the Dock 4.0 software. The FlexiDock software was utilised to perform flexible docking of ligands into receptor binding sites. The default docking parameters were used. After docking, MD simulations of the receptor-ligand complexes were performed and the stabilising interactions of the complexes were examined.

3.7.3 Development of Molecular Mechanical Parameters

When performing MD simulations, it is necessary to assign a MM atom type for each atom in the system, which contains information about its hybridisation state, *e.g.* to distinguish between sp^3-hybridised carbon atoms (which adopt a tetrahedral geometry) and sp^2-hybridised carbons (which are trigonal). The reliability of a force field crucially depends on the accuracy of these parameters. While much effort has focused on the parameterisation of force fields for protein or nucleic acid simulations, problems can arise when considering small organic ligands that may bind to these proteins. It is often necessary to add to the force field parameter set and such an addition necessitates novel bond, angle, dihedral and Lennard-Jones parameters. Parameter fitting involves a reference data set (training set), a set of appropriate potential energy functions and a method to quantitatively compare the computed with the experimental structural data (*e.g.* through a least-squares fit of the deviation between computed and experimental bond lengths, angles etc.).

In this study, a parameter fitting regime was introduced by comparing the additional bond and angular distances in the optimised structures to experimental crystal structures and subsequently fitting the force constants to theoretically determined frequencies (B3LYP/6-31G*). A uniform overestimation of QM harmonic vibrational frequencies has been found and as a result generic frequency scaling factors are often applied. A scale factor of 0.9614 was recommended by Scott *et al.*,[179] for the B3LYP/6-31G* level of theory and was utilised in this work.

3.7.4 Interaction Energies, Binding Energies and Binding Affinities

In this work, MD simulations were performed on a series of ligand/receptor complexes, which were averaged over the last 200 ps of simulation and optimised. The interaction energies (IEs) and binding energies (BEs) of the complexes were then

estimated using the Amber FF. The IEs are evaluated for the receptor and ligand in their bound conformations where E(complex) is the energy of the complex, $E(\alpha_{1A}\text{-}AR_{bound})$ is the energy of the receptor in its bound state and $E(ligand_{bound})$ is the energy of the ligand in its bound state.

$$IE = E(complex) - [E(\alpha_{1A}\text{-}AR_{bound}) + E(ligand_{bound})] \quad \textbf{(57)}$$

Alternatively, the BE of the ligand/receptor complexes takes into account the energy of the conformational change of the ligand and receptor upon complexation and is calculated as the difference between the energies of the bound complex and of the free optimized molecular forms of the receptor and of the ligand. The BE is given by equation 58, where E(complex) is the energy of the complex, $E(\alpha_{1A}\text{-}AR_{uncomplexed})$ is the energy of the receptor in its uncomplexed state and $E(ligand_{uncomplexed})$ is the energy of the ligand in its uncomplexed state.

$$BE = E(complex) - [E(\alpha_{1A}\text{-}AR_{uncomplexed}) + E(ligand_{uncomplexed}) \quad \textbf{(58)}$$

Another form of scoring function are empirically based methods such as XScore.[180] Empirical scoring functions are directly calibrated with a set of protein-ligand complexes with experimentally determined structures and binding affinities (BA) through multivariate regression analysis. This scoring function takes the following form:

$$pK_d = K_0 + K_{vdw} + K_{metal} + K_{HB} + K_{desolvation} + K_{deformation} \quad \textbf{(59)}$$

K_0 is the regression constant which attempts to take account of the translational and rotational entropy loss upon binding. K_{vdw} represents the contribution of vdw interaction between the protein and its ligand, K_{metal} is the contribution of metal-ligand bonding, K_{HB} is the contribution of hydrogen bonding, $K_{desolvation}$ is the contribution of the desolvation effect and $K_{deformation}$ is the contribution due to the deformation effect.

In this chapter, we have outlined all the computational techniques that will be used in this study. We now proceed to the development of the α_{1A}-AR models (Chapter IV), analysis of the ligand compounds (Chapter V) and the study of the receptor-ligand complexes (Chapter VI-VII).

Chapter IV: α_{1A}-Adrenoceptor Models

"Could the search for ultimate truth really have revealed so hideous and visceral looking an object."

Max Perutz, 1964 on protein structure.

4.0 Introduction

The development of homology models of the α_{1A}-AR,[37] allows for structural analysis of the receptor at an atomic level. Although a number of α_{1A}-AR homology models currently exist,[181] they are limited by the suitability and quality of the crystal structure template used. Early α_{1A}-AR homology models[182] were based on the proton pump bacteriorhodopsin, which has a low sequence identity of 11.9 % with the α_{1A}-AR and a similarity of 18.9 %. Although bacteriorhodopsin belongs to a family of 7-TM receptors it is not a member of the GPCR family and hence is not an appropriate template for homology modelling of Class A GPCRs.[183] In 1993, a low resolution, projection density map of bovine Rhodopsin, a Class A GPCR, was obtained by electron cryo-microscopy.[184] From this a model of the Rhodopsin TM core, consistent with the results of a great number of mutagenesis and biochemical studies was developed. Bovine Rhodopsin and α_{1A}-AR have a 17.4 % identity and a similarity of 31.4 %. Furthermore, in 2000, a crystal structure of bovine Rhodopsin (denoted *1f88* in the Brookhaven Protein Data Bank[185]) was elucidated at a 2.8 Å resolution and its frequent use in homology modelling of GPCRs has been reviewed elsewhere.[91] Ongoing crystallographic refinement of bovine Rhodopsin generated the *1hzx* structure at a 2.8 Å resolution,[186] which was further refined in the *1l9h* structure to 2.6 Å.[187] Class A GPCRs consist of a transmembrane heptahelical bundle (TM I-VII), transversing the cell membrane with the helices connected by intracellular (IC) and extracellular (EC) loops (see Figure 4.1).

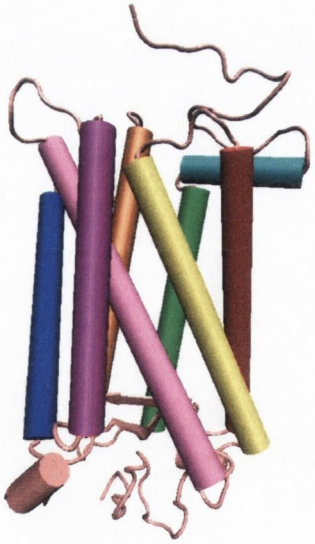


Figure 4.1: Bovine Rhodopsin crystal structure (*1l9h*) illustrating TM-I (red), TM-II (yellow), TM-III (mauve), TM-IV (purple), TM-V (blue), TM-VI (orange), TM-VII (green), TM-VIII (cyan) and loop regions (pink).

The differences between the *1f88* and the *1hzx* rhodopsin structures are mainly located in the cytoplasmic side of Rhodopsin. The *1l9h* structure is quite similar to *1hzx*, with a RMSD of 0.36 Å between Cα atoms.[188] Further experimental refinement resulted in the *1gzm* structure in 2002 and the *1u19* structure in 2004 at a 2.2 Å resolution.

In this present study, to advance our understanding of the molecular structure of the α_{1A}-AR we have built homology models based on the *1l9h*[189] crystal structure of bovine rhodopsin in the inactive state.[190] Once developed, α_{1A}-AR homology models should be further structurally refined, through structural optimizations and/or MD simulations. Pedretti *et al.*[191] performed MD simulations on an α_{1A}-AR model *in vacuo*, which necessitated imposing a series of harmonic restraints on the TM helix backbone. However, a membrane protein is naturally exposed to a heterogeneous environment, where the TM regions are embedded in a phospholipid bilayer and the extramembrane domains are surrounded by water. The restrictions imposed by the lipid bilayer environment on a membrane proteins structure and dynamics are also important for the determination of tertiary structure and should be taken into consideration. Although refining the receptor in its natural lipid environment may result in a more realistic simulation, as we are starting from a homology model rather than a crystal structure, we are sensitive to the possibility that the slow time scale of lipid motion would restrict the receptors mobility. This could mask large-scale conformational drift that might provide evidence for an incorrect receptor fold in the model. In order to reduce the overhead in simulation time, but maintain the biphasic environment, the dynamics of the receptor models were examined in a phase-separated box acting as a membrane mimic. To investigate the effects of different hydrophobic environments on receptor dynamics, simulations in a $H_2O/CHCl_3/H_2O$ and a $H_2O/CCl_4/H_2O$ membrane mimetic were compared.

Comparative MD simulations were then performed on the homology models and the on the *1l9h* rhodopsin structure. As the structural basis for conformational change may lie in the orientation and interactions of the TM helices, a number of features were examined for both the initial homology models and for the average structures after the 1 ns MD simulations.

4.1 Current α_{1A} adrenoceptor models

There are three homology-derived models of the α_{1A}-AR currently deposited in the GPCR database (GPCRDB).[181] The first model[192] was based on a bacteriorhodopsin structure, the second model[193] on the Cα rhodopsin structure of Baldwin *et al.*,[184] and the third model[194] on the *1f88* bovine Rhodopsin structure (see Figure 4.2). All three structures are of the seven TM helices only and do not include loops, tail regions or the short helical segment following TM-VII, termed TM-VIII. This latter helix is predicted for most GPCRs although it remains unknown if it is necessary for structural or functional integrity of GPCRs. Differences in the three helical bundles can be observed specifically in TM-IV where a kink is observed in the Baldwin model, which is absent in the other two models. Construction of structurally variable loop regions is a challenging problem and was not attempted by the depositors.

Figure 4.2: The three α_{1A}-AR models deposited in the GPCRDB.[181] One model is based on the bacteriorhodopsin structure (on left),[192] another model on the Baldwin rhodopsin structure (in centre)[193] and another on the *1f88* bovine Rhodopsin structure (on right).[194] In the figure, TM-I is represented in red, TM-II in yellow, TM-III in mauve, TM-IV in purple, TM-V in blue, TM-VI in orange and TM-VII in green.

The stereochemical analysis (considering the non Gly and non Pro residues) of all three available structures is satisfactory. The bacteriorhodopsin-based model has a total of 206 residues and of the non Gly and Pro residues, none are in sterically disallowed regions, while 154 residues (90.6 %) are in the most favoured regions and 16 residues (9.4 %) are in the generously allowed regions. The rhodopsin α-carbon based model has a total of 184 residues, of which 147 (97.4 %) are in most favoured regions, two (2.0 %) in allowed regions and one (0.6 %) in a disallowed region. Finally, the rhodopsin-*1f88*, based model has a total of 190 residues, with 142 residues (91.6 %) in the most favoured regions and 13 (8.4 %) in generously allowed regions.

As a complete α_{1A}-AR structure is necessary for refinement using MD, we aimed to develop a novel α_{1A}-AR model complete with TM-VIII and loop regions. Recently, complete α_{1A}-AR homology models were developed based on the *1f88* rhodopsin structure,[191,195] but these structures were not available for comparison. In our homology modelling study, the *1l9h* bovine Rhodopsin structure was utilised as the template, as it was the most recent and detailed available at the start of the work.

4.2 Sequence Alignment of Bovine Rhodopsin to the Human α_{1A}-AR

The pursued receptor modelling routine utilized sequence and structure similarities with bovine Rhodopsin for predicting the unknown α_{1A}-AR structure. Sequences of Rhodopsin (PO2699) and the human α_{1A}-AR (P35348) were obtained from the GPCRDB that holds data from the Swiss-Protein and the TrEMBL databanks. The sequence homology within Class A GPCRs is relatively low (in many cases less than 35 % identity with Rhodopsin). Bovine Rhodopsin and α_{1A}-AR have a 17.4 % identity and a similarity of 31.4 %, which is in the 'twilight zone' (less than 30 % identity and 40 % similarity)[43] for homology modelling. When the sequence identity between the target and the template falls below 30 %, accurate sequence alignment is the main difficulty in the homology modelling procedure. More reassuring is the higher identity (18-29 %) observed between the helical regions of the α_{1A}-AR with rhodopsin, TM-III being slightly lower at 11.1 %.[196,197] As related domains are conserved more strongly than amino acid sequences, the presence of highly conserved motifs identifies evolutionary relationships among the Class A receptors.[198] Class A GPCRs share key structural features such as a disulfide bond between TM-III and the extracellular region and a tripeptide Glu/Asp-Arg-Tyr (E/DRY) motif located at the intracellular end of TM-III. There are several other highly conserved residues, such as an Asn-Asp pair located in TM-I and TM-II, respectively, aromatic residues in TM-IV and TM-VI and a common Asn-Pro-X-X-Tyr (NPxxY) motif in TM-VII. In GPCRs, the conservation of Pro residues in TM-V and TM-VI also play key roles in the alignment and modelling of the TM domains.

In this study, two approaches were followed with regards to sequence alignment to the bovine Rhodopsin template. In Model I, helical regions were aligned through a pairwise alignment algorithm, which matches identical/similar amino acids. In Model II, a multiple sequence alignment was performed on 298 sequences of Class A GPCRs,

using the algorithm ClustalW.[47] Multiple alignments involve aligning sequences of a number of proteins with each other and thus identifying conserved regions between all of the alignments. As the average sequence identity in Class A GPCRs is low, multiple sequence alignment programs may not be ideal in aligning this family and in this study models were derived from both approaches. In the alignment process, 365 residues for Model I and 369 residues for Model II of α_{1A}-ARs 466 residues were considered, as the remaining residues lie in the C terminus tail and could not be aligned suitably (bovine Rhodopsins full sequence having only 348 amino acids). In both alignments, the conserved residues in the helical regions are aligned equivalently with the bovine Rhodopsin sequence (Table 4.1).

The main differences in the two sequence alignments emerge in the loop regions, which are more divergent in sequence and probably in 3D structure than the helical regions. It is difficult to accurately model loops by homology, as most can not be seen in an electron density map and those conformations observed may be induced by crystal packing forces. Also, the sequence identity between most GPCRs and bovine Rhodopsin is too low to derive any reliable loop alignment. In this work, the protein database of structures was searched for regions of a suitable sequence, length and geometry at the interface with the structurally conserved regions (SCRs). Upon examination of the determined structures, none were identified having a sequence identity of 20 % or greater, which formed a loop between antiparallel helices. Hence, no loop structures, other than a portion of the rhodopsin structure were utilised in modelling.

From the two sequence alignments, 3D models containing all non-hydrogen atoms were obtained using Modeller.[43,199] One hundred models were generated and the 'best' model was selected by choosing that with the lowest value of the Modeller scoring function.[44] The basic goal of loop modelling is to predict the conformation of a loop that is fixed at both ends of the receptor helical backbone. The loop regions, excluding IC-III at 67 residues, were optimised independently many times using a thorough molecular dynamics/simulated annealing procedure. Twenty-five models were generated by this method. The resulting molecular models were examined and the routine of realignment, modelling and structure validation was repeated until no further improvements to the receptor models were observed.

Table 4.1: TM helical sequence alignment (TM-I to TM-VII) between the bovine Rhodopsin template (*1l9h*) and the human α_{1A} adrenoceptor (*a1aa*). Identical residues are indicated in red and similar residues in blue.

TM-I	**1l9h**	F	S	M	L	A	A	Y	M	F	L	L	I	M	L	G	F	P	I	N	F	L	T	L	Y	V	T	V
	A1aa			L	L	G	V	I	L	G	G	L	I	L	F	G	V	L	G	N	I	L	V	I	L	S	V	
TM-II	**1l9h**		Y	I	L	L	N	L	A	V	A	D	L	F	M	V	F	G	G	F	T	T	T	L	Y			
	A1aa				I	V	N	L	A	V	A	D	L	L	L	T	S	T	V	L	P	F	S	A	I	F	E	V
TM-III	**1l9h**		N	L	E	G	F	F	A	T	L	G	G	E	I	A	L	W	S	L	V	V	L	A	I	E	R	Y
	A1aa	C	N	I	W	A	A	V	D	V	L	C	C	T	A	S	I	M	G	L	C	I	I	S	I	D	R	Y
TM-IV	**1l9h**		A	I	M	G	V	A	F	T	W	V	M	A	L	A	C	A	A	P	P	L	V					
	A1aa			L	M	A	L	L	C	V	W	A	L	S	L	V	I	S	I	G	P	L	F	G	W	R	Q	
TM-V	**1l9h**		F	V	I	Y	M	F	V	V	H	F	I	I	P	L	I	V	I	F	F	C	Y	G				
	A1aa		Y	V	L	F	S	A	L	G	S	F	Y	L	P	L	A	I	I	L	V	M	Y	C				
TM-VI	**1l9h**		V	T	R	M	V	I	I	M	V	I	A	F	L	I	C	W	L	P	Y	A	G	V	A	F	Y	
	A1aa		L	G	I	V	V	G	C	F	V	L	C	W	L	P	F	F	L	V	M	P	I	G	S	F		
TM-VII	**1l9h**		F	M	T	I	P	A	F	F	A	K	T	S	A	V	Y	N	P	V	I	Y	I	M				
	A1aa		V	F	K	I	V	F	W	L	G	Y	L	N	S	C	I	N	P	I	I	Y	P	C	S			

4.3 Structural Analysis of the α_{1A}-AR Homology Models

The differences in the initial structures were measured via the RMSD of the Cα atoms of the receptors. The two models are structurally quite similar with a RMSD of 0.81 Å (see Figure 4.3). The short helical segment TM-VIII was experimentally determined in the Rhodopsin crystal structure and is present in both models. Analysis of residues within each TM domain, allows for the identification of divergent motifs that may be responsible for local structural deviations from the Rhodopsin structure. Pro residues (20 present in the α_{1A}-AR) can have a significant effect in modulating the conformation of TM helices. In TM-V and TM-VI there is a Pro residue, which is conserved amongst most GPCRs and in both models this residue imposes a distortion. Ballesteros *et al.*[91] and Sansom *et al.*[200] examined the role of Pro-kink motifs in modulating the conformation of TM helices and suggested that these motifs can function as flexible hinges inducing a significant bend in the helix. These Pro-kinks may play a functional role by supporting alternative helical conformations, which could be used to signal from one side of a bilayer to the other.[201,202] Distortions of helices can also occur in the absence of Pro residues.[200] The pronounced kink in TM-II of Rhodopsin is caused by flexibility of the Gly-Gly sequence in the middle of this helix. However, this Gly-Gly motif is not present in the α_{1A}-AR sequence and no such kink was observed in our models.

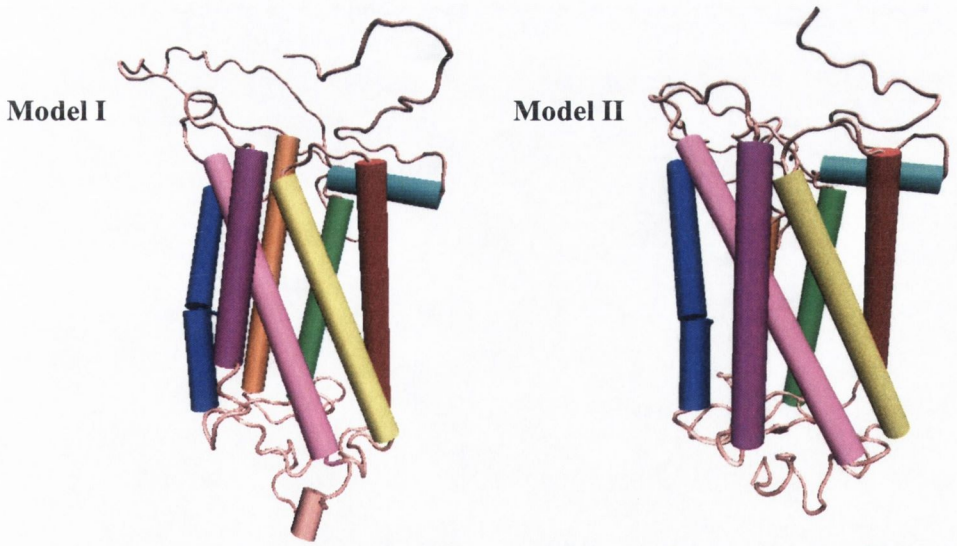


Figure 4.3: α_{1A}-AR homology Model I and Model II, illustrating TM-I (red), TM-II (yellow), TM-III (mauve), TM-IV (purple), TM-V (blue), TM-VI (orange), TM-VII (green), TM-VIII (cyan) and loop regions (pink).

There are also differences noted in the tail regions of the models, where a short helix is present for both Rhodopsin and Model I but not for Model II. Some structural differences were also noted in the models' loop regions particularly, in IC-II and EC-II. The EC-II loop (14 residues in length) is partially inserted into the TM framework as also noted in the recent α_{1A}-AR model of Pedretti *et al.*.[191]

Procheck[203] analysis was utilised to examine the models' stereochemistry. In the resultant Ramachandran plots, the background colour varies according to how appropriate the ϕ and ψ angles are for the residues. A protein is stereochemically satisfactory when a high percentage of residues (over 90 %) are in the core regions (darker shaded red regions). This is largely observed in the ramachandran plots of Figure 4.4.

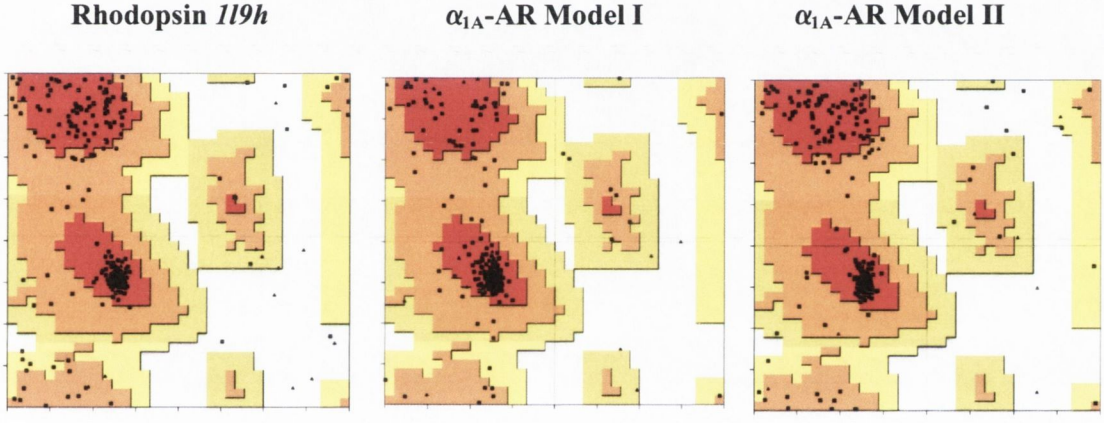


Figure 4.4: Ramachandran plots for the Rhodopsin structure, Model I and Model II.

A summary of the stereochemical information for rhodopsin and the two α_{1A}-AR models is given in Table 4.2. For Rhodopsin, 87.8 % (259 of the non Gly and non Pro residues) are in the most favoured regions, with the remaining 12.2 % (36 residues) in allowed regions. For Model I, a similar stereochemical quality was observed with 82.5 % (261 residues) in the most favoured regions and 1.9 % (six residues) in disfavoured regions with the remaining 15.6 % in generously allowed regions. Finally, for Model II, 85.4 % (275 residues) are in the most favoured regions with 1.2 % (four residues) in disfavoured regions. As previously mentioned, there is greater uncertainty in modelling loop regions, which are greater than 5-6 residues. When the helical structures of the models were considered without the loops or tail regions, residues were not observed in disallowed stereochemical regions. The one exception occurred for Model I with Arg-166 of TM-IV.

Table 4.2: Percentage of residues in the various stereochemical regions, with the corresponding percentage without the loops indicated in brackets.

	Rhodopsin	**α_{1A}-AR Model I**	**α_{1A}-AR Model II**
Most favoured	87.8 (98.5)	82.1 (97.4)	85.4 (94.9)
Allowed	10.5 (1.5)	13.5 (2.2)	11.2 (5.1)
Generously Allowed	1.7 (0.0)	2.5 (0.0)	2.2 (0.0)
Disallowed	0.0 (0.0)	1.9 (0.3)	1.2 (0.0)

Hydrogen bonding is a fundamental driving force in the folding of proteins. Analysis using the Definition of Secondary Structure of Proteins program (DSSP[204]) indicated a large HB network with 229 for Model I and 230 for Model II compared to 248 for rhodopsin. All three residues of the conserved DRY motif are involved in HBs in the homology models. For Model I, Asp-123 forms a HB with Ile-119 and Ile-120; Arg-124 with Asp-123, His-256, Ile-120 and Ser-121; and Tyr-125 with Ser-121 and Ile-122. Similarly, in Model II, Asp-123 forms a HB with Ile-119; Arg-124 with Glu-267, Ile-120 and Ser-121; and Tyr-125 with Ser-121 and Ile-122. Such a bonding network may hold the receptor in its inactive state.

Furthermore, the Rhodopsin structure contains a disulfide bond (Cys-110 to Cys-187) of 2.03 Å. By analogy, a disulfide bond may form between Cys-99 in TM-III and Cys-176 in EC-II of the α_{1A}-AR. In Model I, the S-S distance is 2.02 Å, while for Model II, due to the differences in alignment of the loop regions with the template a disulfide bond did not form, with a S-S distance of 10.15 Å. No other disulfide bonds were formed for either homology model.

The radius of gyration of the models serves as a measure of the compactness of the structures and reflects this differing number of interhelical bonds. For the bovine Rhodopsin template the radius of gyration is 21.96 Å, which is comparable to 22.69 Å for Model I, while Model II is slightly larger at 24.49 Å. The packing of the TM framework is further stabilised by a network of hydrophobic interactions, such as an aromatic cluster formed by Phe-193 and Tyr-194 in TM-V; Trp-285, Phe-288 and Phe-289 in TM-VI; and Phe-308, Phe-312, Trp-313 and Tyr-316 in TM-VII of the α_{1A}-AR models.

4.4 α_{1A}-AR Model Refinement - Molecular Dynamics Simulations

The use of the Amber force field[128] in the study of pure lipid simulations has been examined in only a few cases. This includes united atom studies on the phospholipid dipalmitoylphosphatidylcholine (DPPC) by Smondyrev *et al.*[205] whose parameters were further developed for the palmitoyoleoylphosphatidylcholine (POPC) lipid by Lin *et al.*[206] and the dimyristoylphosphatidylcholine (DMPC) lipid by Moore *et al.*.[207] Simpler, all-atom, membrane mimetics also possess the two essential phases to approximate the hydrophobic/hydrophilic interface found in lipid bilayers. A bewildering variety of solvent environments, with faster equilibration periods than phospholipids, have been used to maintain the integrity of protein structures, see section 1.2.3.[208] Indeed, Wymore *et al.*,[56] concluded that the same qualitative features were achieved in simulations of the biphasic (hexane/H_2O) system to that of phospholipids (DMPC).

In the present study, comparisons are made between simulations of one of the homology models (Model II) in a variety of solvent environments. Simulations were performed in the TIP3P water model (Simulation A), a TIP3P/$CHCl_3$/TIP3P bilayer (Simulation B) and finally a TIP3P/CCl_4/TIP3P bilayer (Simulation C), see Scheme 4.1. Simulation A was performed to examine the effect of a non-hydrophobic environment on the helical region of the receptor. Simulation B and C were performed to mimic a lipid bilayer structure that is hydrophilic on its two outer surfaces and hydrophobic in between. Both CCl_4 and $CHCl_3$ are organic solvents with low dielectric constants (2.238 and 4.806) and are immiscible in water.[209]

		II-A (H_2O)		F-A
Model II	solvent	II-B (H_2O/$CHCl_3$)	MD (1ns)	F-B
	⟶	II-C (H_2O/CCl_4)	⟶	F-C

Scheme 4.1: Computational scheme for examination of the role of solvent in MD simulations in α_{1A}-AR structure refinement.

The positioning of the receptor model into a membrane mimic is an important consideration. Aromatic residues such as Trp and Tyr on the surface of membrane proteins tend to be located in bands that correspond to the locations of the lipid headgroup-water interfaces of the bilayer in which the protein is embedded. The outer surfaces of the α_{1A}-AR model were examined and two rings of Trp and Tyr residues

were determined (consisting of residues 91 and 92 and residues 63, 165, 204 and 326), which marks the presumed interfacial site in the α_{1A}-AR. These arrangements were used to define the placement of the model in the bilayers and such a cell is shown in Figure 4.5. Counter-ions were added to neutralise each cell, replacing a H_2O molecule of high electrostatic potential.

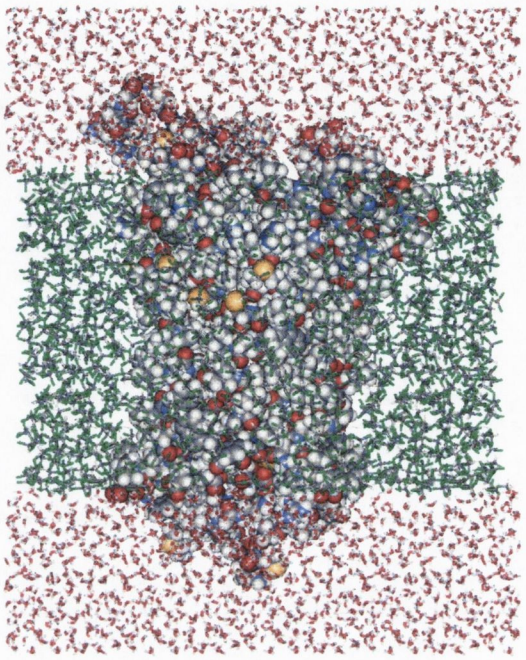


Figure 4.5: Initial simulation cell of Model II in a $H_2O/CHCl_3/H_2O$ solvent box (Simulation B).

4.4.1 Structural Analysis of the α_{1A}-AR over Simulations A-C

MD simulations were utilised to gain insights into the effect of the environment on the receptor and to monitor the movement of the initial 'rhodopsin-like' model towards a native α_{1A}-AR conformation. The analysis has been broken into three stages, the optimisation step, the scaling period, which involves heating and equilibrating and finally the production run. Little structural change occurs over the optimisation steps while larger structural changes occur over the scaling period of all three simulations generating a rise in RMSD. The time dependent RMSD data were monitored over the production runs to examine the stabilisation of these structural changes to the helices over the simulations. Average trendlines (period of 50 ps) of the RMSD values for each helix are plotted for clarity. No helical unravelling was noted over any of the simulations and the overall RMSD analysis indicates that structural stability is achieved.

For simulation A, the largest deviation was observed for TM-II, while the smallest deviations were observed for the short TM-VIII, which underwent little structural change throughout the simulation (see Figure 4.6). In general, the time dependent RMSD data of simulation A indicates a substantial structural change from a 'rhodopsin-like' receptor form occurred over the scaling steps and has largely stabilised over the production run. An exception occurs for TM-II over the last 100 ps.

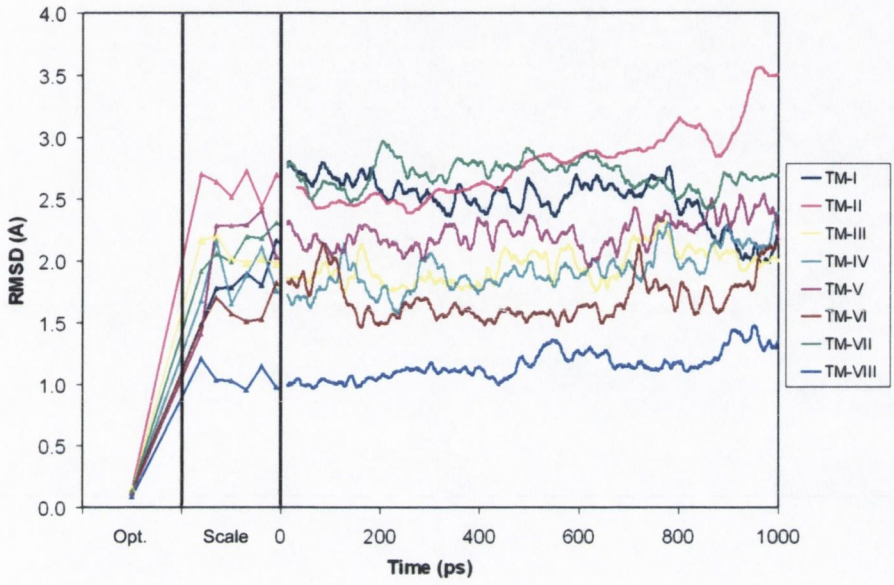


Figure 4.6: RMSDs in Å for simulation A. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

For simulation B (see Figure 4.7), again little structural change occurred over the optimisation step followed by significant structural changes over the equilibration period, indicating a movement from the initial 'rhodopsin-like' homology model. The RMSDs have largely stabilised by 300 ps.

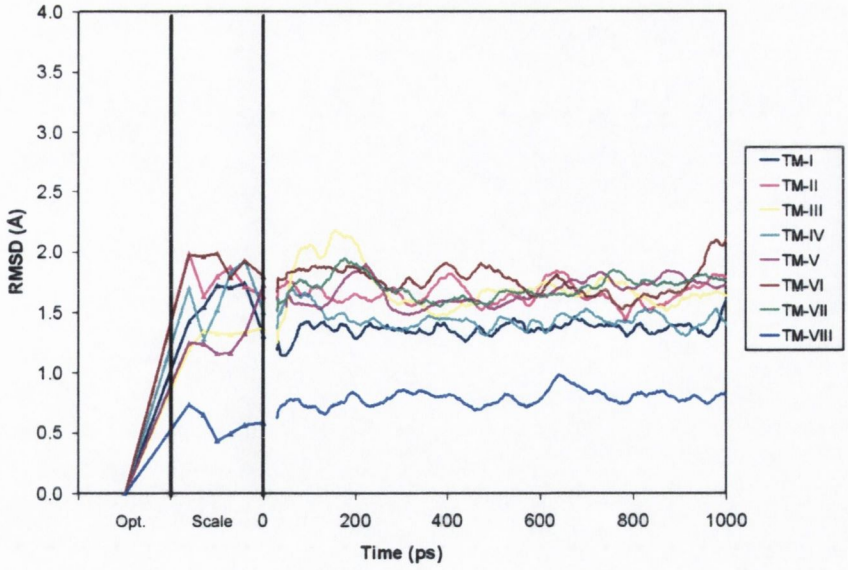


Figure 4.7: RMSDs in Å for simulation B. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

For simulation C, a rise in RMSD is again observed over the course of the equilibration period, before stabilisation at a RMSD level comparable to that of simulation A (see Figure 4.8). Over the production run the largest deviation in RMSD occurs with TM-V. Small structural fluctuations occur at ~200 ps for TM-I and TM-VI. The receptor structure appears to have stabilised by 400 ps, after which the RMSD analysis deviates around the mean values for each helix. An exception is TM-IV, which experiences a small rise in RMSD at 900 ps.

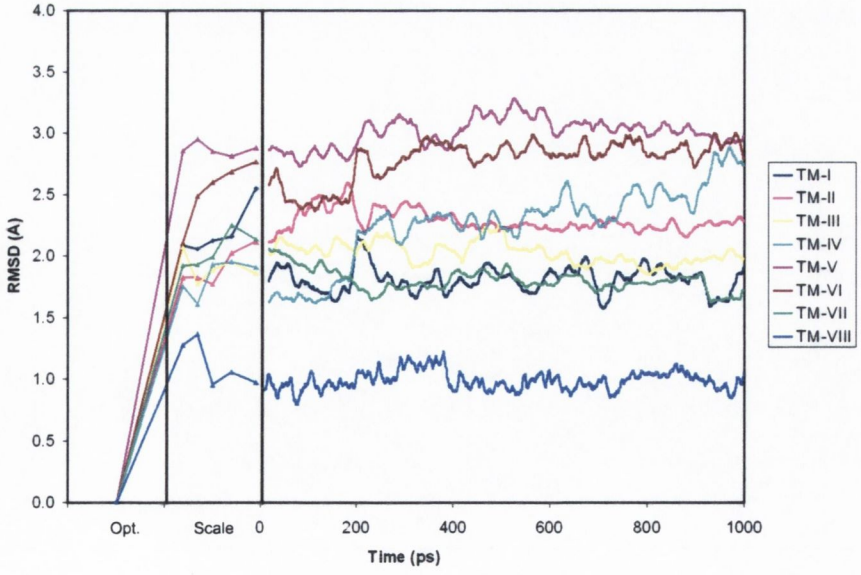


Figure 4.8: RMSDs in Å for simulation C. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

In summary, the largest RMSDs over the production runs were observed for Simulation A. This is not unexpected as the helices are in water, which is not an appropriate membrane mimic to fully maintain helical stability. Structural changes also occurred for simulations B and C from the initial 'rhodopsin-like' state. Simulations B and C are more appropriate mimics for the hydrophilic/hydrophobic nature of a phospholipid membrane and hence may induce more appropriate side chain conformations in the helical regions. For simulation A, there was 20 inter-helical HB sampled with an average of eight present. However, for simulations B and C, with a more suitable membrane mimetic environment a larger number of interhelical HB were observed as the polar residues were found to be orientated towards the inside of the helical bundle and to have more HB partners. For simulation B, there were 30 HBs sampled with an average of 15 present and for simulation C, there were 44 sampled with an average of 23 present. The resultant receptors are believed to represent inactive conformations of the α_{1A}-AR.

4.4.2 Refined Structural Analysis of the α_{1A}-AR Models

The average structures over the last 200 ps of each simulation were optimised and termed F-A, F-B and F-C. The stereochemical quality of the structures were evaluated through Ramachandran plots, F-A having 250 residues (77.6 %) in most favoured regions and four residues (1.2 %) in disallowed regions (Figure 4.9). The remaining 115 residues are in additionally and generously allowed regions. Receptor F-B has 252 residues (78.3 %) in most favoured regions and five residues (1.6 %) in disallowed regions while receptor F-C has 238 residues (73.9 %) in most favoured regions and only two residues (0.6 %) in disallowed regions. However, for receptors F-B and F-C, all the residues in disallowed regions were in loop and tail regions. Given the relatively low percentage of residues having disallowed torsional angles, the stereochemical quality is acceptable for all three final models.

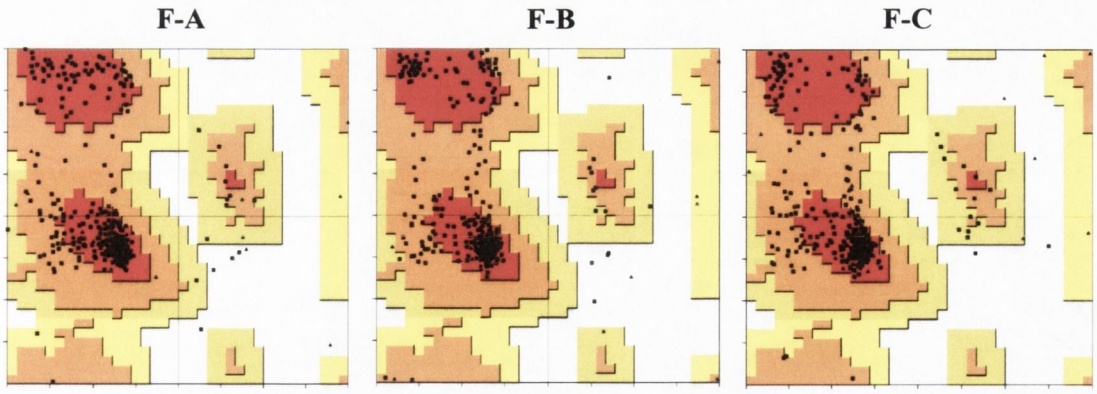


Figure 4.9: Stereochemical analysis of the average structures after molecular dynamics simulations for receptor structures F-A (on left), F-B (in middle) and F-C (on right).

The MD simulations were utilised to gain insights into the effect of the environment on the receptor and at the end of simulation B and C (1 ns) a stable bilayer remained. The use of biphasic models is sufficient to obtain qualitative structural information on the relative stability of various models. The stability of the receptor model in a bilayer environment was examined under the criteria of RMSDs, secondary structure stability, quantity of HBs and stereochemical acceptability. As no significant difference was observed between the resultant receptor structures in the biphasic systems over the 1 ns time period of the study and $CHCl_3$ was the closer of the two to a lipid having hydrogens present, the $H_2O/CHCl_3/H_2O$ cell was chosen for further studies.

4.5 Molecular Dynamics Simulations on Rhodopsin and the α_{1A}-AR Models

MD simulations (1 ns) were performed on Model I (termed Simulation I) and on Model II (termed Simulation II), in a $H_2O/CHCl_3/H_2O$ cell (see Scheme 4.2). As a validation of the MD protocol employed in this study a control simulation was also performed on the *1l9h* Rhodopsin crystal structure. Our MD refined receptor models were termed Final I and Final II.

Model I	Simulation (1ns)	Final I
Model II	⟶	Final II
Rhodopsin	($CHCl_3/H_2O/CHCl_3$)	Final Rhodopsin

Scheme 4.2: Computational scheme for refinement of the α_{1A}-AR homology models and the rhodopsin structure through molecular dynamics simulations.

4.5.1 Structural Analysis for Rhodopsin and the α_{1A}-AR Models over the Molecular Dynamics Simulations

A direct comparison was made of the time dependent RMSD data for the three simulations. Over the course of the Rhodopsin simulation (1 ns), there were helical RMSDs (see Figure 4.10), consistent with the small structural rearrangement expected from the static crystal structure. An exception was for TM-VII, which experiences a rise in RMSD at the optimisation step, which continues over equilibration but starts to decrease over the production run. Structural rearrangements also occurred for TM-VIII, possibly as a consequence of the movement of TM-VII, which have stabilised by 800 ps.

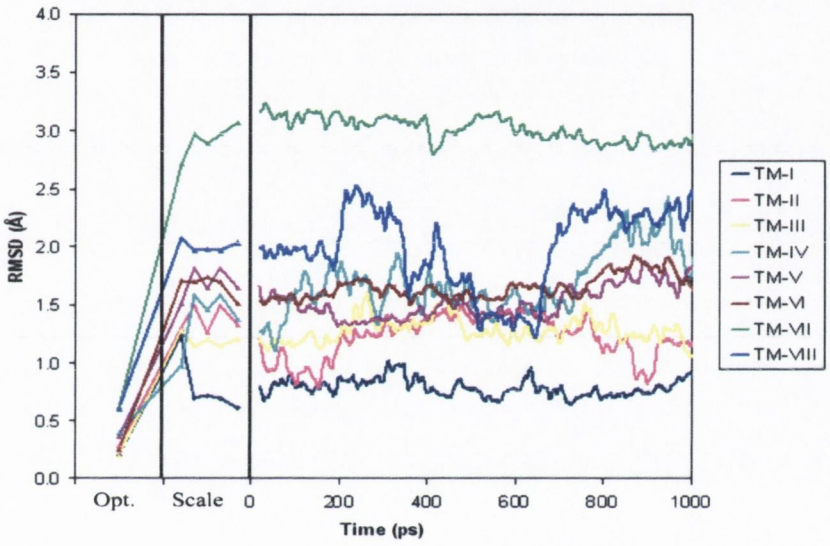


Figure 4.10: RMSDs in Å for the rhodopsin simulation. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

The time dependent RMSD data was analysed for Simulation I (see Figure 4.11), where larger RMSDs were observed than for the rhodopsin simulation. Again this is not unexpected as a structural change from the 'rhodopsin-like' conformation is required to achieve a native α_{1A}-AR structure. The most notable structural changes occur for TM-VI, which rose significantly in RMSD over the scaling step. There was also a rise in RMSD for TM-II at 600 ps, while the helices have largely structurally stabilised by 800 ps.

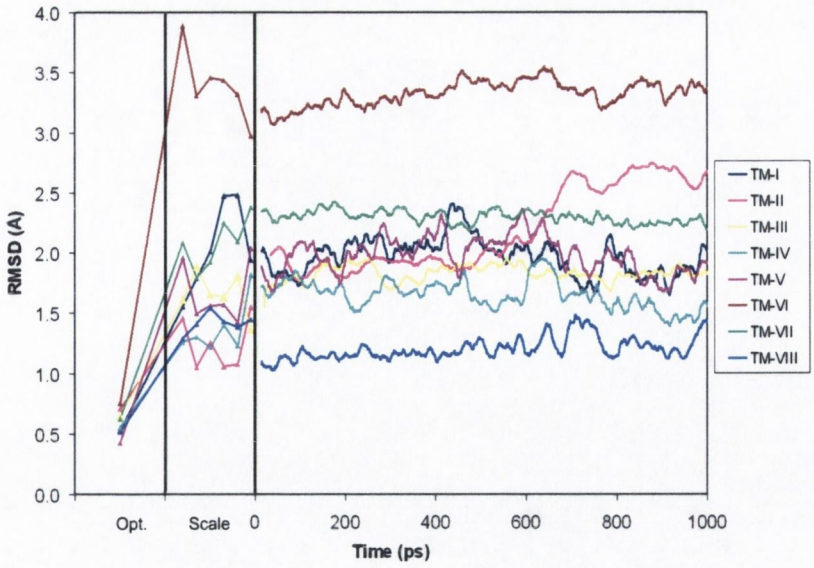


Figure 4.11: RMSDs in Å for the simulation of Model I. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

As previously noted in the discussion of the environment effect on Model II, over simulation II, the largest structural deviation occurred for TM-VI (see Figure 4.12). Monitoring the time dependent RMSD for the helices a structural change for the receptor occurs at ~300 ps, which stabilises by 400 ps.

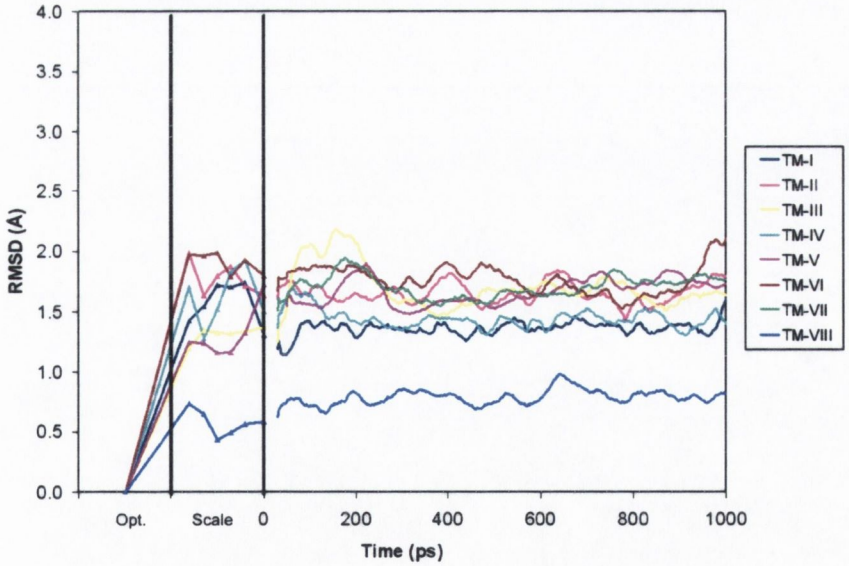


Figure 4.12: RMSDs in Å for the simulation of Model II. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns).

Coupled with the RMSD analysis, geometrical features were monitored over the simulations. One structural template that is potentially shared by most, if not all, GPCRs is an inter-helical network of side chain interactions. During the simulation a larger number of inter-helical HBs occurred with 33 different interactions observed for Simulation I, compared to 30 for Simulation II and 13 for the Rhodopsin simulation. Stabilising features included a salt bridge between Arg-96 in TM-III and Asp-172 (EC-II) that formed for Simulation II, but not Simulation I. For Rhodopsin the Cys-110 / Cys-187 disulfide interaction is at an average distance of 4.13 Å. The two Cys residues postulated to form a disulfide interaction in α_{1A}-AR have drifted apart over the production run of Simulation I at an average distance of 7.56 Å. For Simulation II the corresponding maximum distance is 16.47 Å, similar to its initial distance (10.15 Å).

For Rhodopsin the HB connecting Glu-134 with Arg-135 of the ERY motif has a maximum distance of 3.09 Å. This interaction is thought to be important for maintaining the receptor in its inactive state. The equivalent interaction is also formed for Asp-123 and Arg-124 of the corresponding α_{1A}-AR DRY motif, which was longer for Simulation I at a maximum distance of 5.01 Å and was maintained with a maximum distance of 3.79 Å for Simulation II.

4.5.2 Structural Analysis after Molecular Dynamics Simulations

After the MD simulations (1 ns), the final structures of Rhodopsin and the two α_{1A}-AR models were averaged over the last 200 ps of the production run and optimised. For the final Rhodopsin structure the overall Cα RMSD from the starting structure was 2.65 Å. The helical RMSDs were smaller indicating larger structural changes in the loop regions of the receptor . Comparing the initial and final states, the movement from Model I to Final I had a Cα RMSD of 4.97 Å, while the largest helical movement occurred for TM-I at 3.41 Å (Table 4.3). The Cα RMSD of Final II with respect to Model II was 5.38 Å although this structural deviation was primarily in the loop regions with the largest helical deviation determined for TM-IV at 2.68 Å. TM-VI also exhibited large deviations for both final structures. The Cα RMSD of the two final α_{1A}-AR structures with respect to each other shows a clear deviation of 8.17 Å which is comparable to that found for the initial structures, 8.44 Å. A particularly large RMSD occurs again in the case of TM-VI at 4.89 Å.

Table 4.3: RMSD (in Å) between Initial and Final structures for Simulation I and II.

Helix	Model I vs Final I	Model II vs Final II	Final I vs Final II
TM-I	3.41	1.77	2.39
TM-II	2.46	2.03	4.15
TM-III	2.30	2.33	2.19
TM-IV	2.38	2.68	2.80
TM-V	2.25	2.25	2.21
TM-VI	3.22	2.53	4.89
TM-VII	3.29	1.97	3.57
TM-VIII	0.96	0.72	1.17
C α (all)	4.97	5.38	8.18

The final Rhodopsin structure has a total of 251 HBs including 17 inter-helical interactions. The number of HBs has increased for the α_{1A}-AR models with a total of 252 HBs in Final I (from 229) and 245 for Final II (from 230). This includes a large increase in stabilising inter-helical interactions from seven to 24 for Final I and from four to 15 for Final II. For Rhodopsin the HB connecting the ERY motif is maintained at 2.71 Å as well as the disulfide bond at 2.04 Å. For the α_{1A}-AR models, the HB network of the DRY motif is also maintained while no disulfide interactions or ion-pairs were observed in the final models. The radius of gyration for Final I is 24.32 Å and for Final II is 23.94 Å, indicating that the compactness of the models while larger are still comparable to the Rhodopsin final structure at 21.83 Å.

In summary, both α_{1A}-AR models have remained structurally stable over the course of the simulations and have increased favourable interactions most notably in the form of HBs. The final refined structures represent two models of the α_{1A}-AR in the inactive state. To our knowledge this work represents the first development of an adrenoceptor in the inactive state, refined through molecular dynamics simulations in an appropriate membrane mimetic.[190]

4.6 Discussion / Conclusions

In this study, homology modelling was used to obtain structural information for the human α_{1A}-AR. The structure prediction process consisted of target-template alignment, model building and model evaluation. Although prediction methods are far from replacing high-resolution structures, homology modelling has become increasingly important in prediction of the fold of proteins with homologues of known structures. The Achilles heel of homology modelling is its ultimate dependence on the level and quality of existing structural knowledge and the similarity to the target protein. Since there is only one experimentally known Class A GPCR structure, homology applications lead to TM regions that are very similar to the bovine Rhodopsin template structure. Such models are considered as low resolution structures due to the low sequence identity and the consequent difficulty in aligning templates.

As part of this research, structural refinements from the initial 'rhodopsin-like' conformations of the homology models to more native α_{1A}-AR conformations were pursued. One of the most important properties of the lipid bilayer is a well-defined hydrophilic/hydrophobic interface whose essential properties are preserved in a mimetic system.[210] We have selected a simple biphasic system, $H_2O/CHCl_3$, to facilitate more routine simulations of GPCRs in 'membrane-like' surroundings. In a sense, the use of a $CHCl_3/H_2O$ interface as a model system for a membrane falls between simulations of a lipid bilayer in full atomic detail and continuum electrostatics calculations on membrane receptors. The main advantages of 'membrane mimetic' simulations are

(i) The dynamics of $CHCl_3$ is orders of magnitude faster than that of phospholipids.

(ii) They include a reasonable representation of water and the hydrophobic membrane interior and are expected to be an accurate representation of the membrane.

Limitations to the use of $CHCl_3$ in membrane mimics include the lack of a number of features of real membranes, including phospholipid headgroups, the structured long lipid tails and the non-uniform density and pressure profiles of real lipid bilayers. The principal limitation of the implemented MD simulations are their

relatively short duration, a result of computational limitations for atomistic simulations of relatively large systems (~70,000 atoms).

The MD simulations presented provide the first detailed study of an α_{1A}-AR model inserted into a membrane mimic.[190] The interaction networks of the α_{1A}-AR models were examined for the functionally crucial motifs in GPCRs, such as the E/DRY and NPxxY motifs. The refined α_{1A}-AR models retained their α helical conformations and the systems remained stable during the MD simulations (1 ns). The final structures were satisfactory both sterically and energetically and represented inactive conformations of the receptor.

Chapter V: Examination of α_1 Adrenoceptor Ligands

"We are reaching the stage where the problems we must solve are going to become insoluble without computers. I do not fear computers. I fear the lack of them."

Issac Asimov

5.0 Introduction

In this chapter, we shall computationally examine the two catecholamine agonists, adrenaline and noradrenaline (see Figure 1.2) and a series of α_1-AR antagonists as potential agents in the treatment of BPH. In Chapter I, we reviewed a number of antagonists acting on the α-ARs that are currently used in the treatment of BPH, including prazosin, terazosin, doxazosin, alfuzosin and tamsulosin (see Figure 5.1). The quinazoline antagonists were originally developed as antihypertensives and their ameliorative effects in the treatment of BPH were not observed until their introduction into clinical practice.[211]

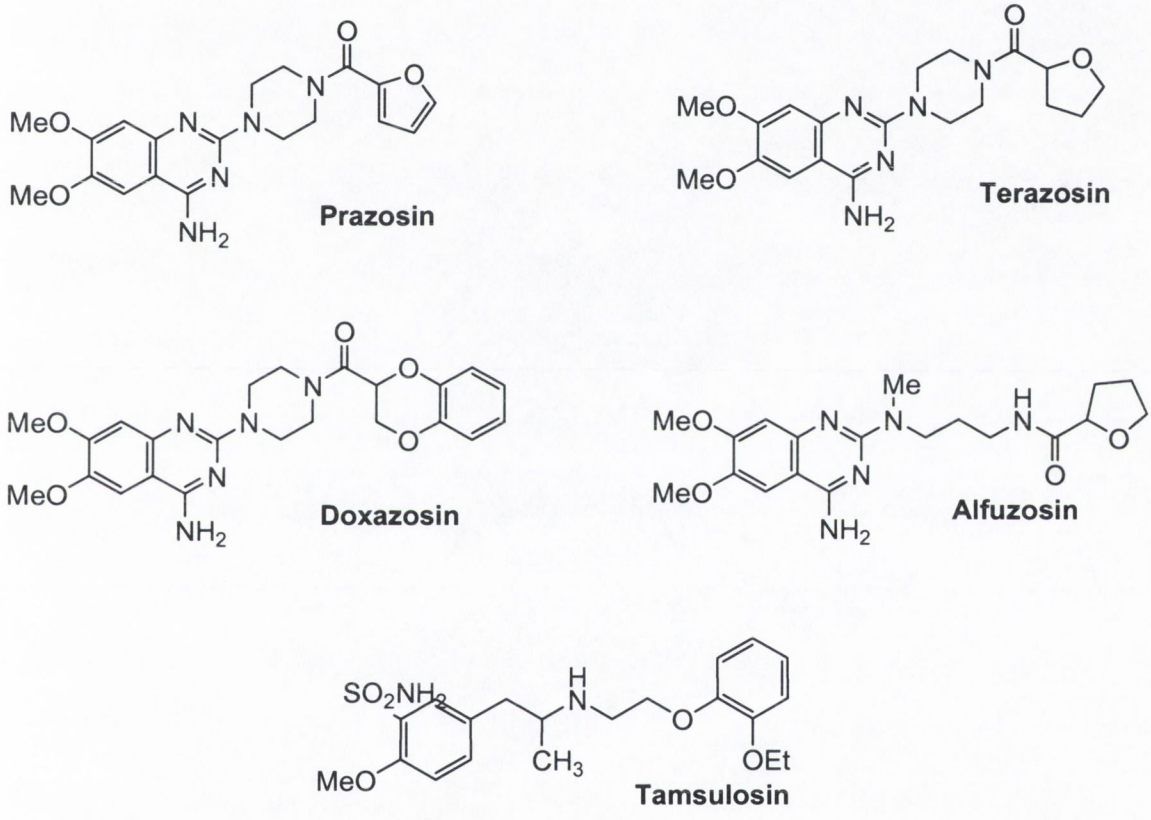
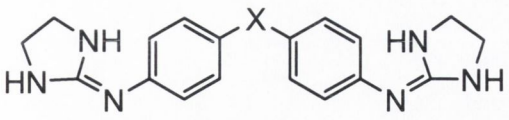


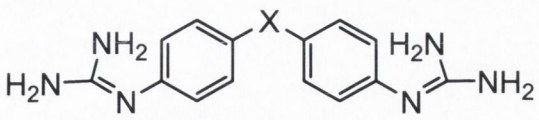
Figure 5.1: Antagonists used in clinical practice for the treatment of BPH (prazosin, terazosin, doxazosin, alfuzosin and tamsulosin).

Ongoing research suggests that use of selective, α_{1A}-AR antagonists offers an advantage over non-selective α_1-AR blockers by allowing for the administration of therapeutic doses with fewer systematic side effects. Previous work investigated two families of compounds with bis-imino-imidazolidinium or bis-guanidinium cations at both ends of a chain formed by diphenyl derivatives that exhibited α_{1A}-AR antagonist activity, compounds **1-8**, in Figure 5.2.[60] Pharmacological studies on slices of human

prostate with BPH showed that guanidinium derivatives are able to inhibit between 90-95 % of the contractions induced by noradrenaline. These results are comparable to the inhibition observed for doxazosin (95 %) and these agents may be useful in the treatment of BPH.[61]



1 (X=NH), **2** (X=CO),
3 ($X=SO_2$), **4** ($X=CH_2$)



5 (X=NH), **6** (X=CO),
7 ($X=SO_2$), **8** ($X=CH_2$)

Figure 5.2: α_{1A}–AR antagonists prepared previously by our group (compounds **1** to **8**).

A literature review of α_1-AR antagonists yielded, amongst others, the compounds described in this section (Figure 5.3). The piperazine pyriminedione derivatives 5-methylurapidyl (**9**)[212] and RS-100,975 (**10**)[213] are antagonists of the α_{1A}–AR and α_{1D}-AR. Screening the R. W. Johnson PRI library of compounds, Kuo *et al.* found that the piperazine oxazoline derivative RWJ-37914 (**11**)[62] has very good selectivity towards the α_{1A}-AR. A number of benzodioxanes were also prepared and tested by Barbaro *et al.*[214] and in particular, compound **12** was determined to fit the α_{1A}-AR pharmacophore model previously developed.[214] Chern *et al.*,[64] prepared tricyclic fused quinazolines of which compound **13** exhibited the largest selectivity for α_{1A}-AR binding over α_{1B}-AR.

A variety of antagonists contain sole nitrogens in an alkyl chain, including the benzodioxan derivatives related to WB4101 (**14**, a known α_1–AR antagonist). Quaglia *et al.*,[65] developed phendioxan (**15**) with a marked drop in affinity towards the α_{1D}-AR and α_{1B}-AR subtypes compared to **14**, while not affecting the affinity for the α_{1A}-AR subtype. Further work in this group resulted in compound **16** as a potent and selective antagonist for the α_{1A}-AR. Finally, Phenoxybenzamine (**17**) is an α-AR antagonist while RS-17053 (**18**), is an α_{1A}–AR subtype selective antagonist, which displays 126- and 50-fold selectivity over human α_{1B}- and α_{1D}- ARs respectively.[215,216]

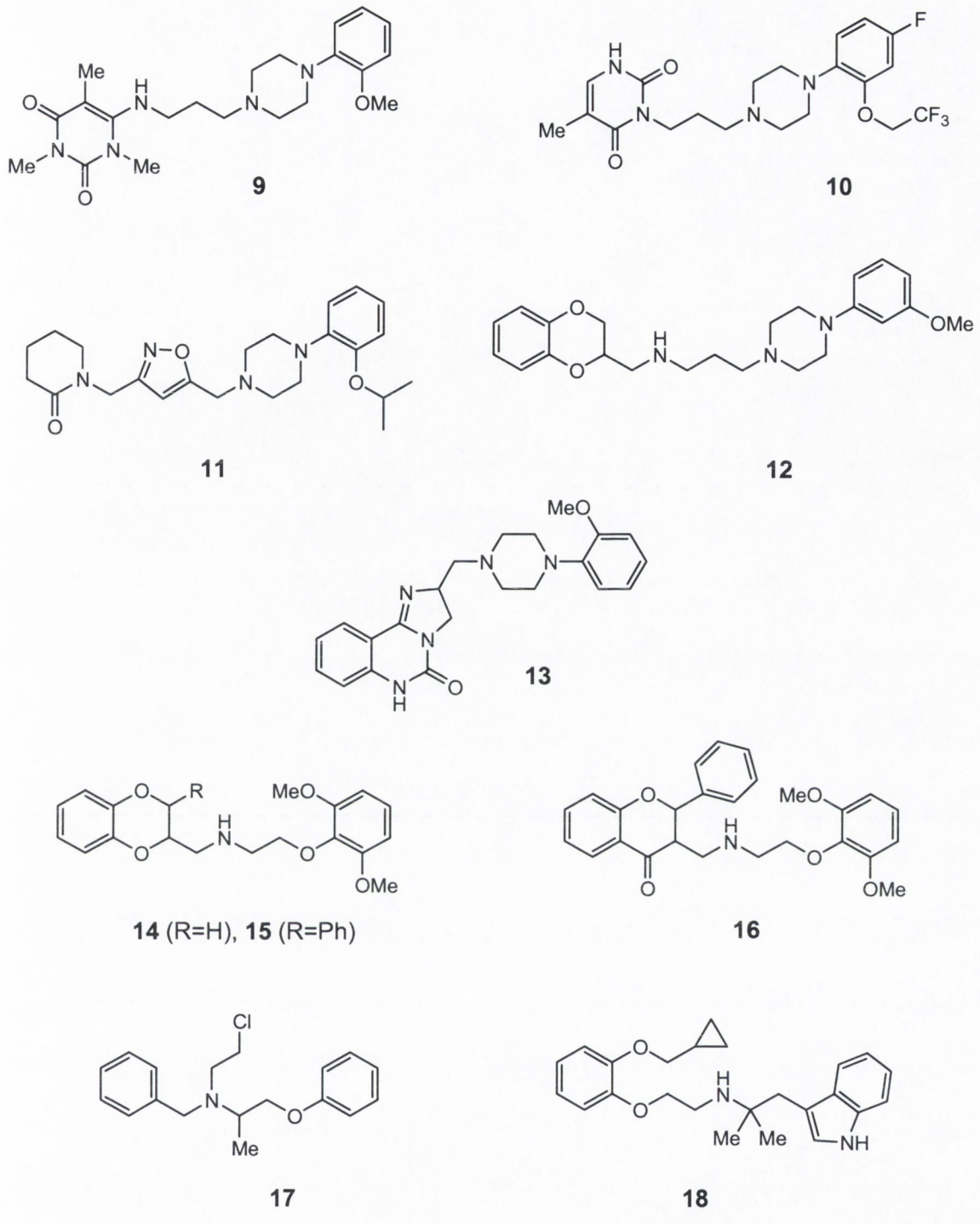


Figure 5.3: Antagonists of the α_{1A}-AR developed by different authors (**9** to **18**).

In order to further investigate the α_{1A}-AR affinity of these compounds and to design new potential drugs, we ultimately aim to model the interaction of these ligands with the α_{1A}-AR models. In this chapter, we will study the conformational freedom and calculate the proton affinities (PA), with and without solvation effects, of this selection of α_{1A}-AR agonists and antagonists to determine possible interacting states with the α_{1A}-AR under physiological conditions.

5.1 Theoretical estimation of Proton affinity

Proton affinity (PA) is defined as the negative of the enthalpy change on reaction of a proton and the chemical species concerned to give the conjugate acid of that species. Techniques such as electrospray ionization, matrix-assisted laser desorption or the extended kinetic method allow for the experimental determination of PA values and several examples of the application of these methods can be found in the recent literature.[217,218] Computational approaches are also valuable tools to determine PAs in the gas phase and a number of recent systematic electronic structure studies on PAs have been performed at high theoretical levels.[219,220,221] Furthermore, DFT studies using the B3LYP functionals with 6-31G* and 6-31+G** basis sets have been shown to produce PA values within 3-5 kcal mol^{-1} of experimental results.[222,223] As the protonated state of a ligand will play an essential role in the ligand/receptor interaction, the predicted PA values for the protonation were examined.

5.1.1 Proton Affinity Determination for the Catecholamine Agonists

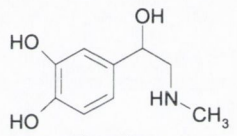
A systematic, conformational analysis was pursued on the unprotonated agonists followed by semi-empirical optimisations (AM1) to rank the conformers in order of energy. Optimised molecular structures of the unprotonated and protonated forms were then computed using the B3LYP hybrid density functional and the 6-31G* basis set. Electronic structure and vibrational frequency calculations were performed using Gaussian98[178] or Gaussian03.[224]

The PA determined for adrenaline was 243 kcal mol^{-1} and for noradrenaline was 240 kcal mol^{-1} (see Table 5.1). The QM energy minimum corresponds to a hypothetical, motionless state at 0K, while experimental measurements are made on molecules at a finite temperature. To compare the two, it is necessary to make appropriate corrections to allow for translational, rotational and vibrational motions. Hence, the gas phase values were thermally corrected to 234 and 231 kcal mol^{-1} for adrenaline and noradrenaline respectively.

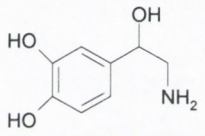
Biology effects are dominated by processes, which occur in aqueous or lipid environments. Hence, accurate simulation of solvation is central to the development of reasonable models for biological simulations.[225] PA values for adrenaline and noradrenaline were additionally determined utilising the Onsager and the PCM

solvation techniques.[148] Consideration of the solvation energy led to a stabilisation from the *in vacuo* energy values (see Table 5.1). Unlike a gas, a liquid is a set of strongly interacting particles and the intra- and inter- molecular interactions can cause substantial changes in the geometry and electronic structure of substrates in comparison with isolated gas phase. Including Onsager solvation effects, there was an increase in the PA values to 259 kcal mol^{-1} for adrenaline and 246 kcal mol^{-1} for noradrenaline at 0 K and 249 and 237 kcal mol^{-1} at 298 K (Table 5.1). Larger PA values were obtained using the PCM solvation technique, with a further increase to 295 kcal mol^{-1} for adrenaline and 289 kcal mol^{-1} for noradrenaline at 0 K and 286 and 281 kcal mol^{-1} at 298 K. Unfortunately, no experimental PA values were available for comparison.

Table 5.1: Proton affinities (PA) in kcal mol^{-1} were estimated as the difference in total energy between the protonated and the corresponding neutral species. No scaling was applied to the obtained frequencies for the calculation of the thermodynamic parameters, which were temperature corrected to 298.15K.



adrenaline



noradrenaline

	In Vacuo		*Onsager*		*PCM*	
Ligand	PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
Adrenaline	243.4	233.9	258.5	248.9	294.5	286.1
Noradrenaline	240.2	231.0	246.2	237.1	289.3	280.6

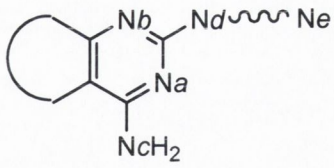
It is worth noting the computational cost of including the various solvation techniques. For the optimisation of protonated adrenaline there was little computational cost in moving from an *in vacuo* (71.2 mins) calculation to an Onsager solvated calculation (72.3 mins), while a marked increase in computational cost occurred for utilising the PCM technique (611.1 mins). Despite its simplicity, the Onsager solvation technique has had considerable success[226,227] and as the method requires little extra computational cost compared to *in vacuo* calculations this solvation technique was included in all future studies. The PCM solvation technique was included in a number of cases for comparison although its computational cost and technical difficulties in its implementation excluded it from full consideration for the larger antagonists.

5.1.2 Proton Affinity Determination for the Clinical Antagonists

Compounds within the 2,4-diaminoquinazoline series, typified by prazosin, terazosin, doxazosin and alfuzosin, differ from one another through changes to the quinazoline side chain. Detailed conformational analyses (random search) were performed on these antagonists followed by semi-empirical (AM1) optimisations on all the conformers to energetically rank them. A subsequent structural optimisation (B3LYP/6-31G*) was performed on the lowest energy conformer and its protonated Nitrogen forms, labelled N*a* – N*e* in the scheme of Table 5.2.

For the α_1–AR antagonist, prazosin the *in vacuo* calculations (Table 5.2) indicate that the nitrogen atom labelled as N*b* was the most likely to be protonated at 254.7 kcal mol^{-1}. As the amine group of the compound is electron donating, N*c,* is *ortho/para* directing towards N*b* and this particular protonated form was previously found to be the most stable protonated species of prazosin.[29] A study by de Benedetti *et al.*[228] highlighted that N*a* of prazosin, which is in the *ortho* position with respect to the amine group, is unlikely to be protonated, which is consistent with our results as the energy difference for protonating at N*a* with respect to N*b* is approximately, 10 kcal mol^{-1} (Table 5.2). For terazosin and doxazosin, the results followed the same trend, with the N*b* nitrogen having the highest PA both at 0 K and 298 K. The second nitrogen to be protonated is the second pyrimidine nitrogen, N*a*, showing an energy difference of ~10 kcal mol^{-1}. In the case of alfuzosin, with a related structure to the previous antagonists, the N*b* atom appears most likely to be protonated in gas phase, followed by N*e* and N*a*.

Table 5.2: Proton Affinity (PA) in kcal mol^{-1}, *in vacuo* and with Onsager and PCM solvation methods, without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the antagonists prazosin, terazosin, doxazosin and alfuzosin.



		In Vacuo		*Onsager*		*PCM*	
Ligand		PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
Prazosin	N*a*	244.4	220.5	247.4	239.1	285.8	279.1
	N*b*	254.7	231.0	258.4	250.0	292.4	284.8
	N*c*	222.8	198.2	235.2	226.6	270.3	262.2
	N*d*	240.0	215.2	241.0	231.9	277.7	269.9
	N*e*	233.7	209.7	251.7	243.5	272.2	264.9
Terazosin	N*a*	244.1	235.8	248.7	240.4	286.8	278.4
	N*b*	254.9	246.4	259.7	251.1	293.5	284.8
	N*c*	222.8	214.5	235.1	226.3	270.6	261.9
	N*d*	239.4	231.1	241.6	232.5	278.1	270.4
	N*e*	231.4	223.4	244.9	237.0	267.3	260.1
Doxazosin	N*a*	242.1	233.9	255.5	247.3	292.3	278.4
	N*b*	253.0	244.5	264.6	256.3	285.8	280.7
	N*c*	221.1	212.3	242.2	233.6	270.3	262.3
	N*d*	242.4	233.5	241.6	232.5	277.3	269.1
	N*e*	214.7	206.6	218.4	210.4	259.5	252.2
Alfuzosin	N*a*	255.8	247.3	255.9	247.4	#	#
	N*b*	261.3	252.8	263.7	255.1	297.2	289.3
	N*c*	235.8	227.3	236.4	227.9	#	#
	N*d*	242.0	232.9	243.9	234.8	#	#
	N*e*	257.0	248.3	271.7	263.0	292.9	284.0

a number of the optimisations failed to converge.

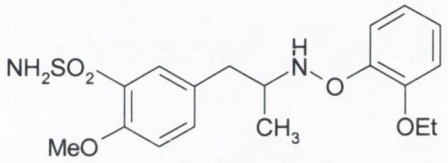
With both solvation techniques similar trends in stabilisation were observed, although a larger stabilisation of the protonated forms was again observed with the PCM solvation technique. Higher PA values were determined over *in vacuo* calculations, ranging over 47 kcal mol^{-1} (218-265 kcal mol^{-1}) for the Onsager solvation method and over 36 kcal mol^{-1} (259 to 297 kcal mol^{-1}) for the PCM technique. When thermally corrected the PA values range over 46 kcal mol^{-1} (210-256 kcal mol^{-1}) with Onsager solvation and over 28 kcal mol^{-1} (261 to 289 kcal mol^{-1}) for the PCM solvation.

For prazosin, terazosin and doxazosin the same order of PA is detected when solvation effects are included, N*b* has the highest affinity for a proton followed by N*a*. Curiously, the inclusion of Onsager solvent effects when calculating the PA of the

different N atoms of alfuzosin produces a change in the protonation order for this molecule. Thus, the most likely nitrogen to be protonated in an aqueous environment is N*e* followed by N*b* with a difference in PA of ~8 kcal mol^{-1} for the Onsager technique and N*b* followed by N*e* with a difference of 4 kcal mol^{-1} with the PCM technique.

For the antagonist tamsulosin, the lowest energy structure exhibits an intramolecular HB, which forces the molecule towards a folded conformation. In the gaseous phase the phenethyl amine moiety, N*a* was most likely to protonate. However, the sulphonamide, N*b,* is most likely to protonate when solvated, although the energy difference is quite small at ~3 kcal mol^{-1}, Table 5.3.

Table 5.3: Proton Affinity (PA) in kcal mol^{-1}, in gas phase and with the Onsager solvation methods without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the α_{1A}-AR antagonist tamsulosin.



		In Vacuo		*Onsager*	
Ligand		PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
Tamsulosin	N*a*	249.7	240.3	252.9	243.5
	N*b*	228.6	220.1	256.1	247.8

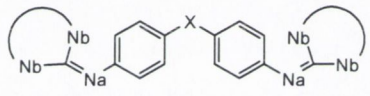
5.1.3 Proton Affinity Prediction for the α_{1A}–AR antagonists (1-8)

For the symmetric bisphenyl derivative compounds, **1-8**, there are two unique N atoms which can be protonated in the guanidinium moieties. They have been labelled N*a* and N*b* as shown in the scheme of Table 5.4. In addition, when the bridge group (X) is the NH group an additional protonation site is available, N*c*. The gas phase protonation energies, of the N*a* atoms range from 246 to 263 kcal mol^{-1}, in a 17 kcal mol^{-1} interval (Table 5.4). These differences are due to the various minimum energy conformations considered. In the gas phase, compound **1** shows, overall, the greatest protonation energy for the N*a* position, and, generally, the PA values obtained for those atoms are in a similar range to those obtained in the previous section. The thermally corrected PA values for the N*a* atom falls in the range between 238 and 254 kcal mol^{-1}. The bridging N*c* position of compound **1** and **6** was next likely to protonate. Regarding the outer N*b* position, lower PAs were observed than for the inner N*a* position. The PAs

obtained for N*b* are in intervals of 19 and 18 kcal mol^{-1} without and with thermal correction.

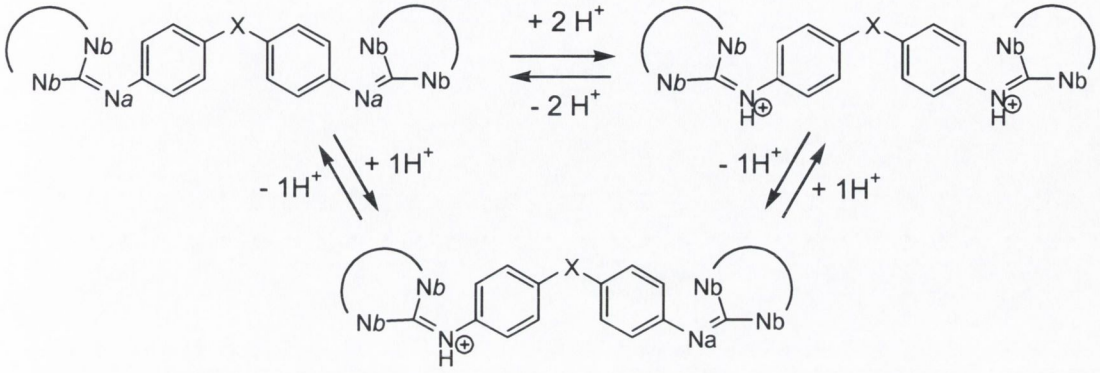
Generally, in these compounds, when solvation is considered, a stabilization of the PA for the protonations of N*a* is observed and the range of protonation energies slightly increases (Table 5.4). Thus, considering Onsager solvation without thermal correction there is a range of 20 kcal mol^{-1} (274 - 294 kcal mol^{-1}) and the values thermally corrected span a similar range (266 - 285 kcal mol^{-1}). A smaller range was observed with the PCM solvation technique of 7 kcal mol^{-1} (294 - 301 kcal mol^{-1}), which was thermally corrected to 11 kcal mol^{-1} (283 - 294 kcal mol^{-1}). However, solvation does not seem to affect the protonation of the bridge N*c* atoms possibly as these atoms are less exposed to the effects of solvent than those in the guanidinium or 2-iminoimidazolidinium groups. In summary, the N*a* position will be the first one to be protonated, for compounds **1** and **5** the bridge N*c* will be the second position to be protonated and N*b* will be the last one.

Table 5.4: Proton Affinity (PA) in kcal mol^{-1} in gas phase without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the α_{1A}–AR antagonists **1-8**.



		In Vacuo		*Onsager*		*PCM*	
Ligand		PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
1	N*a*	262.5	253.8	292.5	284.5	298.5	290.3
	N*b*	230.6	221.7	274.3	266.4	272.1	263.2
	N*c*	246.2	236.9	244.0	234.8	275.4	267.4
2	N*a*	254.1	245.5	284.1	276.4	293.7	283.0
	N*b*	223.5	214.6	264.8	257.0	232.1	216.9
3	N*a*	251.0	242.5	274.0	266.0	296.0	288.1
	N*b*	220.0	211.5	252.1	244.5	273.9	265.7
4	N*a*	258.8	250.1	293.5	285.4	293.8	285.0
	N*b*	225.5	216.8	265.2	257.3	266.2	257.4
5	N*a*	262.5	253.8	289.7	281.3	299.7	292.8
	N*b*	225.3	217.3	243.5	234.6	278.1	271.5
	N*c*	242.2	233.6	241.6	233.0	280.6	272.4
6	N*a*	250.0	242.1	287.6	278.9	294.6	287.8
	N*b*	216.0	207.7	275.4	267.2	269.0	261.7
7	N*a*	246.0	238.3	283.5	275.1	297.3	293.7
	N*b*	212.0	204.0	261.6	268.2	272.7	266.4
8	N*a*	255.1	246.8	292.0	284.1	300.5	293.0
	N*b*	221.0	212.2	267.6	259.7	300.3	293.6

However, as these molecules are symmetric a double protonation is feasible. The double PA at both N*a* positions was calculated as the difference between the unprotonated species and the doubly protonated ($2H^+$) form as shown in Scheme 5.1; additionally the difference from the singly protonated compound, (NaH^+) and the doubly protonated compound were determined, labelled ($NaH^+ + H^+$) in Table 5.5.



Scheme 5.1: Protonation scheme for doubly protonating compounds **1-8**.

Compound **1**, again presents the highest PA for double protonation as with single protonations. However, differences were observed between the resultant PAs derived using the two solvation methods. The results indicate that the positively charged derivatives (NaH^+) have less affinity towards a second protonation than the neutral compounds towards the first protonation when considered with the Onsager solvation technique (see Table 5.5).

However, the second PAs are larger when considering the PCM solvation technique, which would indicate that the positively charge derivatives have a higher affinity towards a second protonation.

Table 5.5: Proton Affinity (PA) in kcal mol^{-1}, in gas phase and solvated with Onsager and PCM techniques, without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the α_{1A} antagonists **1-8** (Scheme 5.1).

		In Vacuo PA		*Onsager PA*		*PCM PA*	
Ligand		(0 K)	(298 K)	(0 K)	(298 K)	(0 K)	(298 K)
1	$2\,H^+$	480.6	463.1	479.1	461.7	594.7	576.6
	$NaH^+ + H^+$	218.2	209.4	235.1	226.9	319.3	309.2
2	$2\,H^+$	466.9	449.9	467.9	450.8	584.2	565.5
	$NaH^+ + H^+$	212.8	204.3	199.3	190.5	290.6	286.1
3	$2\,H^+$	461.5	444.4	473.9	456.7	594.1	594.6
	$NaH^+ + H^+$	210.5	201.9	240.5	228.1	298.1	306.5
4	$2\,H^+$	477.1	459.7	478.4	460.9	593.1	576.5
	$NaH^+ + H^+$	218.3	209.6	192.9	183.4	298.1	306.6
5	$2\,H^+$	473.9	457.2	474.7	458.1	600.6	584.7
	$NaH^+ + H^+$	201.3	205.1	175.8	174.7	300.9	291.9
6	$2\,H^+$	459	442.5	472.7	455.5	590.4	574.8
	$NaH^+ + H^+$	208.6	200.4	185.0	176.8	295.8	286.9
7	$2\,H^+$	453.0	436.8	474.3	457.1	594.9	581.9
	$NaH^+ + H^+$	207.0	198.5	185.1	176.6	295.8	286.9
8	$2\,H^+$	470.7	453.7	470.8	454.3	#	#
	$NaH^+ + H^+$	215.6	206.9	178.8	170.3	#	#

a number of the optimisations failed to converge.

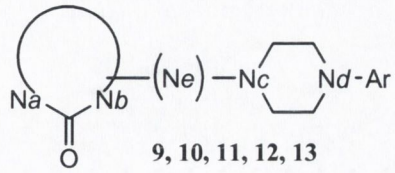
5.1.4 Proton Affinity Determination for α_{1A}–AR antagonists 9 to 18

Compounds **9** to **13** are represented by a generic scheme shown in Table 5.6, in which three moieties can be identified. At one end of the molecule a heterocyclic moiety containing a N*a*-CO-N*b* group, at the other end a piperazine ring, labelled N*c* and N*d* in Table 5.6 and in between, a heterogeneous bridge including a N atom (labelled N*e*). In compounds **9**, **10** and **13** the N atoms with the largest PA, in the gas phase, are those of the piperazine rings (N*c* and N*d*) with values ranging between 246 and 257 kcal mol^{-1} without thermal correction and between 237 and 255 kcal mol^{-1} at 298 K. For compounds **11** and **12**, the C=N*e* group, is most likely to protonate and the PA values obtained in the gas phase are similar (254 - 270 kcal mol^{-1} at 0 K and 236 - 261 kcal mol^{-1} with thermal corrections, respectively).

When solvation is included a general increase in the PA values are observed both at 0 K and 298 K (Table 5.6) indicating stabilisation of the protonated state. Similar patterns are detected between the PA of the solvated molecules and those in the gas phase. It was determined in the gas and solvated phases, that the N atoms in the piperazine ring (N*c* and N*d*) of compounds **9** to **13** are generally protonated, while those

of the heterocyclic ring (N*a* and N*b*) show the lowest PA values both at 0 K and 298 K. Exceptions are compounds **11** and **12** which show larger PA values at the N*e* atoms than at the piperazine N atoms.

Table 5.6: Proton Affinity (PA) in kcal mol^{-1}, in gas phase and solvated with Onsager and PCM techniques, without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the α_{1A}-AR antagonists **9-13**.



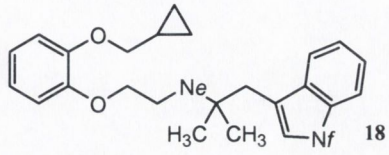
9, 10, 11, 12, 13

		In Vacuo		*Onsager*	
Ligand		PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
9	N*a*	207.6	199.4	222.2	214.6
	N*b*	207.4	199.6	220.1	212.7
	N*c*	246.3	236.8	251.0	241.6
	N*d*	246.0	236.3	276.8	267.8
	N*e*	219.5	211.1	228.5	220.3
10	N*a*	189.3	181.6	246.9	237.5
	N*b*	196.9	189.2	248.1	241.6
	N*c*	245.2	235.5	251.2	241.6
	N*d*	246.1	236.5	267.9	261.5
11	N*b*	220.1	211.1	244.1	218.6
	N*c*	239.9	231.9	241.3	233.3
	N*d*	246.9	236.9	245.9	236.3
	N*e*	263.8	236.8	262.4	253.3
12	N*c*	249.9	240.3	256.0	246.5
	N*d*	245.4	236.1	229.6	224.6
	N*e*	270.4	261.2	270.1	260.9
13	N*a*	204.9	196.7	241.5	233.8
	N*b*	206.7	198.8	231.1	223.3
	N*c*	256.9	248.2	265.9	257.4
	N*d*	245.8	235.9	246.0	236.1
	N*e*	253.8	243.8	258.9	249.2

For compounds **14** to **17** there is only one N atom present and for compound **18** there are two N atoms, one in an alkyl chain and the other in the indole ring (labelled N*e* and N*f*, respectively in the scheme of Table 5.7). The mono amine derivatives (**14** to **17**) follow the same trend when solvated as for the gas phase studies with compound **15** having the largest PA (Table 5.7), which can be explained by the presence of a near phenyl ring and the dioxane O atoms. Compound **16,** which has a slightly smaller PA than that of compound **15**, has a piranone system instead of the dioxane ring. The PA

values obtained for the first protonation of these compounds are in a similar range as the previous compounds. For compound **18**, under gaseous and solvated conditions the alkyl chain Nitrogen was most likely to protonate.

Table 5.7: Proton Affinity (PA) in kcal mol^{-1}, in gas phase and solvated with the Onsager technique, without and with thermal correction at 298 K calculated at B3LYP/6-31G* level for the α_{1A}-AR antagonists **14-18**.



18

		In Vacuo		*Onsager*	
Ligand		PA (0 K)	PA (298 K)	PA (0 K)	PA (298 K)
14	N	236.6	227.2	245.4	236.1
15	N	264.7	254.7	265.1	255.1
16	N	246.4	237.1	246.3	243.5
17	N	239.2	229.7	242.3	232.6
18	N*e*	241.9	232.5	258.8	249.2
	N*f*	229.3	220.1	254.9	246.0

5.1.5 Correlations between Proton Affinities and experimental biological affinity towards the α_{1A}-AR

A correlation has been found between the experimental affinities (K_i, nM) for α_{1A}-ARs antagonists, prazosin, alfuzosin, doxazosin, terazosin, **9** (5-methylurapidil), **10** (RS-100,975), **14** (WB4101) and **18** (RS-17053)[29] and the PA at 298 K without solvent effects (

Figure 5.4). The high correlation coefficient, $R^2 = 0.96$ (n=8), observed with the unsolvated PA values may indicate that the actual interaction between the ligand and the active site of the α_{1A}-AR occurring in the absence of water molecules.

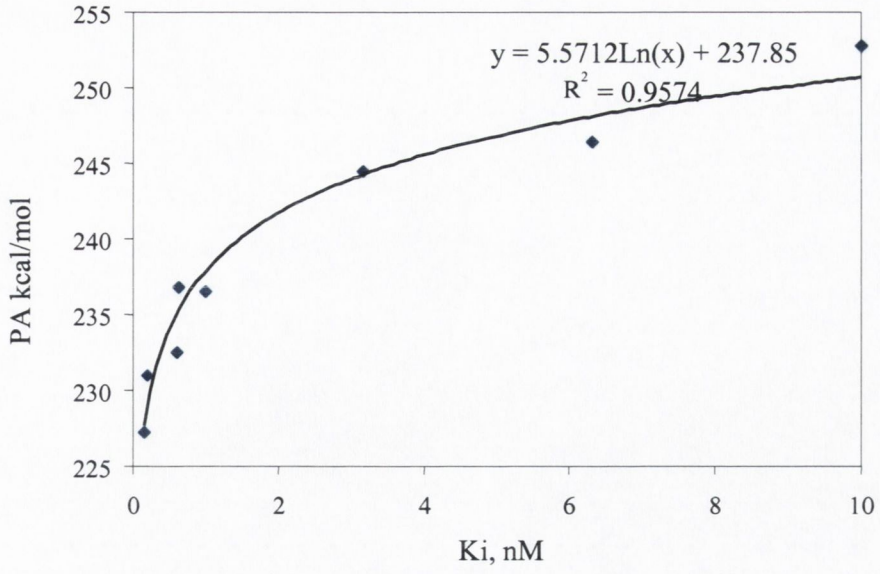


Figure 5.4: Correlation of the (K_i, nM) for α_{1A}-ARs of prazosin, alfuzosin, doxazosin, terazosin, **9** (5-methylurapidil) **10** (RS-100,975), **14** (WB4101) and **18** (RS-17053) with the thermally corrected PA without solvent effects.

Correlations between the PA values and the activity data for antagonists **1-8** were also studied. No clear correlation was found perhaps due to the nature of the biological data obtained. As no direct K_i over the α_{1A}-AR was measured effects from the different tissues could be used in testing the final K_i values.

5.2 Theoretical pKa Calculations

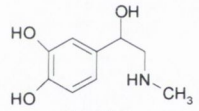
In aqueous solution, the more acidic proton will be lost first with increasing solution pH, so that the experimental pKa value corresponds to the preferred ionization site of a compound. A computational methodology for pKa predictions for small molecules can be based on an *ab initio* description of a generalized thermodynamic cycle, relating pKa to the gas-phase proton basicity (ΔG^0_{gas}), via the solvation energies (ΔG^0_{solv}) of the products and the reactants (see section 3.6). pKa predictions hinge on both accurate solvated PA values as examined in the previous section and on values for the solvation free energy of a proton, $\Delta G^0_{solv}(H^+)$. In water, experimental values of $\Delta G^0_{solv}(H^+)$ range from -254 to -262.5 kcal mol^{-1},[166] while recent computational works report a wider range of values between, -244.9 to -264.0 kcal mol^{-1}.[229] Theoretical studies into pKa prediction have explored the use of different computational methods, basis sets and solvation techniques. In the following sections we shall examine the values of $\Delta G^0_{solv}(H^+)$ required to replicate the available experimental pKa values for

our chosen ligands, utilising the B3LYP functional, the 6-31G* basis set and the Onsager and PCM solvation techniques.

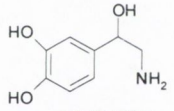
5.2.1 Theoretical Determination of $\Delta G^0_{solv}(H^+)$

Ligands that bind to catecholamine receptors contain an amino function that has a pKa greater than seven so that the ligand will be partly or fully protonated at neutral pH.[230] The experimental pKa values for adrenaline and noradrenaline are similar at 8.59[231] and 8.58[232] respectively. We examined the required value of $\Delta G^0_{solv}(H^+)$ to reproduce the experimental pKa values for the catecholamines utilising the thermodynamics cycle (see Table 5.8). At the B3LYP/6-31G* level of theory, utilising both the Onsager and PCM solvation techniques, the obtained $\Delta G^0_{solv}(H^+)$ values are outside the experimentally determined range.

Table 5.8: Comparison of theoretical $\Delta G^0_{solv}(H^+)$ in kcal mol^{-1}, obtained from structural optimisations using the Onsager and PCM solvation techniques to obtain the experimental pKa values for adrenaline and noradrenaline.



adrenaline



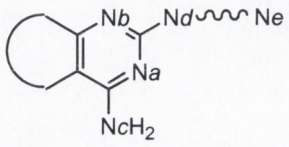
noradrenaline

	$\Delta G^0_{solv}(H^+)$	
Ligand	**Onsager**	**PCM**
Adrenaline	-220.4	-184.8
Noradrenaline	-224.9	-182.9

To further study the prediction of $\Delta G^0_{solv}(H^+)$ we examined the experimental data for our antagonist set. The experimental pKas of the clinical antagonists, terazosin (7.1)[233], prazosin (7.0), doxazosin (6.9)[234] and alfuzosin (8.13)[235] were obtained from the relevant pharmaceutical companies. For tamsulosin the experimental pKa value was not obtained, as it is under patent until 2006. The necessary values of $\Delta G^0_{solv}(H^+)$ to replicate these experimental pKa results for the N*b* position were within the experimental range, from -261.7 kcal mol^{-1} for doxazosin to -254.2 kcal mol^{-1} for prazosin, when utilising the Onsager solvation technique (see Table 5.9). However,

when utilising the PCM solvation technique, the $\Delta G^0_{solv}(H^+)$ values ranged from -280.0 to -290.1 kcal mol^{-1}, which is outside the predicted limits. Hence, our research has highlighted that the method by which the solvent (water) was treated is a contributing factor for the discrepancy between the pKa values.

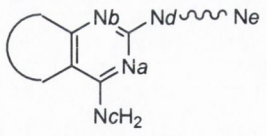
Table 5.9: Comparison of theoretical $\Delta G^0_{solv}(H^+)$ in kcal mol^{-1}, obtained from structural optimisations using the Onsager and PCM solvation techniques to obtain the experimental pKa values for the clinical antagonists.



	$\Delta G^0_{solv}(H^+)$	
Ligand	**Onsager**	**PCM**
Alfuzosin	-257.6	-290.3
Doxazosin	-261.7	-280.0
Prazosin	-254.2	-285.6
Terazosin	-255.0	-286.8

Each determined value for $\Delta G^0_{solv}(H^+)$ using the Onsager solvation technique was utilised in the thermodynamic cycle to predict theoretical values for the pKa of each clinical antagonist. The predicted pKa values for the N*b* position were in the range of 2-12 (see Table 5.10). It is evident that small changes in $\Delta G^0_{solv}(H^+)$ even within the accepted experimental range have significant effects on the predicted pKas. Furthermore, the theoretical calculations incorrectly predict the relative order of the pKas; with doxazosin having the largest predicted pKa followed by alfuzosin, terazosin and prazosin. Alternatively, the determined experimental order is alfuzosin followed by terazosin, prazosin and finally doxazosin at 6.9.

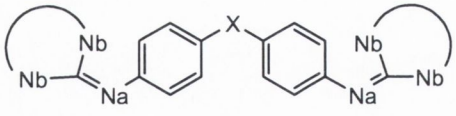
Table 5.10: Experimentally determined pKa values, and those calculated at B3LYP/6-31G* level for the N*b* position of the quinazoline α_1–AR antagonists, with the Onsager and PCM solvation methods.



	Exp. pKa	Predicted pKa			
Ligand		**Alfuzosin (-257.6)**	**Doxazosin (-261.7)**	**Prazosin (-254.2)**	**Terazosin (-255.0)**
Alfuzosin	8.13	8.13	9.9	4.5	5.2
Doxazosin	6.9	5.1	6.9	1.5	2.2
Prazosin	7.0	10.6	12.4	7.0	7.7
Terazosin	7.1	10.0	11.8	6.5	7.1

The experimental pKas for our antagonist set (**1-8**)[236] were determined following the procedure outlined in Aggawal *et al.*[237] The required values for $\Delta G^0_{solv}(H^+)$, covered a large range between -232 and –286 kcal mol^{-1} when utilising the Onsager solvation technique (see Table 5.11). However, the determined value for $\Delta G^0_{solv}(H^+)$ of compound **5**, at -250.5 kcal mol^{-1} was in a theoretically determined range for $\Delta G^0_{solv}(H^+)$.[229] The required values for $\Delta G^0_{solv}(H^+)$ using the PCM solvation technique at -277 to -294 kcal mol^{-1} were outside the experimental range.

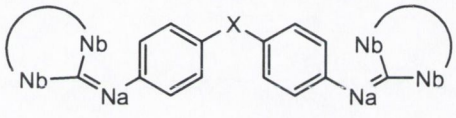
Table 5.11: Comparison of theoretical $\Delta G^0_{solv}(H^+)$ in kcal mol^{-1}, obtained from structural optimisations using the Onsager and PCM solvation techniques to obtain the experimental pKa values for the antagonists **1-8**.



		Expt. pKa	$\Delta G^0_{solv}(H^+)$	
Compound			**Onsager**	**PCM**
1	N*a*	10.80	-283.0	-288.9
	N*c*	2.80	-233.1	-277.0
2	N*a*	8.60	-273.1	-285.8
4	N*a*	9.80	-286.0	-294.0
5	N*a*	10.80	-250.5	-291.3
	N*c*	2.90	-232.1	-281.1
6	N*a*	9.40	-277.5	-286.5
8	N*a*	11.30	-281.4	-292.4

As a $\Delta G^0_{solv}(H^+)$ value of -250.5 kcal mol^{-1} reproduced the pKa of compound **5** and was within the accepted range of $\Delta G^0_{solv}(H^+)$ values, it was chosen for this class of compounds. Utilising this value in the thermodynamic cycle we again observed a poor agreement with the experimental results for this series of compounds with both solvation techniques (see Table 5.12). The pKas predicted, utilising the Onsager solvation technique are in the range of -2 to 36, while those predicted using the PCM solvation technique lie between 22 and 42. Such values are notably larger than the experimental pKa range, at 3 to 11.

Table 5.12: Experimentally determined pKa values, (Metrohn 632 pH-meter, Ω metrohm electrode, pH 0-14 / 0-80 °C, 3M KCl) and those calculated at B3LYP/6-31G* level for the α_{1A}–AR antagonists **1-8**, with the Onsager and PCM solvation methods.



			Predicted pKa	
Compound		**Expt. pKa**	**Onsager**	**PCM**
1	N*a*	10.80	34.7	39.0
	N*c*	2.80	-1.6	22.3
2	N*a*	8.60	25.2	34.5
4	N*a*	9.80	35.9	41.8
5	N*a*	10.80	10.8	40.8
	N*c*	2.90	-2.4	26.1
6	N*a*	9.40	29.3	35.9
8	N*a*	11.30	34.0	42.1

In summary, there are a number of limitations to the current implementation of theoretical pKa calculations, most notably in the accuracy of theoretical solvation techniques and in the lack of accurate experimental knowledge of $\Delta G^0_{solv}(H^+)$. The calculation of absolute pKa values is a challenging task since a difference of only 1.361 kcal mol^{-1} in the aqueous free energy of reaction results in an error of one pKa unit.[238] A value for $\Delta G^0_{solv}(H^+)$ can not be theoretically reproduced within accepted the experimental range for all cases. The approach of combining DFT and continuum solvation theory has been shown to accurately predict pKa in some cases[239] but to yield errors in calculated relative to experimental pKa values for other structural classes. Hence, it does not appear possible, at least at this level of theory and current solvation techniques, to extrapolate this technique for predicting pKas to all systems.

5.3 Conformational Analyses of Ligands

When the exact 3D structure of the receptor is unknown and a series of compounds that show an activity of interest have been identified, a form of *in silico* screening can be performed through pharmacophore searching. Determining conformations of antagonists, which selectively fit one receptor subtype but not the other subtypes, can be useful in antagonist design for more selective drugs. Bremner *et al.*,[73] employed the Apex-3D program and the Catalyst software to treat molecular structures as templates consisting of strategically positioned chemical functions that would bind to a receptor. In the α_1-AR pharmacophore models of Bremner *et al.*,[73] the main features involve a polar group, an aromatic group and basic nitrogen (Figure 5.5). The critical structural feature for α_{1A}-AR selectivity over the other α_1-AR subtypes is the distance between the basic nitrogen and the centre of an aromatic ring system, (aromatic distance of 5.2-5.8 Å for α_{1A}-AR and 6.2-7.8 Å for α_{1B}-AR) and to a polar group (polar distance of 6-8 Å for α_{1A}-AR and 5-6 Å for α_{1B}-AR. Later, Bremner *et al.*,[74] published the first pharmacophore model for α_{1D}-AR, polar distance of 4.5 Å and an aromatic distance of 5.4 Å is required.

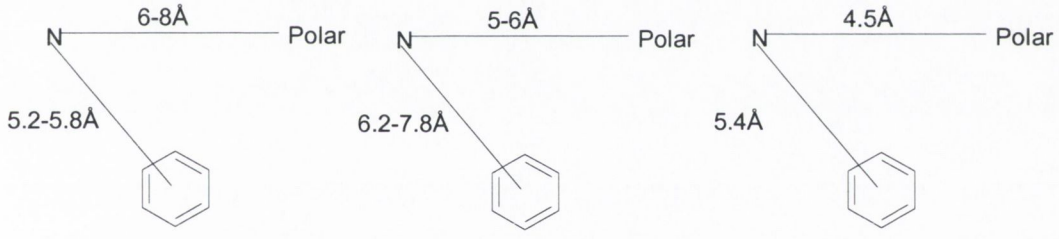


Figure 5.5: Pharmacophore requirements for the α_{1A}-AR, α_{1B}-AR and α_{1D}-AR subtypes, indicating the necessary polar and aromatic distances for subtype selectivity.

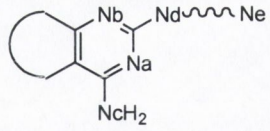
Barbaro *et al.*,[240] also determined an analogous pharmacophore model for α1-ARs utilising the Catalyst software, which consisted of a positive ionisable portion (*i.e.* 'basic N'), three hydrophobic features (*i.e.* 'aromatic ring') and a HB acceptor group (*i.e.* 'polar centre'). In addition, Li *et al.*[241] has recently generated a pharmacophore model for the α_{1A}-AR using a training set of thirty significant antagonists. This pharmacophore consisted of one positive ionizable element (*i.e.* 'basic N'), one hydrogen-bond donor, one aromatic element (*i.e.* 'aromatic ring'), and one hydrophobic group.

The key issue in pharmacophore screening is conformational flexibility of the ligands. As solvation effects can have an important influence on the conformation of a ligand the conformational analysis of the protonated form was followed by AM1 and subsequent B3LYP/6-31G* optimisations with the Onsager solvation technique. As the binding conformation of the ligands can not be assumed to coincide with the global minimum conformation, 20-30 of the low energy conformations were taken into consideration for each molecule.

5.3.1 Pharmacophore Study – Clinical Antagonists

The four clinical antagonists are experimentally 'universal' antagonists (not α_1–AR subtype selective) and hence it may be expected that the conformations should fulfil more than one pharmacophore subtype model. The nitrogens with the highest PA determined previously were taken as the protonation centre. Doxazosin, terazosin and prazosin have the protonated nitrogen (N*b*) as part of the quinazoline ring, while tamsulosin has the protonated nitrogen on the linker chain. A variety of polar groups were considered in the pharmacophoric measurements such as carbonyls, methoxy groups, trifluoro carbons and other nitrogens. The aromatic element was taken as the centre of aromatic rings. Amongst the quinazoline-type α_1-AR antagonists (excluding doxazosin), the antagonists have the correct 'polar distance' for α_{1A}-AR requirements in agreement with the pharmacophore model (see Table 5.13). Prazosin and terazosin exhibit no subtype selectivity according to the pharmacophore requirements while doxazosin appears to better fit the α_{1B}-AR and α_{1D}-AR pharmacophores than the α_{1A}-AR. Only alfuzosin and tamsulosin also have the correct range for the 'aromatic distance' requirement for the α_{1A}-AR. Due to its large conformational range tamsulosin appears to fulfil the requirements for all three subtypes, while alfuzosin fulfils that of α_{1A}-AR and α_{1B}-AR). In no instance, do any of these antagonists prefer one subtype over the other two subtypes indicating an *in silico* non-selectivity of the compounds in agreement with experimental findings. Screening of further receptor subtypes, e.g. the α_2-AR subtype, was prohibited due to a lack of pharmacophore models.

Table 5.13: Range of distances for polar and aromatic distance pharmacophore requirements for the clinical antagonists with the three α_1-AR subtypes. A **Y** denotes a successful fit, while **N** denotes a lack of fit for that subtype.



	Polar Range (Å)	Aromatic Range (Å)	α_{1A}	α_{1B}	α_{1D}
Prazosin	6.2 – 6.4	5.9 – 9.6	N	N	N
Terazosin	6.2 – 7.7	2.9 – 3.8	N	N	N
Doxazosin	2.0 – 5.1	2.9 – 11.3	N	Y	Y
Alfuzosin	5.7 – 12.5	4.6 – 11.7	Y	Y	N
Tamsulosin	2.9 – 8.3	4.5 – 7.9	Y	Y	Y

5.3.2 Pharmacophore Study - Imidazolidinium/Guanidinium Antagonists (1-8)

For compounds **1** to **8**, the N with the highest PA is that connected to the aromatic ring (N*a*). However, if this is considered as the 'basic N', some of these compounds will not fulfil the α_1-AR pharmacophores. Thus, both the N*a* and N*b* atoms from the guanidine moiety were considered as possible 'basic N'. Two possible scenarios are outlined in Figure 5.6 both of which take the bridging group X as the 'polar centre' and the further phenyl ring as the 'aromatic element'. Compounds **4** and **8**, having a CH_2 bridge were included for a non-polar comparison. Consequently, for derivatives **2** to **4** and **6** to **8**, the N*a* atoms were taken as the 'basic N' whereas for derivatives **1** and **5**, N*b* atoms were considered (see Figure 5.6).

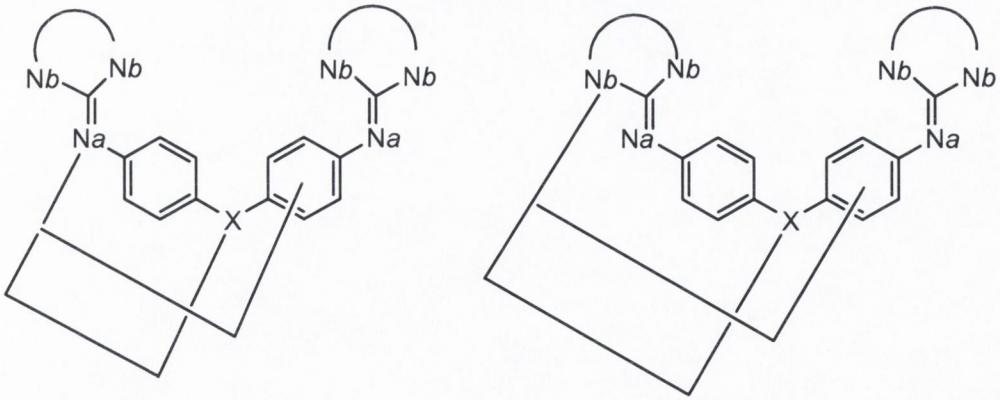
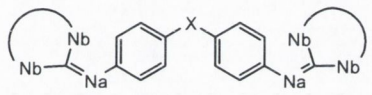


Figure 5.6: Antagonist models (**1-8**) with pharmacophore distances indicated.

Firstly, it should be noted that all compounds match the α_{1A}-AR pharmacophore although compounds **4** and **8**, do not fulfil the pharmacophore requirements as the CH_2

bridging group is not a polar centre. Secondly, for compounds **1** and **5** the 'polar distances' are in the correct range for the α_{1A}-AR pharmacophore (see Table 5.14) while not matching the α_{1B}-AR or α_{1D}-AR pharmacophores which is in agreement with the α_{1A}-AR affinity observed in the experimental study. Compound **2**, which exhibited the highest percentage inhibition of noradrenaline in our pharmacological studies, also selectively fulfils the α_{1A}-AR pharmacophore model (see Table 5.14). However, for compounds **3**, **6** and **7** along with fulfilling the α_{1A}-AR pharmacophore the α_{1B}-AR pharmacophore requirements are fulfilled, which indicates a possible non-selectivity.

Table 5.14: Range of distances for polar distance and aromatic distance pharmacophore requirements for possibly α_{1A}-AR selective antagonists. A **Y** denotes a successful fit, while **N** denotes a lack of fit for that subtype.



	Polar Range (Å)	Aromatic Range (Å)	α_{1A}	α_{1B}	α_{1D}
1	6.4 – 7.9	5.4 – 7.9	Y	N	N
2	6.4 – 7.1	3.8 – 9.2	Y	N	N
3	5.4 – 7.9	5.2 – 8.6	Y	Y	N
4	6.6 – 8.0	5.4 – 9.1	Y	N	N
5	6.3 – 9.6	5.4 – 10.4	Y	N	N
6	5.7 – 6.5	5.1 – 10.3	Y	Y	N
7	5.4 – 7.9	5.2 – 8.6	Y	Y	N
8	6.4 – 7.9	5.1 – 8.7	Y	N	N

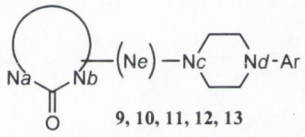
In summary, six of the bis-imino-imidazolidinium and bis-guanidinium antagonists fit the α_{1A}–AR pharmacophore and none the α_{1D}-AR. However, compounds **1**, **2** and **5** appear more selective for the α_{1A}-AR subtype as these antagonist conformations fit the α_{1A}-AR subtype and not the α_{1B}-AR and α_{1D}-AR subtypes and have an appropriate bridging polar group.

5.3.3 Pharmacophore Study – Antagonists 9-16

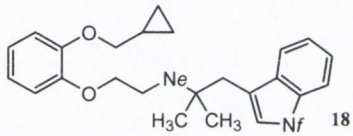
For the remaining antagonists, the basic nitrogens were taken as those with the largest PA as discussed earlier. The polar groups examined were the remaining nitrogens, polar oxygens and the CF_3 moiety, if present. Compounds **9** and **10** are thought to be selective for the α_{1A}-AR and were used in the initial training set for the α_{1A}-AR pharmacophore development. Nevertheless, from our pharmacophore screening

only compound **9** appears to fulfil the α_{1A}-AR and α_{1D}-AR subtype criteria while compound **10** fulfils all three pharmacophore models (see Table 5.15). For compound **11**, both the α_{1A}-AR and α_{1B}-AR pharmacophore requirements are fulfilled, despite experimental α_{1A}-AR selectivity. Compounds **12** and **13** fit all three pharmacophores and hence interactions at the other AR subtypes are equally likely under these criteria. Compounds **14, 16** and **17** do not fulfil any of the pharmacophore requirements, which indicates that these compounds may be non-selective antagonists, although compound **16** exhibits experimental α_{1A}-AR selectivity. Addition of another phenyl group in compound **15** results in α_{1B}-AR and α_{1D}-AR selectivity by increasing the range available for the aromatic distance. Finally, compound **18** an experimentally α_{1A}–AR subtype selective antagonist fulfils the α_{1A}-AR and α_{1B}-AR but not the α_{1D}-AR requirements.

Table 5.15: Statistical information for polar distance and aromatic distance pharmacophore requirements for α_1-AR antagonists **9-18**. A **Y** denotes a successful fit, while **N** denotes a lack of fit for that subtype.



9, 10, 11, 12, 13



18

Compound	Polar Range (Å)	Aromatic Range (Å)	α_{1A}	α_{1B}	α_{1D}
9	4.2 – 8.1	2.9 – 5.7	Y	N	Y
10	3.8 – 7.2	4.3 – 12.5	Y	Y	Y
11	4.8 – 9.7	3.5 – 10.4	Y	Y	N
12	3.2 – 6.7	2.7 – 10.5	Y	Y	Y
13	2.9 – 8.8	2.7 – 9.2	Y	Y	Y
14	2.6 – 5.6	2.1 – 5.1	N	N	N
15	2.8 – 5.3	1.4 – 6.5	N	Y	Y
16	2.8 – 5.9	2.0 – 5.1	N	N	N
17	3.1 – 3.2	3.4 - 7.3	N	N	N
18	4.8 - 7.6	3.1 - 6.6	Y	Y	N

In summary, the conformational flexibility of the antagonists was examined coupled with their fit to α_1–AR pharmacophores, to theoretically examine *in silico* α_{1A}–AR selectivity. The quinazoline antagonists and many of the literature compounds (**10** to **18**) were not found to selectively fulfil one of the different subtype pharmacophores, which would explain their non-selectivity in clinical studies. Compounds **1-8** were found to be active for the treatment of BPH in pharmacological tests. For compounds **1**, **2** and **5**, the α_{1A}-AR pharmacophore was selectively fulfilled, which may explain the good activity determined experimentally. Compound **9** exhibited α_{1A}-AR and α_{1D}-AR selectivity, which may be useful in the treatment of BPH.

5.4 Discussion / Conclusions

In this chapter, we presented an in depth study of a series of non-selective α_{1A}-AR and selective α_1-AR antagonists. Our aim was to elucidate the possible protonation and conformational states of these ligands upon reaching the binding site of the α_{1A}-AR. The PA values obtained for the first protonation of the ligands are ~250 kcal mol^{-1}. This included the 1-nitrogen atom of the quinazoline ring, the inner nitrogen (N*a*/N*c*) of the bis-imino-imidazolidinium and bis-guanidinium compounds, a nitrogen of the piperazine ring for compounds **9, 10** and **13,** at the heterogeneous bridge for compounds **11** and **12** and at the sole nitrogen for compounds **14-17**, and in the indole ring of compound **18**. In future chapters the nitrogen with the highest determined PA was protonated for all antagonists.[242] In addition, a good correlation was found between the PA values corresponding to the first protonation in gas phase of some of the compounds studied and their corresponding experimental affinity constants K_i, which measure the actual affinity of the compounds for the α_{1A}-AR.

Accurate, computational, solvation methods are hampered by two interconnecting problems, the system size and the complex molecular environment. While the Onsager solvation technique has the advantage of simplicity and rapid computation times there are some limitations to its use. The Onsager method can fail for compounds were the electron distribution is poorly described by a dipole moment. The Onsager method is also only appropriate for relatively compact molecules that might reasonably be placed in a spherical cavity. The inclusion of solvation by means of the Onsager and PCM models, yielded stabilization in the PA values obtained. Therefore the protonation of these compounds in a solvated medium will be energetically favoured in all the cases.[243] While the PCM solvation technique yields a more satisfactory description of the solute molecule it is for the larger molecules prohibitively computationally expensive and often associated with technical difficulties.

The theoretical procedure to determine pKas involved DFT calculations to characterise the gas-phase system and then solvation by coupling to continuum solvent using self-consistent reaction field cycle. The pKa values obtained using the proposed thermodynamic cycle for the current clinical antagonists were in agreement with some experimental results, but not for other series of compounds. pKa calculations required reliable and highly accurate gas-phase protonation/deprotonation energy calculations as

well as solvation energy calculations for both products and reactants. The theoretical method employed in the present calculations (B3LYP/6-31G*) is, by necessity, more approximate than the most extended methods applied to study smaller systems. Furthermore, it is important to mention that there is no solvation model, to the best of our knowledge complete enough to ensure the proper description of pKa for all compounds.

Since the global minimum energy conformation of a ligand does not necessarily need to be the pharmacologically active one, the conformational freedom of the ligands under investigation was taken into consideration. The current clinical antagonists were not found to selectively fulfil the α_{1A}-AR pharmacophore model, which would theoretically explain their non-selectivity in clinical studies. Another possibility is that the antagonists are interacting in a different site than those of the training set for the pharmacophore and hence a fit to this pharmacophore may not be enough to estimate the viability of the antagonist models. For compounds **1** and **5**, once the alternative nitrogen site (N*b*) was used the α_{1A}-AR pharmacophore model was fulfilled, which may explain the good activity determined experimentally. Compounds **2** and **6** exhibited the best activity and fitted the α_{1A}-AR pharmacophore model, although compound **6** also fulfilled the α_{1B}-AR pharmacophore. However, compounds **3** and **7** also fulfilled the α_{1A}-AR pharmacophore model while exhibiting poor activity. Compounds **4** and **8** exhibited high activity even though the α_{1A}-AR pharmacophore criteria were not met *i.e.* there was no bridging polar group in these compounds. Of antagonists **9-18**, compound **9** exhibits selectivity, in terms of pharmacophore screening for the α_{1A}-AR and α_{1B}-AR subtypes over the α_{1B}-AR. Compound **9** may prove an interesting agent as antagonist inhibition at the α_{1A}-AR subtype is useful in alleviating obstructive symptoms in cases of BPH, while the blockade of α_{1D}-AR may also be important for patients with irritative symptoms by reducing bladder overactivity.[17] The remaining antagonists were not found to selectively fulfil the α_{1A}-AR pharmacophore model.

All this information will be useful for the design and future synthesis of new and more selective α_{1A}-AR antagonists that could find application for the treatment of BPH.

Chapter VI: Agonist / α_{1A}-AR Complexes

– Towards Receptor Activation

"In theory, there is no difference between theory and practice.;
In practice, there is."

Chuck Reid

6.0 Introduction

Obtaining structural information about the interactions of agonists with α_{1A}-ARs has been hampered by the lack of a GPCR crystal structure in the activated state.[244] Molecular modelling studies can facilitate the integration of experimental observations and biophysical data into a mechanistic scheme for receptor structure and function and shall now be pursued. In this chapter, we will utilise our homology models developed in Chapter IV to examine the interactions of agonists with the α_{1A}-AR and the induced conformational changes.

The main features of the catecholamine agonists include a basic nitrogen, a *meta*, *para* and *beta* hydroxyl and for Adrenaline a N-methyl group (see Figure 6.1). Mutagenesis studies of the α_{2A}- and β_2- ARs suggest that the amino group of the catecholamines forms an electrostatic interaction with the carboxylate side chain of an Asp in TM-III (Asp-106 in α_{1A}-AR), which is highly conserved in all aminergic GPCRs.[245,246] More recent studies have focused on the molecular interactions of the catecholamines with the α_{1A}- and α_{1B}- AR subtypes,[90,247,248] for which Ser residues in TM-V may interact with the catechol hydroxyl groups, Ser-188 with the *meta* hydroxyl group and Ser-192 with the *para* hydroxyl group.[100]

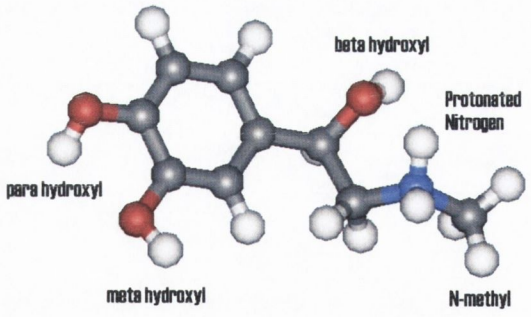


Figure 6.1: Diagram indicating the principal interacting features in Adrenaline.

The detailed activation mechanism for Class A GPCRs is still unknown, but has been postulated for a number of receptors, including the α_{2A}- and β_2- ARs (for reviews[249,250]). Agonist binding may induce specific rearrangements in Pro kink regions or rigid body motions of TM-V. Studies also support an important role for TM-VI in the network of conformational constraints required to maintain the receptor in the inactive state and a movement of TM-VI relative to TM-III may result from agonist induced receptor activation. [251]

The direct use of the inactive bovine Rhodopsin crystal structure (*1l9h*) as a template for modelling the active receptor conformation remains questionable.[99,191] To circumvent this issue, Bissantz *et al.*[252] manually rotated TM-VI by 30° anticlockwise around its helical axis to produce an activated form of the dopamine D_3, the β_2-AR and the δ opioid receptors. Recently, an *in silico* active form of the Rhodopsin structure was developed by isomerisation of the 11-cis-retinal chromophore, followed by constrained dynamics. From this structure a homology model of the 5-HT_{2A} receptor was developed.[253] However, Carmine *et al.*,[254] suggested that the agonist-binding pocket of the β_2-AR is not rigid but is dynamically formed as the agonist builds an increasing number of contacts with the receptor. As the receptor is "en route" to the active form, it could induce or change the configuration of the agonist-binding cavity. Therefore, the key agonist binding interactions and the active receptor conformation may develop in parallel, necessitating a dynamical approach.

In this work, we have developed a stepwise protocol to examine the interactions between the catecholamine agonists and our developed α_{1A}-AR models. Our computational approach to model the α_{1A}–AR conformation induced by the interaction with a ligand involves:

1) Docking of agonists (AD and ND) to examine their interactions with our previously developed homology-derived models of the α_{1A}-AR in Chapter IV.

2) Performing MD simulations (1 ns) in a $H_2O/CHCl_3/H_2O$ membrane mimic to refine the agonist/α_{1A}-AR complexes. Exploration of the structural changes that occur during the molecular dynamics simulations.

3) Further docking of the agonists into the 'agonist-induced' receptor forms developed in step 2.

4) Further MD simulations (2 ns) of the agonist/α_{1A}-AR complexes to determine agonist binding modes and the agonist induced receptor structures.

A repeated cycling of steps 3 and 4 can be used to further refine the binding site.

6.1 Docking of agonists with the α_{1A}-AR Receptor Models

A minimum requirement for agonist activity is that the ligand sterically fit the region of buried volume inscribing the active site cavity, with some allowance for induced fit and conformational changes. From experimental mutagenesis studies the catecholamines are predicted to bind in a pocket formed by residues in TM-III through TM-VI, anchored by an interaction between the side chain of Asp-106 in TM-III and the protonated nitrogen of the catecholamine. The probes of the PASS utility represent potential binding sites for ligands of arbitrary shape. As can be observed in Figure 6.2, there is a marked difference in the probes of the two initial α_{1A}-AR models in this site. In Model I, the probes form a branched conformation to extend to the three residues (Asp-106, Ser-188 and Ser-192), while Model II has a more compact binding site. As conformational differences have emerged in the two binding sites both α_{1A}-ARs (Model I and Model II) were used for subsequent docking studies, as the different start conformations allowed for greater exploration of binding possibilities.

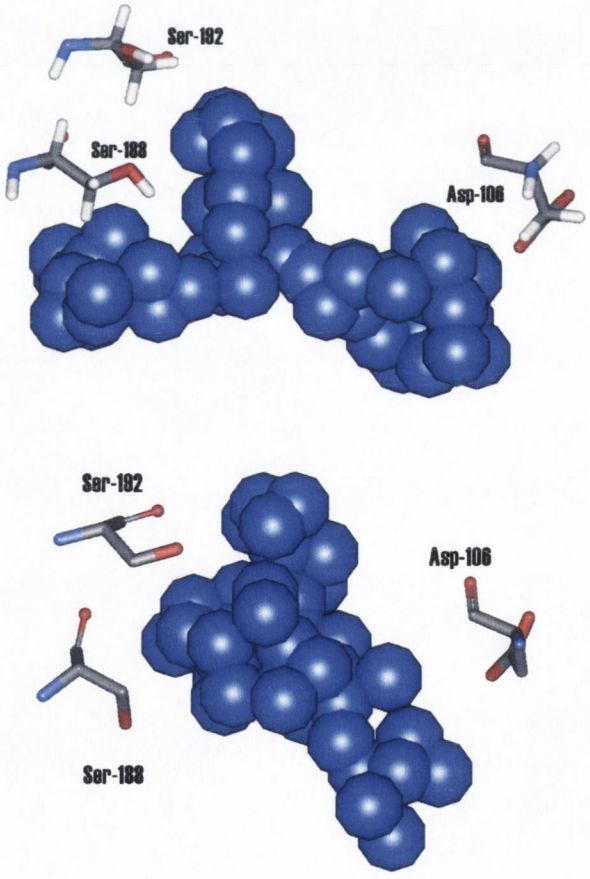


Figure 6.2: PASS probes for the binding site of Model I (upper) and Model II (lower).

The conformation of the cavities in the refined, α_{1A}-AR structures (Final I and Final II, from Chapter IV), were also examined but were determined to be unsuitable for ligand docking. In Final I, the binding site has closed completely with interactions forming between Asp-106 and residues Ser-188 and Ser-192. Alternatively, with Final II, the site has expanded compared to the homology model and there are probes clustered about Asp-106 and Ser-188 (see Figure 6.3).

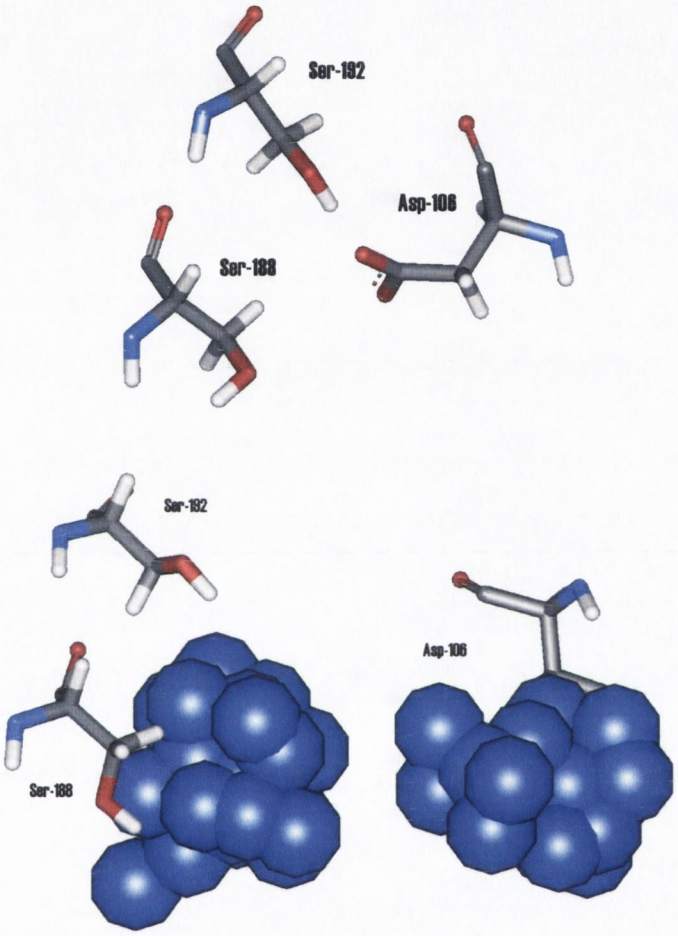


Figure 6.3: PASS probes for the binding site of Final I (upper) and Final II (lower).

To produce 'agonist-induced' receptor conformations the following protocol was pursued (see Scheme 6.1) utilising the outlined docking procedures (see Chapter III). The complexes were further refined through MD simulations (1 ns) to produce four 'agonist-bound' receptor conformations, whose interactions were monitored through the formation and breaking of HBs and whose structural changes were analysed through RMSD analysis.

Model I	Dock	I-AD	MD	FI-AD
	⟶	I-ND	⟶	FI-ND
Model II	+AD	II-AD	(1ns)	FII-AD
	+ND	II-ND		FII-ND

Scheme 6.1: Computational pipeline illustrating the development of agonist-bound receptor structures in Chapter VI.

6.1.1 Hydrogen Bond Analysis for the initial Agonist / α_{1A}-AR Complexes

For Model I complexed with adrenaline (termed I-AD) a number of interactions consistent with mutagenesis studies were observed. The docked HB distance from the protonated nitrogen to the carboxylic group of Asp-106 is 3.23 Å (see Figure 6.4). The *meta* hydroxyl is 4.02 Å from the oxygen of Ser-188, while the *para* hydroxyl is 2.16 Å from Ser-188.

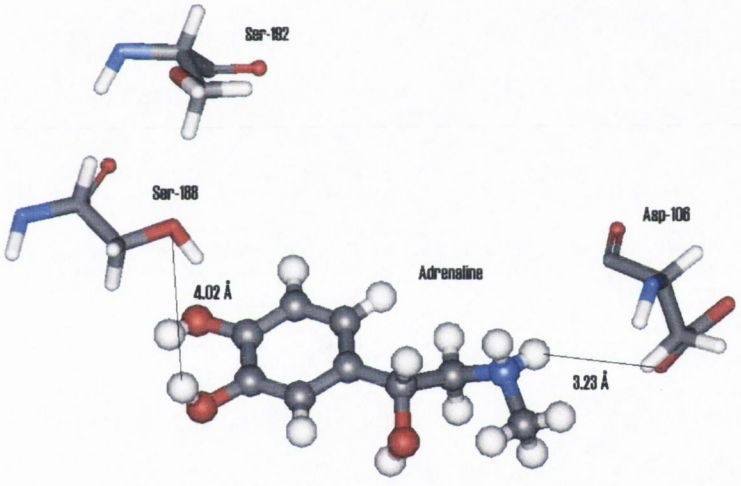


Figure 6.4: Adrenaline docked with Model I, termed complex I-AD.

For the alternative Model II complexed with adrenaline (termed II-AD) the HB distance from the protonated nitrogen to the carboxylic group of Asp-106 is 1.14 Å (see Figure 6.5). The *meta* hydroxyl is 2.81 Å from the oxygen of Ser-188, while the *para* hydroxyl is 3.20 Å from the oxygen of Ser-192. However, the *para* hydroxyl has a closer interaction with Ser-188, at a distance of 2.07 Å while the *beta* hydroxyl is close to Asp-106 (1.99 Å) but is not appropriately orientated.

closer interaction with Ser-188, at a distance of 2.07 Å while the *beta* hydroxyl is close to Asp-106 (1.99 Å) but is not appropriately orientated.

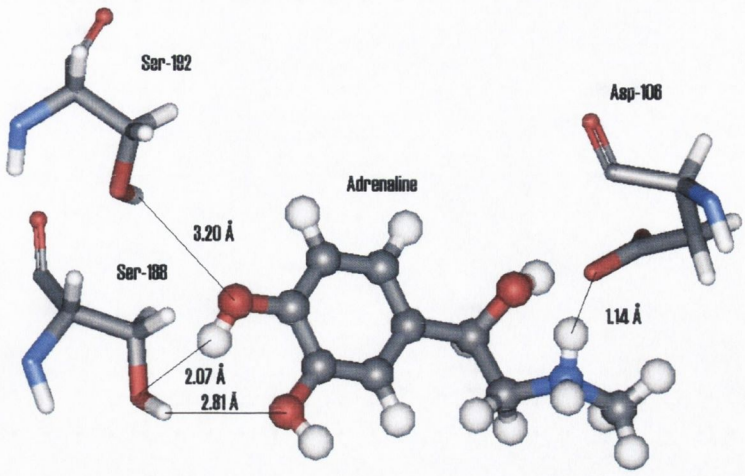


Figure 6.5: Adrenaline docked with Model II, termed complex II-AD.

Considering the noradrenaline complexes, for Model I complexed with noradrenaline (termed complex I-ND), the docked distance from the protonated nitrogen to the carboxylic group of Asp-106 is long at 6.09 Å, as the carboxylic group is orientated away from the ligand (see Figure 6.6). The *para* hydroxyl is 2.93 Å from the oxygen of Ser-188. For the I-ND complex there is an additional interaction between Ser-188 and Ser-192 at 2.09 Å.

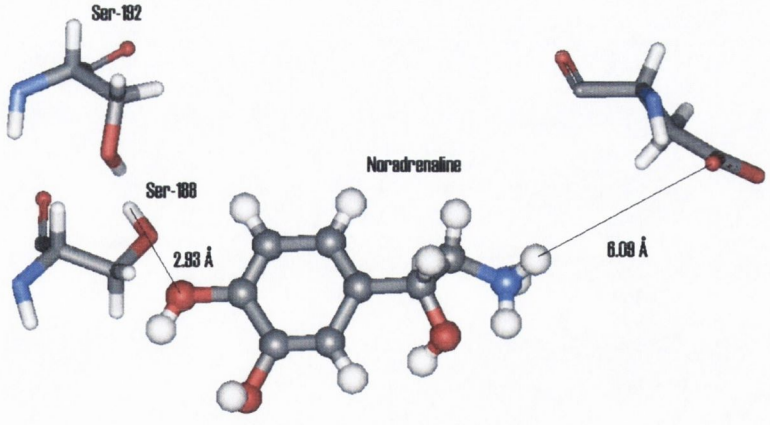


Figure 6.6: Noradrenaline docked with Model I, termed complex I-ND.

Finally, for Model II complexed with noradrenaline (termed complex II-ND), the HB distance from the protonated nitrogen to the carboxylic group of Asp-106 is 2.16 Å. The *meta* hydroxyl is 3.67 Å from the oxygen of Ser-188, while the *para* hydroxyl is

1.92 Å from the oxygen of Ser-188 and 5.00 Å from the oxygen of Ser-192 (see Figure 6.7).

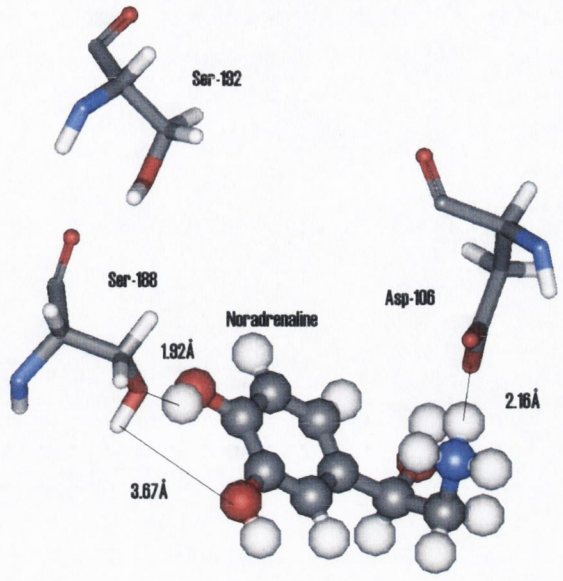


Figure 6.7: Noradrenaline docked with Model II, termed complex II-ND.

In summary, for three of the docked complexes (the exception being complex I-ND) interactions occur with Asp-106 of TM-III and Ser residues in TM-V in agreement with the available mutagenesis data. Our docking analysis is also in agreement with Pedretti *et al.*,[191] who recently determined that Asp-106 and the Ser residues in TM-V of α_{1A}-AR, formed a network of HBs with noradrenaline. Furthermore, our model complexes indicate that although EC-II is inserted into the TM region towards the active site, it does not form interactions with the docked agonists.

6.2 Molecular Dynamics Simulations of agonist/α_{1A}-AR Complexes

The initial four agonist/α_{1A}-AR complexes provided starting structures for further structural refinement through MD simulations (1 ns). These simulations allow for full receptor flexibility and optimisation of the initial complexes.

6.2.1 Force Field Parameters for Agonists

The MM atom types and parameters concerned with the agonists were obtained from the General Amber Force Field (GAFF[255]).[i] Two additional angular terms, *ca-c3-*

[i] GAFF parameters used for adrenaline / noradrenaline included *ca* - an Sp^2 C in pure aromatic systems, *ha* - a H bonded to an aromatic carbon, *oh* - an oxygen in a hydroxyl group, *ho* - a

oh and *ca-c3-h1*, were included by analogy to similar parameters already present in GAFF. In addition, two improper dihedral terms, (*ca-ca-ca-ha, ca-ca-ca-oh*) were added to enforce the planarity of the aromatic rings (1.1 kcal mol^{-1}, 180.0°). The two equilibrium angles and the corresponding force constants were fitted to best reproduce the geometries and IR frequencies obtained from DFT calculations. The fitted values for *ca-c3-oh* were 67.1 kcal mol^{-1}, 110.7° and for *ca-c3-h1* were 47.3 kcal mol^{-1}, 109.1°. To test the validity of the agonist parameters, a comparison was made between experimentally (Expt.) and theoretically (DFT and MM) determined structures and frequencies.[256,257] The experimental structures of adrenaline and noradrenaline were optimised both at a DFT (B3LYP/6-31G*) and at a MM (Amber FF[258]) level. A high structural correlation was exhibited with the GAFF parameters reproducing geometric distances to a RMSD of 0.10 Å from the crystal structures (see Table 6.1).

Table 6.1: RMSD (in Å) of the expt. catecholamine structures optimised with DFT (B3LYP/6-31G*) calculations and fitted MM calculations.

Ligand	**Expt. vs DFT**	**Expt. vs MM**
Adrenaline	0.08	0.10
Noradrenaline	0.06	0.07

There was also a high correlation between the frequencies derived from the MM calculations (Amber FF[258]) and the experimental information available for adrenaline, with an RMSD of 23.18 cm^{-1} from the optimised structure (MM) to experiment. Dynamical simulations were used to further assess the quality of the parameters (150 ps for each agonist in a water box) and both systems remained stable throughout. Such an examination of the MM parameters aims to ensure that the MM calculations produce appropriate molecular geometries for the compounds.

6.2.2 Hydrogen Bond analysis for Agonist / α_{1A}-AR Complex Simulations

Over the MD simulations (1 ns), HBs forming and breaking between the ligand and the receptor were monitored. For simulation I-AD, Figure 6.8 indicates the trends in HB occurrence between adrenaline and selected residues in the receptor. Despite a lengthening of the donor-acceptor distance (d[D...A]) in the equilibration period the

---

Hydroxyl group, *c3* - a Sp^3 C, *h3* - a H bonded to aliphatic carbon, *n4* – a Sp^3 N with four connected atoms and *hn* - a H bonded to nitrogen atoms.

protonated nitrogen interacts steadily with Asp-106, throughout the simulation and sporadically with Cys-176. The *meta* hydroxyl of AD forms an interaction with Ser-188, over the equilibration period which breaks and reforms by 300 ps. The *para* and *beta* hydroxyls do not form lasting interactions. The presence of the N-methyl group in adrenaline may improve the binding affinity compared with noradrenaline by making additional correlated contacts after 300 ps with Thr-174 of EC-II and Trp-102 of TM-III.

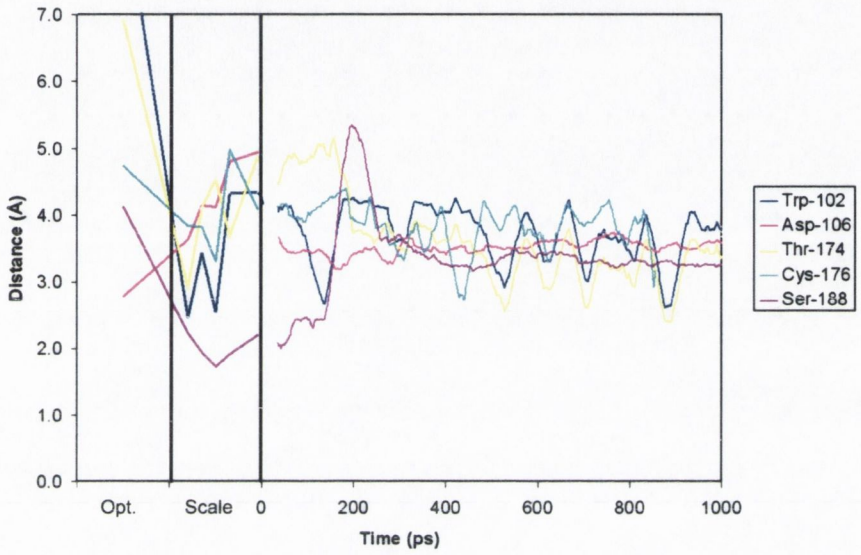


Figure 6.8: Hydrogen Bond analysis in Å for simulation I-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Trp-102 is d[C...O], for Asp-106 is d[N...O], for Thr-174 is d[C...O], for Cys-176 is d[N...O] and for Ser-188 is d[O...O].

For simulation II-AD, HB interactions occur between the protonated nitrogen and Asp-106, which shortens after 400 ps (see Figure 6.9). The protonated nitrogen also forms an interaction with Glu-180 of EC-II over the equilibration period. This is coupled with a lengthening of the interaction involving Ser-192 of TM-V. The RMSDs of TM-V lengthened in the region of 300-500 ps, and correspondingly an interaction between Ser-188 and the *meta* hydroxyl, lengthens after 300 ps, while the *para* hydroxyl interacts with Ser-192, after this point. Subsequently, for simulation II-AD, most HB interactions have stabilised by 500 ps.

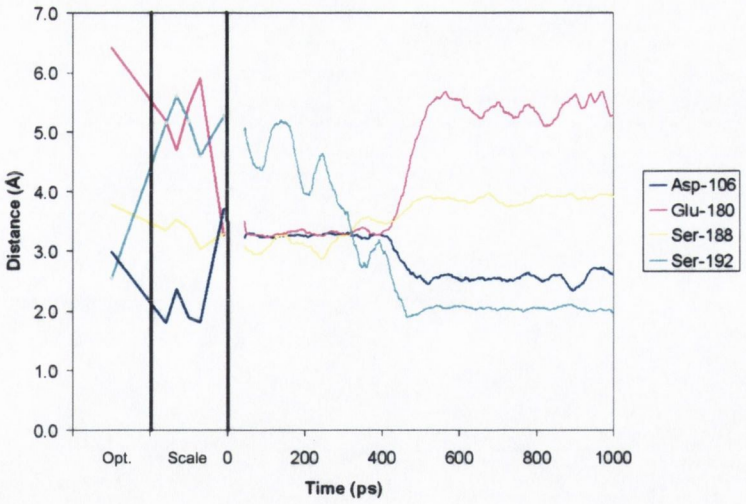


Figure 6.9: Hydrogen Bond analysis in Å for simulation II-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Asp-106 is d[N...O], for Glu-180 is d[N...O], for Ser-188 is d[O...O] and for Ser-192 is d[O...O].

For simulation I-ND, the protonated nitrogen of Noradrenaline interacts steadily with both carboxylate oxygens of Asp-106, (termed Asp-106a and Asp-106b in Figure 6.10). The protonated nitrogen also forms interactions with Cys-176 and Gln-177 (labelled Gln-177a, in the figure), of EC-II indicating their close involvement in the binding site. The *meta* and *para* hydroxyls form no substantial interactions over the simulations. However, the interaction between the *beta* hydroxyl and Gln-177 (labelled Gln-177b), lengthens at ~400 ps, while interactions of the *beta* hydroxyl with Asp-106 gradually increase at this point. This interaction is in agreement with the study of Nyronen *et al.*,[93] where the β-OH group is positioned to form an equivalent HB with Asp-113 of the α_{2A}-AR.

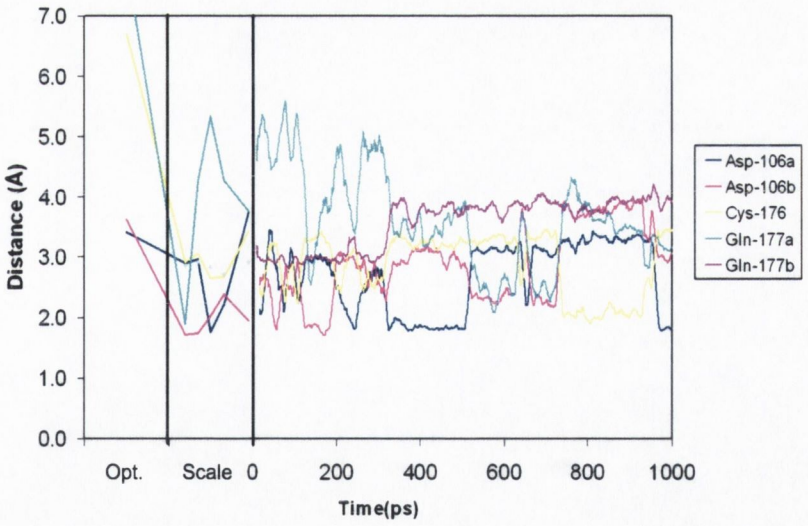


Figure 6.10: Hydrogen Bond analysis in Å for simulation I-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Asp-106 is d[N...O], for Cys-176 is d[N...O] and for Gln-177 is d[O...O].

Similarly, for simulation II-ND, an interaction involving the protonated nitrogen and Asp-106, forms over the equilibration period and interacts over the production run, see Figure 6.11. The protonated nitrogen also interacts with Ser-188, which gradually lengthens after 400 ps. The *meta* and *para* hydroxyl do not form any lasting interactions, while the *beta* hydroxyl interacts with Cys-176.

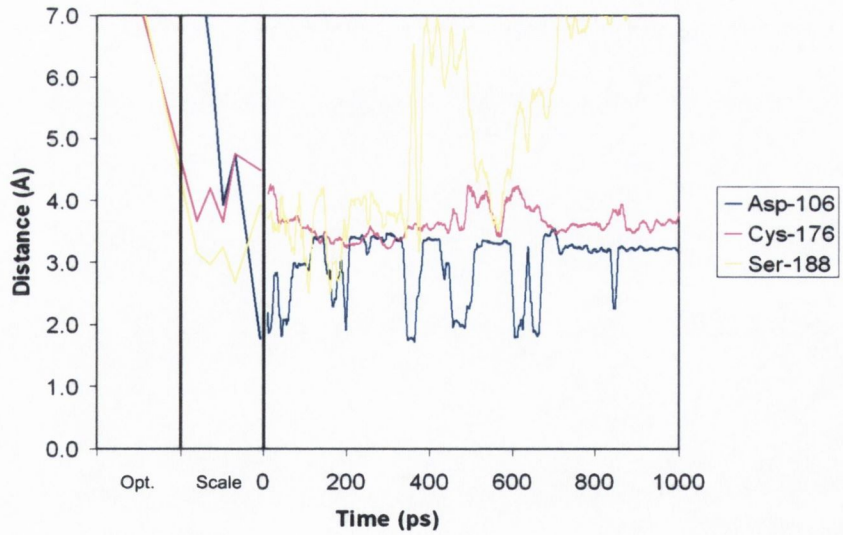


Figure 6.11: Hydrogen Bond analysis in Å for simulation II-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Asp-106 is d[N...O], for Cys-176 is d[O...O] and for Ser-188 is d[N...O].

To summarise, there are interesting time dependent changes to the HB distances as the complexes restructure to accommodate the ligand. The main receptor regions involved in interactions include TM-III, TM-V and EC-II residues. The most consistent interaction occurs between the positively charged amine group of the agonists and the negatively charged side chain carboxyl of Asp-106 in TM-III of the α_{1A}-AR. In general, this interaction is strong and anchors the amine group close to Asp-106. Interactions are developed with Ser-188 for simulations I-AD and II-ND, while that with Ser-192 is lost. At least one interaction with an EC-II residue, Cys-176, Gln-177 or Ile-178 is formed over each of the simulations, which again highlights the importance of this loop in the binding site.

6.2.3 Structural Analysis over the Molecular Dynamics Simulations of agonist / α_{1A}-AR complexes

The time dependent RMSD data of the TM helices were considered to examine structural changes induced by the agonists on the receptor. Little structural change occurred over the optimisation step, although this is often the only procedure used in computational binding studies.[252] A marked increase in the RMSD was observed for all simulations over the heating step of the equilibration runs, as the receptor structures moves from a 'rhodopsin-like' conformation to a ligand induced conformation. In the following analysis, the trendlines for each helix are considered over the production runs. No marked structural changes from their original inactive states occurred for TM-VIII over the production runs of the four simulations, with greater deviations observed in the transmembrane helical regions (see Figure 6.12 - Figure 6.15).

For simulation I-AD, the initial structural changes through optimisation and heating begin to stabilise over the equilibration periods. However, a number of helices experience ongoing drifts in RMSD over the production run, for example, TM-I underwent a decrease in RMSD between 100-300 ps, TM-IV underwent an increase in RMSD between 600 – 800 ps, while TM-VII drifted between 300 – 400 ps (see Figure 6.12). As noted in the previous section a HB interaction forms between Ser-188 and the *meta* hydroxyl at ~100 ps, which corresponds to a slight movement of TM-V at this time. The overall RMSD analysis indicates that a structural rearrangement has occurred due to the presence of an agonist, which has stabilised by 800 ps of simulation.

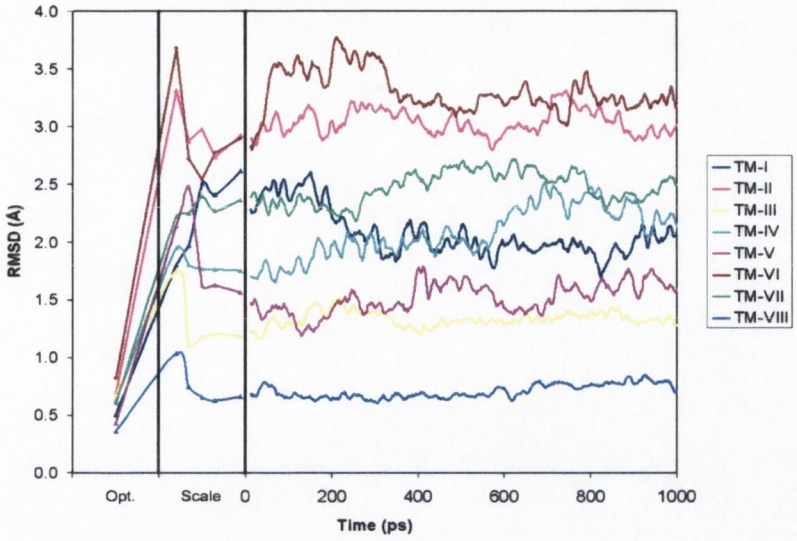


Figure 6.12: RMSDs in Å for simulation I-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

The structural deviations were not as marked over the equilibration period for simulation II-AD. An exception is TM-IV which rises sharply in RMSD over equilibration, exhibits a small decrease over the first 200 ps of the production run, before stabilising after 600 ps (see Figure 6.13). Helical structural fluctuations were observed for simulation II-AD, with TM-I, and TM-V also structurally changing between 300-600 ps, which coincides with the changes in the HB patterns as observed in Figure 6.9. The remaining helices fluctuate around mean RMSD values throughout and have structurally stabilised by 800 ps.

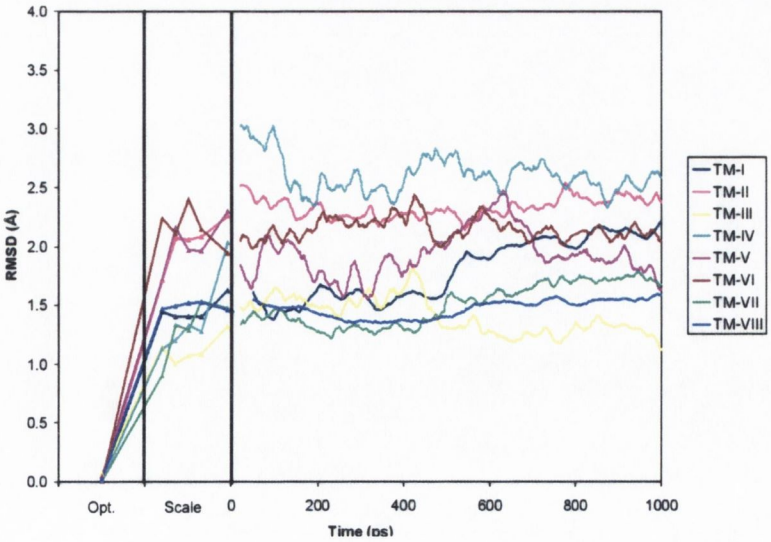


Figure 6.13: RMSDs in Å for simulation II-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

For simulation I-ND, after equilibration, most of the structural rearrangements occur in the region of 200 – 500 ps (see Figure 6.14). Small rises in RMSD were observed for TM-IV and TM-V in this region. TM-III undergoes a gradual RMSD drift over the first 500 ps of the simulation, corresponding to an increase in interactions of Asp-106 with the *beta* hydroxyl of noradrenaline at this time. The RMSD changes for most helices have structurally stabilised by 800 ps, at which time a stable ligand-induced conformation has developed.

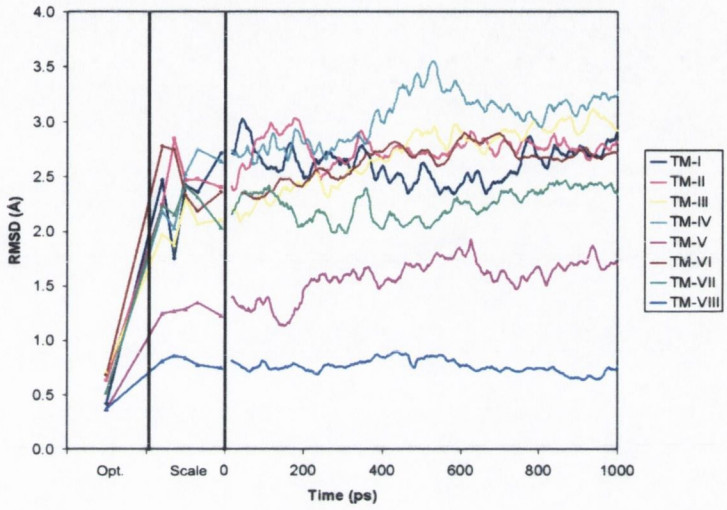


Figure 6.14: RMSDs in Å for simulation I-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

Finally, for simulation II-ND, the largest structural deviations were observed for TM-VI over the equilibration period (Figure 6.15). TM-III and TM-V exhibit gradual structural changes over the first 400 ps of the production run, which corresponds to the lengthening of the interaction between noradrenaline and Ser-188. All the remaining RMSDs for the helices deviated around their mean values and have stabilised by 800 ps.

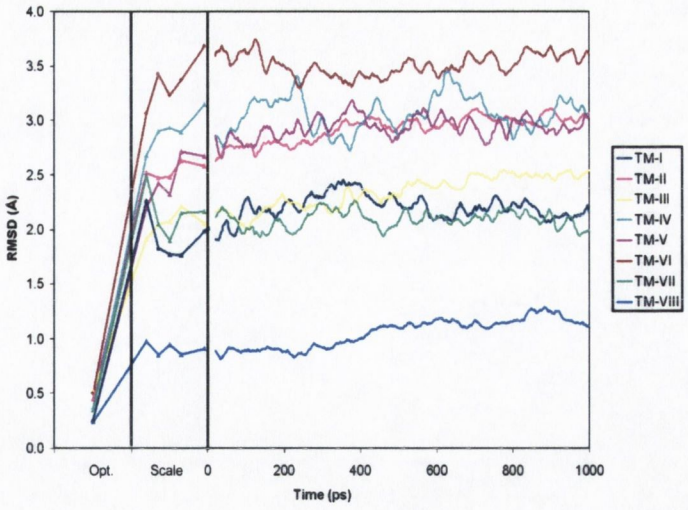


Figure 6.15: RMSDs in Å for simulation II-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

In summary, over the course of the four production runs the deviations from the initial uncomplexed structures were measured via the RMSD. In general, for all four simulations structural changes have occurred and stabilised over the 1ns production runs corresponding to a stabilised induced fit effect of the ligand on the receptor structure.

6.2.4 Structural Analysis after Molecular Dynamics Simulations of agonist / α_{1A}-AR complexes

After 1 ns of MD simulations, the final structures of the four complexes were averaged over the last 200 ps of the production runs, optimised and termed FI-AD, FI-ND, FII-AD and FII-ND. Both final adrenaline complexes contain interactions involving the protonated nitrogen; complex FI-AD contains a HB to Asp-106 and Cys-176, while complex FII-AD interacts with Asp-106 and Ile-178. Hence, both final complexes generated novel interactions with EC-II residues but have lost Ser interactions with TM-V. For the final noradrenaline complexes FI-ND, the agonist interacts with Asp-106, Cys-176 and Gln-177, while for complex FII-ND interactions involve Asp-106 and the only remaining TM-V interaction with Ser-188. As with the adrenaline complexes, the novel interactions involve EC-II residues. Hence, a major change in the final binding sites lies in the role of the EC-II loop, which further inserts into the TM framework and is involved in binding interactions.

The final structures of the two uncomplexed systems, Final I and Final II, as reported in Chapter IV, were compared with those of the agonist complexed structures developed in this study. The radius of gyration of FI-AD (23.3 Å) and for FI-ND (22.8 Å) is comparable to that of Rhodopsin (21.96 Å) and more compact than Final I (24.32 Å). Alternatively, the radius of gyration of FII-AD (24.51 Å) and FII-ND (25.03 Å), have increased from the 23.94 Å of Final II. The structural movement from the final uncomplexed to the final complexed receptor forms was again monitored through the RMSDs. A large overall RMSD is observed for the Cα atoms, in the range of 8.03 - 8.56 Å (see Table 6.2). Hence, structural changes have occurred for the agonist-bound receptor structures relative to the uncomplexed receptor forms. When focusing on the helices, for both FI-AD and FI-ND the maximum deviations are for TM-VI at 4.34 Å and 4.39 Å respectively, while for FII-AD it is TM-IV at 5.29 Å and TM-II at 5.39 Å for FII-ND, that exhibit the largest RMSD values.

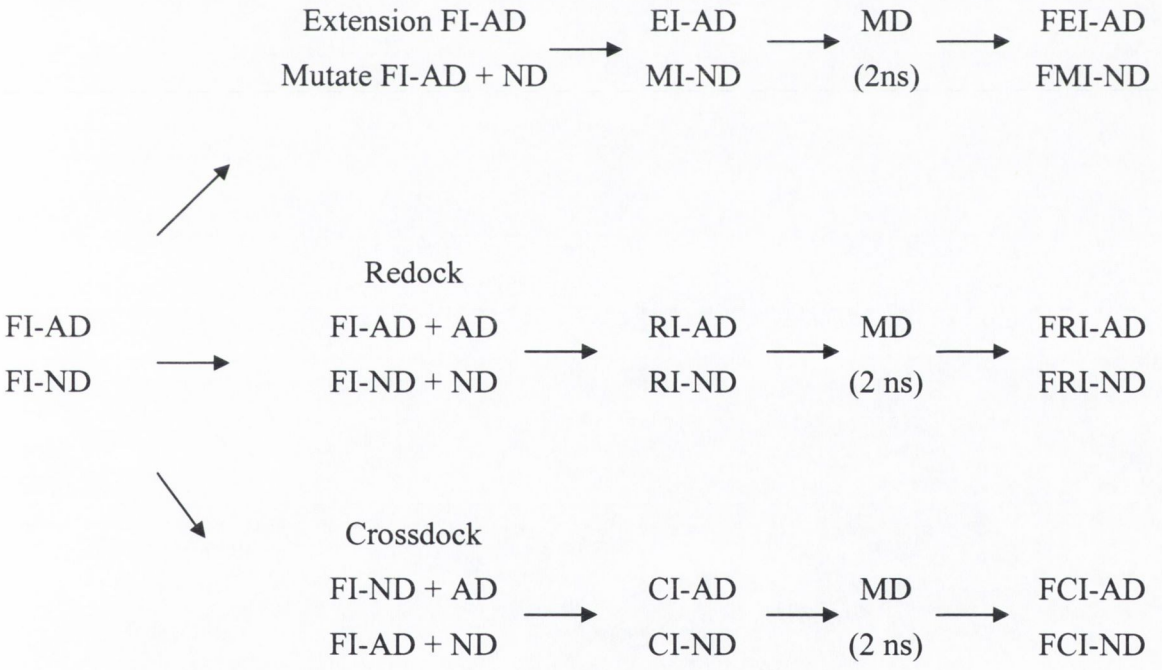
Table 6.2: RMS differences (in Å) for helices of the final agonist complexes against the uncomplexed receptors Final I and Final II.

Helix	FI-AD	FI-ND	FII-AD	FII-ND
TM-I	2.17	3.21	3.83	3.67
TM-II	4.02	3.64	4.79	5.39
TM-III	2.33	2.03	3.68	4.31
TM-IV	2.68	3.65	5.29	4.86
TM-V	2.33	2.03	4.87	4.16
TM-VI	4.34	4.39	4.93	4.77
TM-VII	2.28	2.67	4.84	4.31
TM-VIII	1.13	0.95	4.63	3.38
C α (all)	8.22	8.03	8.56	8.06

In the following sections we consider the α_{1A}-AR models, FI-AD and FI-ND as 'agonist-induced' structures. These receptor structures were chosen as there was a loss of inter-helical interactions and HBs relative to the uncomplexed structure, coupled with a large movement of TM-VI and a lengthening of the donor-acceptor distance for the HB interaction of the DRY motif.

6.3 Utilisation of developed agonist-induced receptors to produce further agonist / α_{1A}-AR complexes

Further MD simulations (2 ns) were performed on a series of additional agonist / α_{1A}-AR complexes (see Scheme 6.2). Firstly, the adrenaline in the FI-AD complex was mutated to noradrenaline and termed MI-ND. The original complex simulation (FI-AD) was also extended and termed EI-AD. The extended and mutated complexes (EI-AD and MI-ND) will illustrate the developed differences between the two agonist complexes commencing from the same starting mode. Next, adrenaline was redocked into the I-AD receptor structure, while noradrenaline was redocked into the I-ND receptor structure. The redocked complex simulations (RI-AD and RI-ND), were performed to further analyse the agonist-induced receptor structures. Finally, adrenaline was crossdocked into the I-ND receptor structure while noradrenaline was crossdocked into the I-AD receptor structure. The cross-docked complex simulations (CI-AD and CI-ND), examined the use of the receptor structures in virtual screening.



Scheme 6.2: Computational pipeline illustrating the flow of 'agonist-bound' receptor structures through extension, mutation, redocking and cross-docking of agonists in Chapter VI. Subsequent MD simulations (2 ns) were performed on each complex in a $H_2O/CHCl_3/H_2O$ membrane mimetic environment.

6.3.1 Molecular Dynamics simulations of the extended adrenaline / α_{1A}-AR complex and the mutated noradrenaline / α_{1A}-AR complex

The comparison of the extended and mutated complex simulations (EI-AD and MI-ND) will illustrate any developed differences between the two agonist complexes commencing from the same starting mode. For simulation EI-AD, the protonated nitrogen interacted in a correlated manner with Trp-102 and Asp-106 (labelled Asp-106a) over the first 1 ns of the simulation (see Figure 6.16). However, as can be observed, the Trp-102 interaction lengthens after 1.2 ns, which corresponds to a lengthening of the Asp-106 interaction. The N-methyl group of adrenaline interacts with residues Asp-106 (labelled Asp-106b) and Cys-176.

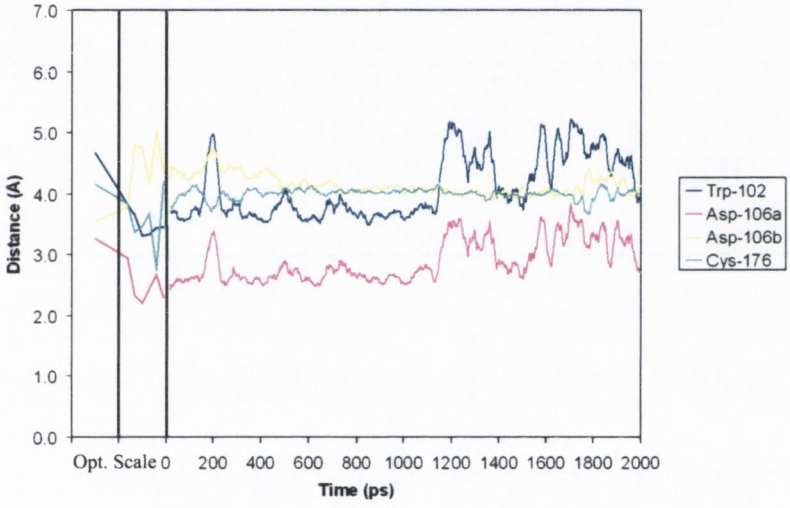


Figure 6.16: Hydrogen Bond analysis in Å for simulation EI-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Trp-102 is d[N...O], for Asp-106 is d[N...O] and for Cys-176 is d[C...O].

Similarly, over simulation MI-ND, the protonated nitrogen interacts in a correlated way with Trp-102, Asp-106 and additionally Cys-176, see Figure 6.17. An interesting movement occurs at 1600 ps when there is a marked decrease in the HB distance for Asp-106, Trp-102 and Cys-176 corresponding to a decrease in the HB distance between the protonated nitrogen and Thr-174. By 1650 ps, there is a stabilisation of the HB interactions with the formation of a HB with Thr-174, which remains for the rest of the simulation.

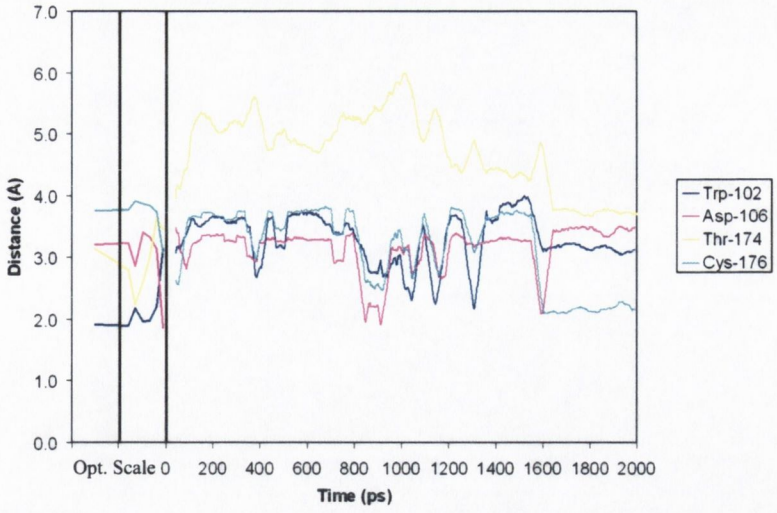


Figure 6.17: Hydrogen Bond Analysis in Å for simulation MI-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Trp-102 is d[N...O], for Asp-106 is d[N...O] for Thr-174 is d[N...O] and for Cys-176 is d[N...O].

In many 'Rhodopsin-like' GPCRs, agonist binding to a cluster of aromatic residues in TM-VI may promote receptor activation by altering the configuration of TM-VI.[259] In both simulations, there are three aromatic residues (Trp-285, Phe-288 and Phe-289) in TM-VI clustered around the binding mode. This aromatic cluster is hypothesised to further induce or stabilize an altered configuration of the side chains within this cluster that could promote receptor activation.[260]

6.3.2 Helical Movement over the extended adrenaline / α_{1A}-AR complex and the mutated noradrenaline / α_{1A}-AR complex

Over the MD production runs (2 ns), the maximum Cα RMSD was 4.09 Å for the adrenaline complex (EI-AD), while the noradrenaline complex (MI-ND) was lower at 3.68 Å. The minimum helical deviation occurred for TM-II in MI-ND at 0.86 Å and the maximum helical deviation occurred for TM-VI of EI-AD at 2.23 Å. This movement is significantly smaller than for the simulations of the initial docked agonist / α_{1A}-AR structures.

The time dependent RMSDs for simulation EI-AD were examined (Figure 6.18). The largest structural change occurs for TM-VI at 800 ps before structurally stabilizing at ~1.2 ns. TM-V gradually decreases in RMSD through the production run until ~1300

ps. There is almost no drift of the helical RMSD after ~1300 ps, by which time adrenaline has perturbed the receptor and optimised its interactions.

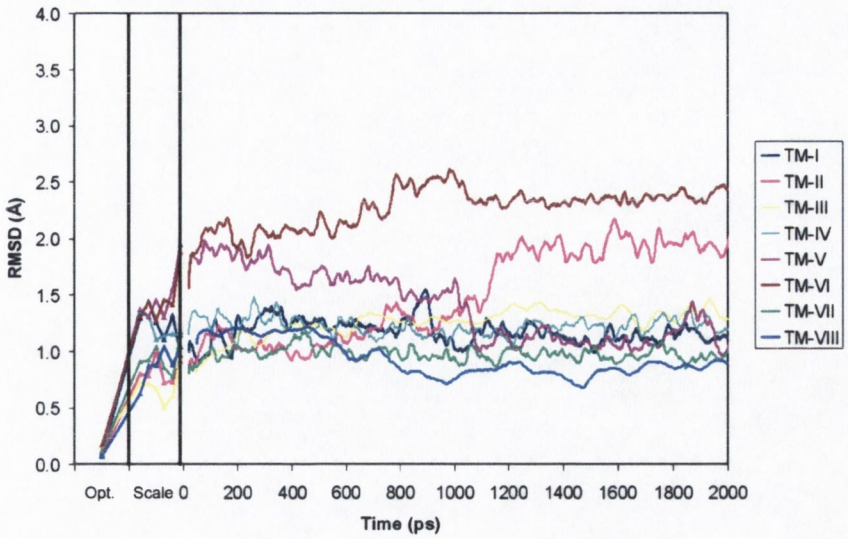


Figure 6.18: RMSDs in Å for simulation EI-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

For simulation MI-ND, TM-IV and TM-VI experience the largest rise in RMSD over the equilibration period but are structurally stable over the production run. Little helical movement is observed until 400 ps of the production run when larger RMSD changes are observed for TM-I and TM-VII than for simulation EI-AD (Figure 6.19). The remaining helices fluctuate about mean RMSDs and the helical RMSD have stabilized by 1600 ps, which corresponds with a stabilisation of the HB interactions at this time.

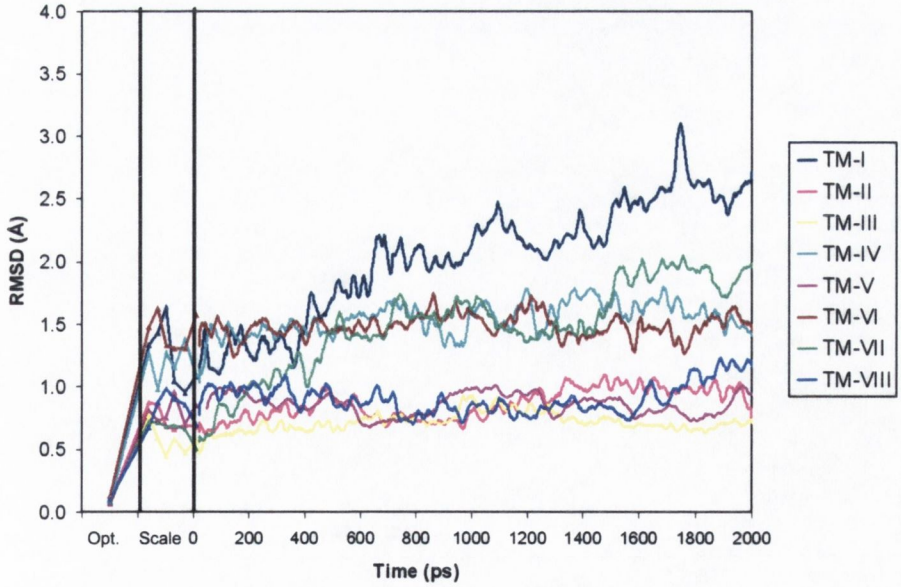


Figure 6.19: RMSDs in Å for simulation MI-ND. Optimisation and scaling steps, followed by moving average trendlines (period of 50 ps) over the production run (2 ns).

As the two agonist complexes commenced from the same initial binding mode any structural differences can be considered to be a consequence of differences induced by the agonists. The main differences in the time dependent RMSD over the two simulations involved TM-I and TM-VII, which deviated more over simulation MI-ND than for simulation EI-AD, while TM-VI and TM-II in simulation EI-AD deviated more than in simulation MI-ND.

6.3.3 Structural analysis after the extended adrenaline / α_{1A}-AR complex and the mutated noradrenaline / α_{1A}-AR complex Simulations

The averaged structures over the last 200 ps of the simulations were optimised and termed, FEI-AD and FMI-ND for the complexes. Two different structural conformations of the receptor emerge, with the overall average Cα RMSD of the final receptors being 3.94 Å and with the largest helical RMS differences occurring for TM-I (1.79 Å), TM-V (1.34 Å) and TM-VII (1.99 Å). As described earlier, the mobility of the EC-II loop, which includes residues Thr-174 and Cys-176, also has a strong effect on the shape of the binding site as it inserts more into the helical bundle over the simulation.

The overall analysis of the agonist extended and mutation studies indicates that the two agonists (adrenaline and noradrenaline) result in different final binding modes. For FEI-AD the final binding mode interactions involved the protonated nitrogen with Trp-102 (3.42 Å) and the methyl linker with Cys-176 (3.99 Å). For FMI-ND a larger number of principal interactions involved the protonated nitrogen with Trp-102 (3.31 Å), Asp-106 (3.37 Å), Thr-174 (3.63 Å) and Cys-176 (3.71 Å), while the *beta* hydroxyl interacted with Thr-174 (3.80 Å) and Tyr-316 (3.61 Å). The superimposed final binding modes of the mutated complexes indicate that although they started from the same binding mode a different conformation was adopted for noradrenaline with the *beta* hydroxyl group orientated towards the TM-III (Figure 6.20).

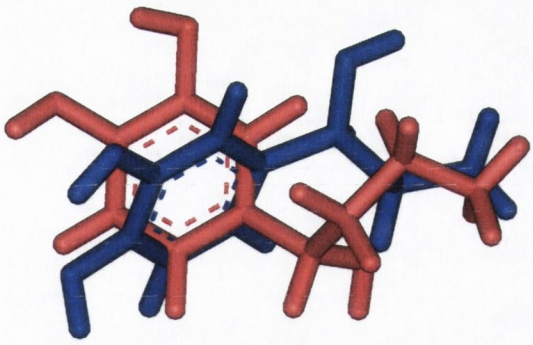


Figure 6.20: Superpositioning of the final binding modes, for FEI-AD (in red) and FMI-ND (in blue).

6.3.4 Molecular Dynamics simulations of redocked agonist/α_{1A}-AR Complexes

The redocked complex simulations (RI-AD and RI-ND) were performed to further analyse the agonist-induced receptor structures. For simulation RI-AD, the hydrogens of the protonated nitrogen formed interactions with Asp-106, and Cys-176 of EC-II, which were the strongest interactions over the course of the simulation (2 ns), Figure 6.21. The *meta* hydroxyl interacted initially with Ser-188, but switched to interact with Ser-192, after ~850 ps. The N-methyl group formed occasional correlated interactions with Trp-102 and Thr-174.

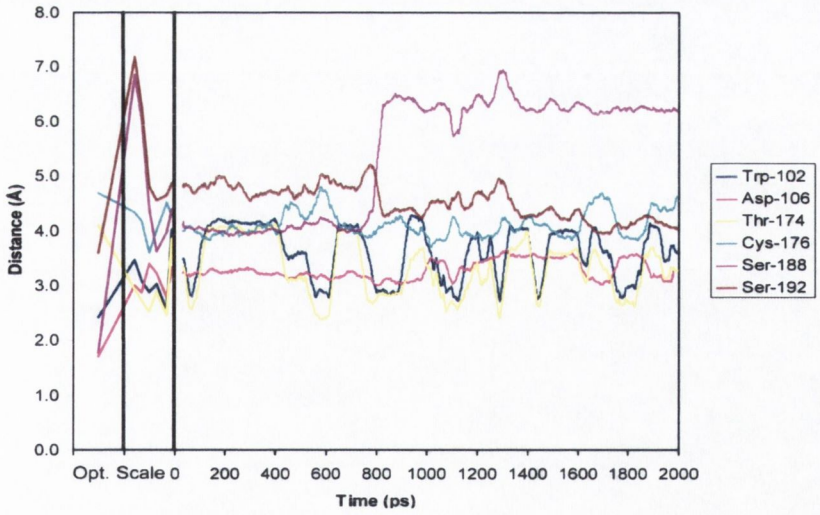


Figure 6.21: Hydrogen Bond distances in Å (donor--acceptor) for simulation RI-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Trp-102 is d[C...O], for Asp-106 is d[N...O], for Thr-174 is d[C...O], for Cys-176 is d[N...O], for Ser-188 is d[O...O], for Ser-192 is d[O...O].

Noradrenaline was redocked into the FI-ND receptor structure and termed (RI-ND). The HB interactions optimise over the equilibration period and interactions largely occur between the protonated nitrogen and Asp-106, and with Gln-177, throughout the simulation, see Figure 6.22. Occasional interactions occurred between Ser-188, and the *beta* hydroxyl.

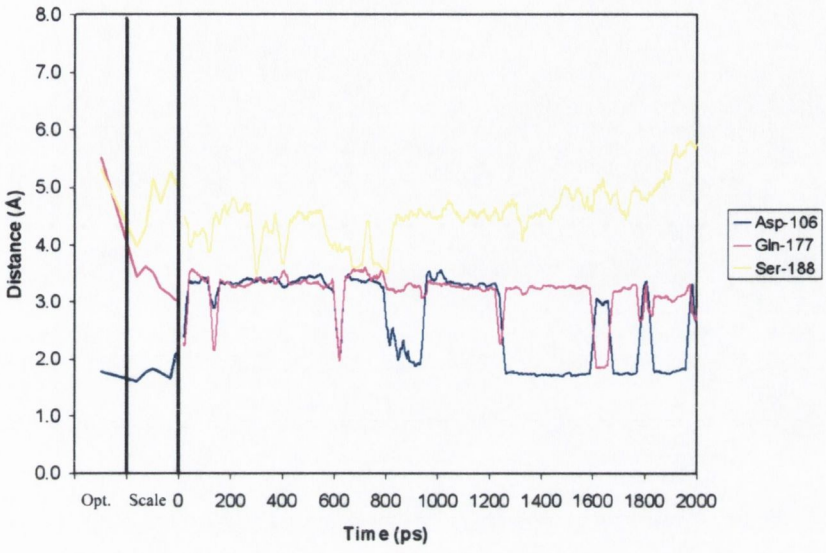


Figure 6.22: Hydrogen Bond distances in Å (donor--acceptor) for simulation RI-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Asp-106 is d[N...O], for Gln-177 is d[N...O] and for Ser-188 is d[O...O].

6.3.5 Helical Movement over the redocked agonist / α_{1A}-AR Complexes

For simulation RI-AD, the RMSD analysis indicated an initial rise in RMSD for the helices over the equilibration period followed by stable RMSDs over the production run. Two exceptions included a gradual structural drift for TM-I over the simulation (see Figure 6.23) and two slight rises in RMSD at 0.9 ns and 1.7 ns for TM-VI.

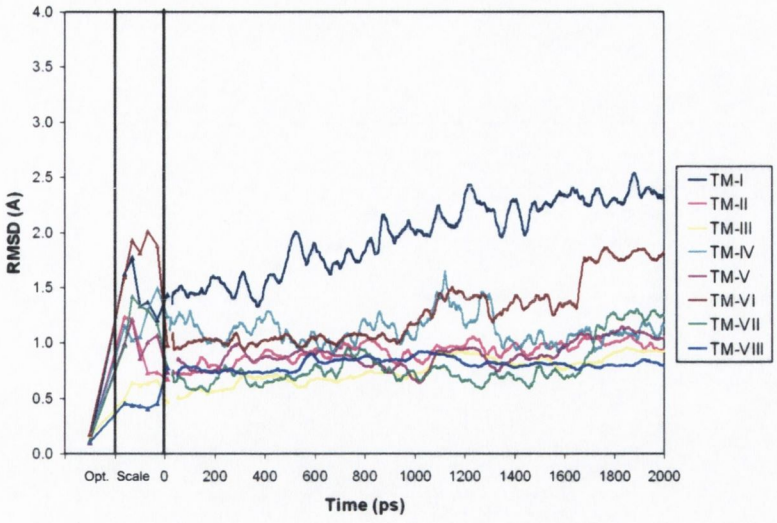


Figure 6.23: RMSDs in Å for simulation RI-AD. Optimisation and scaling steps, followed by moving average trendlines (period of 50 ps) over the production run (2 ns).

In general, for simulation RI-ND, the helices undergo little structural change and have structurally stabilised by the end of the production run (Figure 6.24). For this simulation, TM-VII exhibited the largest RMSD, which rose sharply over the course of the equilibration period but was largely stabilised over the production run, while TM-I underwent a dip in RMSD between 200-1200 ps before rising in RMSD again.

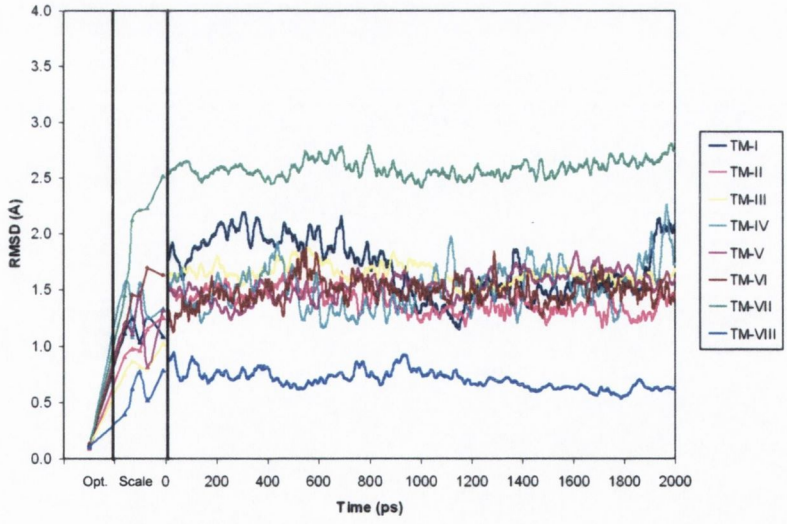


Figure 6.24: RMSDs in Å for simulation RI-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

6.3.6 Final Redocked agonist / α_{1A}-AR Complex Binding Modes

The final binding mode of simulation RI-AD, has both the *beta* hydroxyl (1.61 Å) and the protonated nitrogen (1.78 Å) orientated towards the carboxylate group of Asp-106, while the N-methyl group is close to Thr-174 (2.62 Å). A movement of TM-V over the simulation facilitates the formation of two Ser interactions in the final bound mode where the *meta* hydroxyl is close to Ser-188 (1.80 Å) and the *para* hydroxyl is 1.94 Å from Ser-192 (see Figure 6.25).

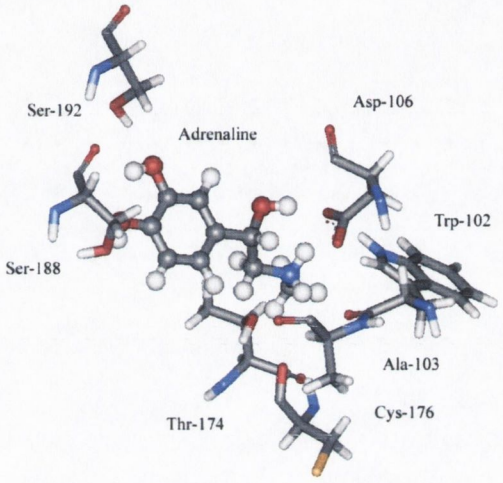


Figure 6.25: Final binding mode for RI-AD after 2ns MD simulations.

A marked difference is observed in the conformation of the agonist after simulation RI-ND as the final binding site is quite closed, with a HB formed between Asp-106 and Ser-188 (1.6 Å). In the final binding mode, the protonated nitrogen was involved in two interactions again with Asp-106 (3.59 Å) and an EC-II residue, in this instance Gln-177 (3.23 Å), see Figure 6.26. However, unlike the RI-AD simulation, the catechol hydroxyl groups are orientated away from the Ser residues of TM-V.

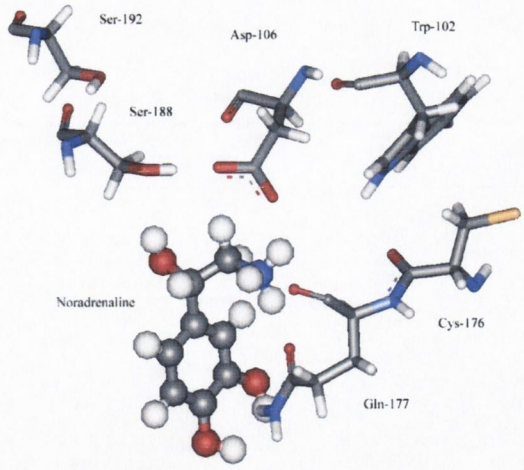


Figure 6.26: Final binding mode for RI-ND after 2ns MD simulations.

6.3.7 Molecular Dynamics Simulations of the crossdocked agonist/α_{1A}-AR Complexes

The cross-docked complex simulations examined the use of the receptor structures in virtual screening. Adrenaline was crossdocked into the FI-ND receptor structure and termed CI-AD. Over the course of the production run, there where steady interaction distances involving the protonated nitrogen with Asp-106 and occasionally Thr-174, Figure 6.27. Further interactions developed with residues in EC-II, including Cys-176, with the N-methyl group and temporarily Thr-174 with the *beta* hydroxyl.

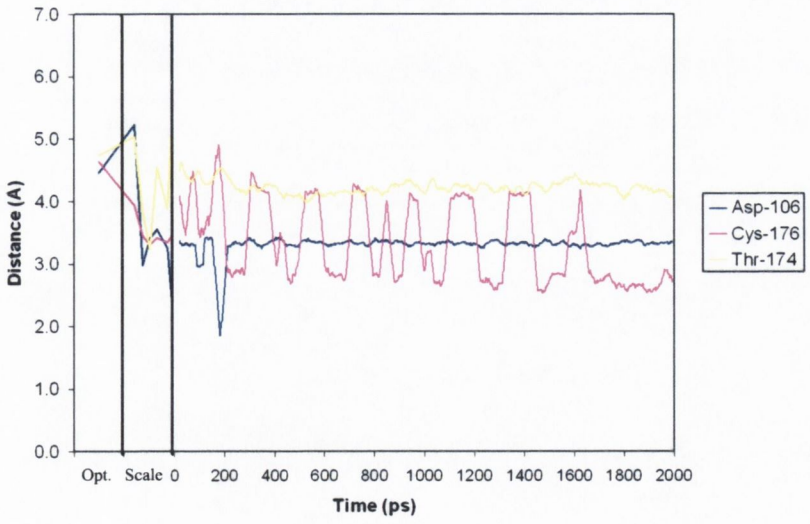


Figure 6.27: Hydrogen Bond distances in Å for simulation CI-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Asp-106 is d[N…O], for Cys-176 is d[C…O] and for Thr-174 is d[O…N].

Noradrenaline was cross-docked into the FI-AD receptor structure and termed CI-ND. For complex CI-ND, over the 2 ns production run, the protonated nitrogen interacted steadily and in a correlated motion with Asp-106 (labelled Asp-106a) and Trp-102. After 400 ps, the protonated nitrogen also interacted occasionally in a correlated manner with Cys-176 and Ala-103, Figure 6.28. Steady interaction distances were observed between the methyl linker and Asp-106 (labelled Asp-106b) and the nearby *beta* hydroxyl with Thr-174.

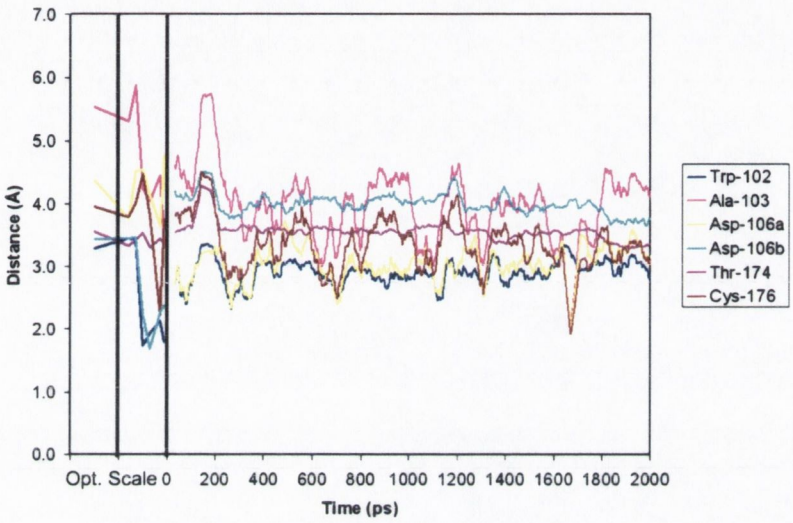


Figure 6.28: Hydrogen Bond distances in Å for simulation CI-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (2 ns). The indicated distance for Trp-102 is d[N...O], for Ala-103 is d[N...O], for Asp-106 is d[N...O], for Thr-174 is d[O...O] and for Cys-176 is d[N...O].

6.3.8 Helical Movement over the crossdocked agonist/α_{1A}-AR Complexes

Over simulation CI-AD, all the structural movement is in the equilibration period, although the movement is quite small (see Figure 6.29). There was little effect exerted by the crossdocked agonist on the receptor, other than on TM-I, which has the largest RMSD at 2.8 Å. Similarly, a large RMSD of TM-I was observed over simulation I-ND.

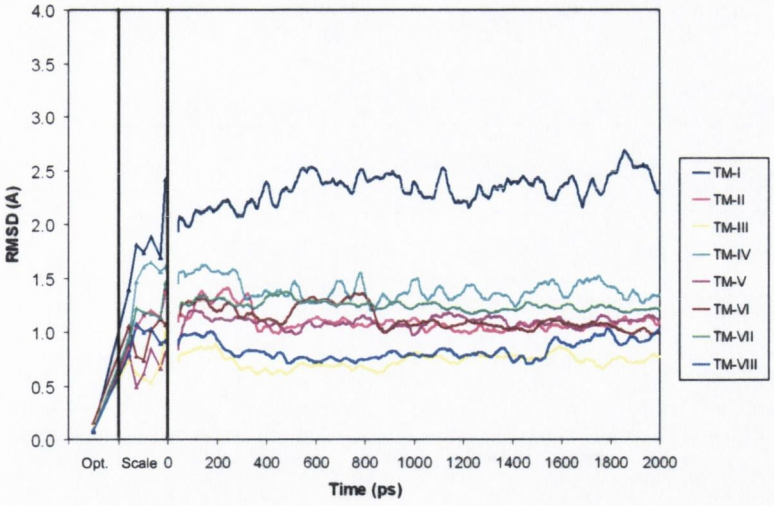


Figure 6.29: RMSDs in Å for simulation CI-AD. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

No major structural changes occur over the time-dependent RMSD data for simulation CI-ND (see Figure 6.30). The helical RMSDs experience only thermal fluctuations although TM-VIII has an interesting structural profile for this simulation with a marked decrease in the region of 900-1100 ps.

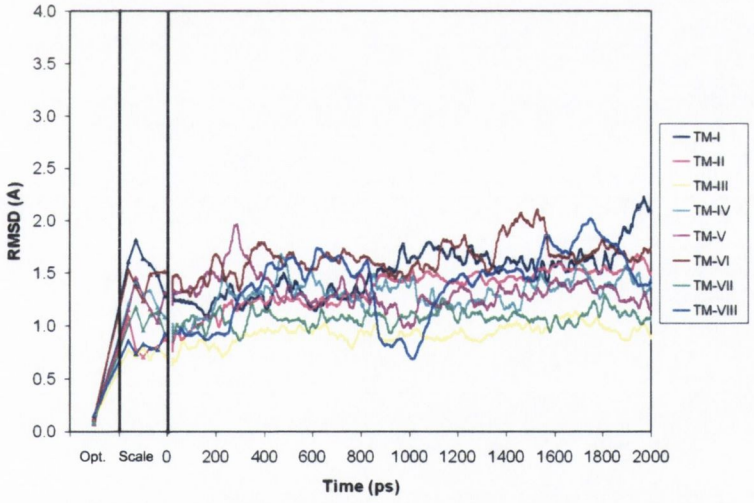


Figure 6.30: RMSDs in Å for simulation CI-ND. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run.

6.3.9 Final Cross-docked agonist / α_{1A}-AR Complex Binding Modes

In the final binding mode of complex CI-AD, an interaction occurred between the protonated nitrogen and Asp-106 at a distance of 3.30 Å. The orientation of the agonist in the binding site is indicated in Figure 6.31.

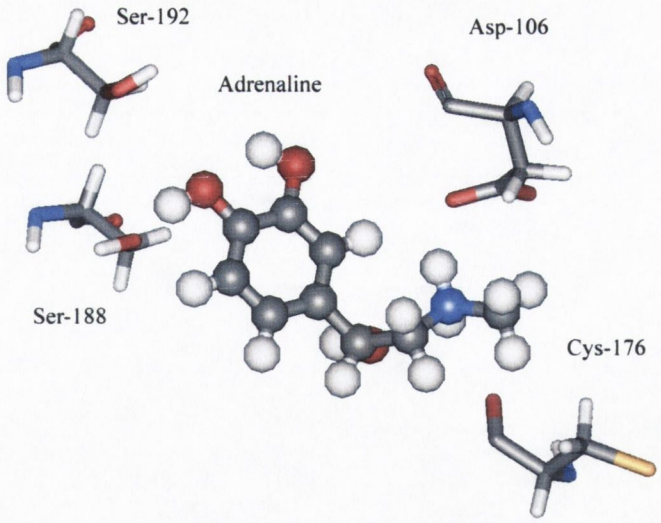


Figure 6.31: Final crossdocked binding mode of CI-AD with Asp-106, Cys-176, Ser-188 and Ser-192 after MD simulation (2 ns).

For complex, CI-ND, the main interaction formed was between the protonated nitrogen and Asp-106 at 1.80 Å and the *beta* hydroxyl with Asp-106 at 1.9 Å (Figure 6.32). It appears that noradrenaline has twisted in the active site so that the *para* hydroxyl is orientated away from the Ser residues in the final binding mode. No other close hydroxyl interactions were determined see Figure 6.32.

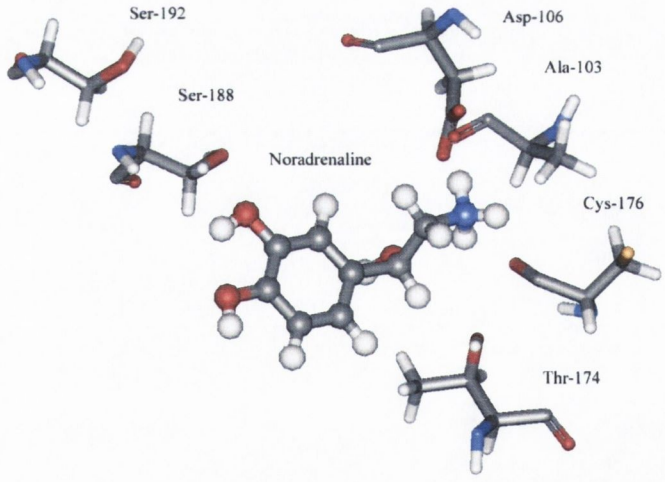


Figure 6.32: Final crossdocked binding mode of CI-ND with Asp-106, Thr-175, Gln-177, Ser-188 and Ser-192 after MD simulation (2 ns).

6.4 Comparison of Final agonist/α_{1A}-AR Binding Modes

The three final bound conformations of adrenaline were compared, namely FEI-AD, FRI-AD and FCI-AD, see Figure 6.33. The extended and crossdocked conformations of adrenaline were conformationally the closest, as can be observed in Figure 6.33 there is a great deal of overlay between these two conformations. The marked difference lies in the positioning of the *beta* hydroxyl group of the redocked adrenaline which is orientated for an interaction with TM-III coupled with the positioning of the *meta* and *para* hydroxyls for interactions with Ser residues in TM-V. As a consequence the RMSD from the redocked to the extended structure is 1.62 Å and to the crossdocked structure is 1.74 Å.

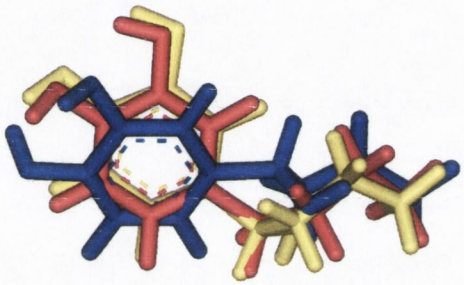


Figure 6.33: Superpositioning of the FEI-AD (red), FRI-AD (blue) and FCI-AD (yellow) binding modes.

The three final bound conformations of noradrenaline, namely FMI-ND, FRI-ND and FCI-ND, were also compared in Figure 6.34. The RMSDs of the three conformations revealed that the mutated and crossdocked conformations of noradrenaline were also the closest at 0.68 Å. However, the binding site of the redocked noradrenaline was relatively closed as a HB formed between Asp-106 and Ser-188 (1.6 Å). As a consequence the RMSD from the redocked to the mutated structure is 1.74 Å, and to the crossdocked structure is 1.76 Å.

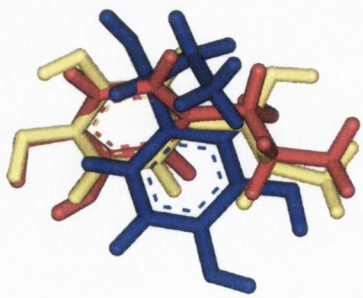


Figure 6.34: Superpositioning of the FMI-ND (red), FRI-ND (blue) and FCI-ND (yellow) binding modes.

From our analysis the most representative agonist complexes that have good binding interactions, are complexes RI-AD and CI-ND. The HB analysis suggested a larger conformational fluctuation for noradrenaline (CI-ND) than adrenaline (RI-AD) over the course of the simulations. There is also a notable difference in the conformation of the chain in the final binding modes of adrenaline and noradrenaline (Figure 6.35). It is possible that adrenaline is doubly anchored by the additional interactions of the N-methyl group.

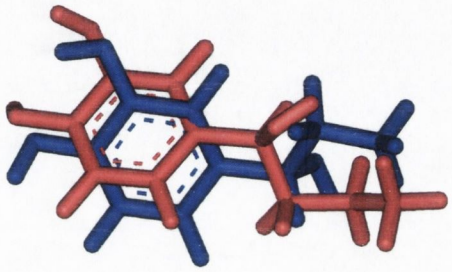


Figure 6.35: Superpositioning of the FRI-AD (red) and FCI-ND (blue) binding modes.

Both FRI-AD and FCI-ND, 'agonist-induced' structures proved to be stereochemically satisfactory after 2 ns, with 316 residues (99.3 %) in allowed regions with a further two loop residues (0.7 %) in disallowed regions for FRI-AD; while for FCI-ND, four residues were in disallowed regions (three of which were loop residues) as shown in Figure 6.36.

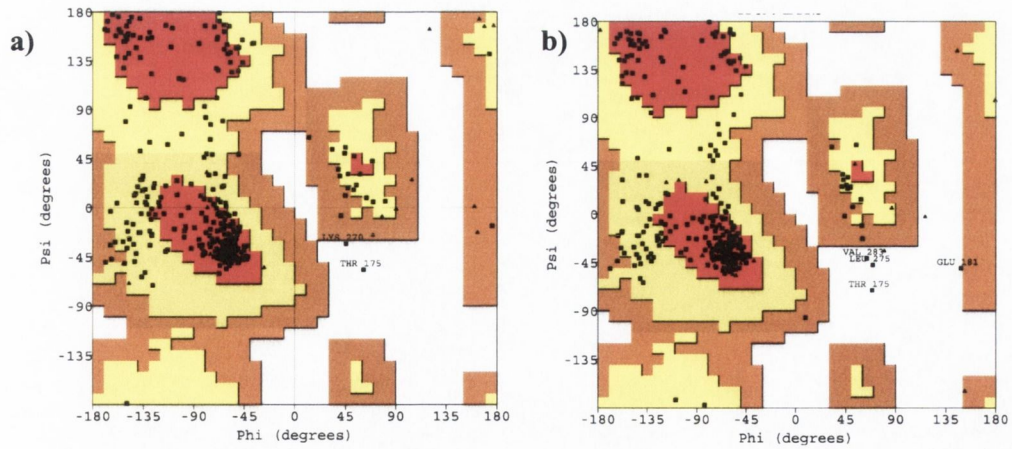


Figure 6.36: Stereochemical analysis for the final complexed structures; (a) FRI-AD receptor and (b) FCI-ND receptor.

6.5 Comparison between the Uncomplexed and Agonist Induced Receptor Conformations

So far, comparisons have been made with respect to our homology model with which we obtained the initial binding modes through docking. This has allowed us to examine the degree of rearrangement during the dynamics and hence determine when structural change has occurred. However, determination of the differences between the inactive and active forms of the receptor requires comparisons with an inactive form which has been equilibrated under the same conditions, see Chapter IV. In this study, two components of receptor activation deduced from prior experimental and computational studies of GPCRs were investigated for the α_{1A}-AR, namely, (1) conformational changes in the 'agonist-bound' receptor form compared to the inactive form and (2) changes in the interactions of the conserved DRY and NPxxY motifs.

6.5.1 Conformational changes in the 'agonist-bound' receptor form

A comparative structural analysis was performed between the uncomplexed receptor form produced in Chapter IV (Final I) and the two representative 'agonist-bound' receptor forms. The two overall Cα RMSD differences, between the uncomplexed and complexed forms, are quite similar at 5.68 Å for FRI-AD and 5.25 Å for FCI-ND (Table 6.3). In the helical regions, TM-I has the highest RMSD difference for the FRI-AD structure to the uncomplexed at 4.42 Å, while for FCI-ND it is TM-VII at 3.99 Å.

Table 6.3: Comparison of uncomplexed (Final I) and agonist complexed receptor structures via RMSD (in Å).

Helix	FRI-AD	FCI-ND
TM-I	4.42	3.28
TM-II	2.79	2.93
TM-III	2.64	2.58
TM-IV	2.55	2.76
TM-V	2.19	2.55
TM-VI	4.09	3.70
TM-VII	3.97	3.99
TM-VIII	0.86	0.89
C α (all)	5.68	5.25

In the uncomplexed Final I, the active site is closed with interactions forming between Ser-192 and residues Asp-106 and Ser-188 (see Figure 6.3). Hence, the binding site is closed in the inactive receptor form and can not accommodate an agonist, highlighting the need for our protocol to develop an 'agonist-induced' receptor structure. The close interactions in the uncomplexed binding site between TM-III and TM-V are lost for both the 'agonist-bound' receptors as the active site opens to accommodate the ligand. Hence, with the adrenaline complex (FRI-AD), Ser-188 interacts with Tyr-184 (3.02 Å) and Val-185 (2.65 Å) of TM-V and Ser-192 interacts with Ala-189 (3.74 Å) of TM-V but not with a TM-III residue. Similarly, for the noradrenaline complex (FCI-ND), Ser-188 interacts with Tyr-184 (3.21 Å) and Val-185 (3.05 Å) of TM-V while Ser-192 does not form any interactions.

Agonist binding results in a movement of TM-V away from TM-III on the intracellular side, with a subsequent repositioning of the Ser residues of TM-V, see Figure 6.37, the residues of the uncomplexed Final I are indicated in grey, while the residues of the complexed receptor are in colour.[261,262] This movement was monitored by measuring the distances between the alpha carbons of internal facing residues on each turn of these helices. The intracellular opening between TM-III and TM-V is most marked for the helical turn on the intracellular side of Pro-196 in TM-V. For the uncomplexed receptor the reference separation distance between the Cα of Ile-199 in TM-V and of Leu-117 in TM-III is 6.96 Å, while in the complexes it is notably larger at 11.42 Å for the AD complex and 10.55 Å for the ND complex. Such an intracellular separation may be important for receptor activation as a change in the intracellular position of TM-V may cause a change in the IC-III loop connecting TM-V to TM-VI. It is generally understood that IC-III plays an important role in receptor activation[263,264] and a change in IC-III may possibly facilitate G-protein interaction.

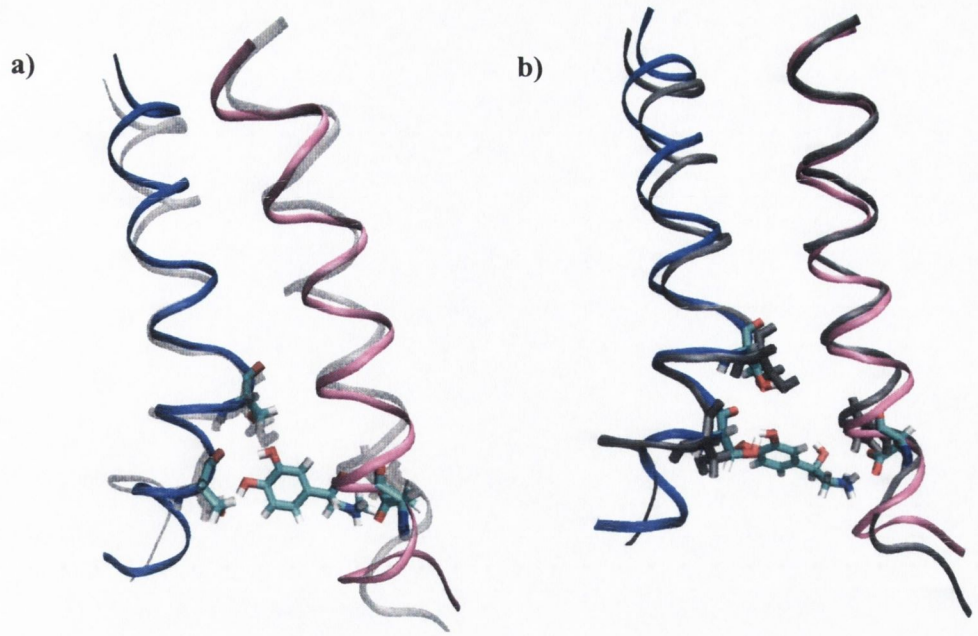


Figure 6.37: The original uncomplexed structure (Final I) is in grey, while after 2 ns TM-III is in mauve and TM-V is in blue for the FRI-AD complex (a) and the FCI-ND complex (b). TM-III and TM-V were superimposed using the fitting protocol of Swissprot.

In rhodopsin and in the inactive α_{1A}-AR, TM-VI exists in a kinked configuration, with its cytoplasmic end near to the cytoplasmic end of TM-III.[37] In our complexed receptor, flexibility about a Pro residue (Pro-287) in TM-VI results in a movement of the cytoplasmic end of TM-VI away from TM-III although different motions are observed for the two agonist complexes. We observe a marked lengthening of the distance separating Leu-117 in TM-III from Val-278 in TM-VI from 9.48 Å in the uncomplexed receptor to 12.84 Å for the ND complex, while the AD complex remains at 10.02 Å. Conversely, for the distance separating Ile-120 in TM-III from Gly-275 in TM-VI, the uncomplexed distance is 13.24 Å while the AD complex distance is longer at 15.19 Å and the ND complex distance is similar at 13.45 Å. Furthermore, the extracellular separation distances have shortened as illustrated by the distance between the Cα of Asp-106 in TM-III and of Phe-288 in TM-VI, at 15.85 Å for the uncomplexed receptor, to 13.62 Å for the AD complex and 14.55 Å for the ND complex. Hence, our α_{1A}-AR results are consistent with a movement of TM-VI away from TM-III, which was identified for Rhodopsin[265,266] and in the β_2-AR indicating similar activation steps.[267,268,269]

6.5.2 Role of the conserved motifs DRY and NPxxY in Receptor Activation

As a change in the interaction of conserved motifs such as the DRY or NPxxY motifs may facilitate activation of the receptor such interactions were monitored over the uncomplexed and complexed simulations. In the uncomplexed receptor structure (Final I), Asp-123 of the DRY motif donates to both neighbouring Arg-124 and Tyr-125. However, for the complexed receptors, novel interactions are formed with Asp-123 donating to Ile-119. Arg-124, a residue entirely conserved in all 'Rhodopsin-like' GPCRs, also interacts with residues in TM-I, TM-III and IC-III for the uncomplexed structure and for the adrenaline induced structure while in the noradrenaline induced structure, interactions between Arg-124 and TM-I and IC-III are lost, with interactions only occurring with other TM-III residues. Noradrenaline is hence hypothesised to promote conformational changes releasing Arg-124 of the DRY motif so that it can reach a position suitable for G protein binding.[270]

In the uncomplexed receptor, Asn-322 of the NPxxY motif[271,272] forms various interactions with Asp-72 of TM-II, Ser-113 of TM-III and Ser-319 and Cys-320 of TM-VII. Again in the complexed simulations novel interactions are formed. For the AD complex, Asn-322 only interacts with its neighbouring Ile-321, while for the ND complex this residue interacts with Asp-72 of TM-II, forms a novel interaction with Val-277 of TM-VI and with Asn-318 of TM-VII.

Furthermore, in our analysis, a Ser-113/Asn-322 interaction is formed in the uncomplexed receptor and broken in the complexed forms, possibly freeing Asn-322 for other interactions such as with other TM-VII residues. This is consistent with the postulate of Bruysters *et al.*[273] that a Ser residue of TM-III acted as a molecular switch in the activation of the Histamine H_1 receptor. Such a movement may change the orientation of residues in IC-II and IC-III that allows for increased affinity of coupling to the G-protein.

6.6 Discussion / Conclusions

The only available structural template, bovine Rhodopsin, has been crystallized in its inactive state that is conformationally different from an activated state. As a result, it is questionable if the inactive Rhodopsin crystal structure is appropriate to use in modelling agonist/GPCR complexes.[274] We have developed a protocol to produce

agonist induced receptor conformations and have examined the four initial complexes by docking and through MD refinement. MD simulations were conducted to induce different conformational states, and over 1 ns the main structural change involved a closing of EC-II down over the helical barrel, perhaps by recognising the exposed side of the agonist.

A further series of MD simulations were performed utilising the agonist induced receptor structures to investigate the conformation and binding modes of the complexes. The extended and mutation study (EI-AD and MI-ND) indicated that the two agonists, although structurally similar, adopted different binding modes over the course of the simulations. The redocked and crossdocked simulations model the reorganisation of the native receptor structure induced by an agonist. The analysis of the interactions of the catecholamine complexes confirms the mutagenesis experiments and suggests new residues involved in the agonist-binding site as the active form of the receptor. In the model complexes the most critical interactions for the binding of the agonists exist between the ligands and residues in TM-III, TM-V, TM-VI and TM-VII, which is consistent with many earlier reports. The molecular modelling results indicate that the contact residues for the protonated amine is Asp-106 in TM-III, for the catechol hydroxy groups are Ser residues in TM-V, while for the β-OH it is Thr-174 for noradrenaline or Asp-106 in TM-III for adrenaline. The contact site for the extra methyl group of adrenaline is less clear but may involve TM-III or EC-II.

A comparison of the final bound conformations revealed marked differences in the positioning of the *beta* hydroxyl group of the redocked adrenaline when compared to the extended and crossdocked conformations. The *beta* hydroxyl in FRI-AD is orientated towards TM-III, were it can form an interaction with Asp-106. For the noradrenaline conformations the RMSD analysis revealed that the mutated and crossdocked conformations of noradrenaline were also quite close in conformation. The marked difference lies in the positioning of the redocked noradrenaline, which had a relatively closed binding site as a HB formed between Asp-106 and Ser-188 (1.6 Å).

A number of steps may be envisaged when moving from agonist binding to G protein activation, including the movement of TM-V away from TM-III. The observed changes in the intracellular end of TM-V and TM-VI are consistent with conformational changes in activation that have been detected by computational[275,276,277] and

experimental studies.[96,278,279,280,281] Upon examining the conserved residues in GPCRs, a change to the environment of Asp-123 of the DRY motif may cause Arg-124 to move out of the TM helical bundle and change the orientation of residues in IC-II and IC-III that allows for increased affinity of coupling to the G-protein. However, 3 ns simulations performed in this study is a small fraction of the ~ μs time believed for receptor activation, although the initial time period would be expected to be quite significant. Interpretation of motions is also complicated by the use of homology models, as imperfectly placed side chains may not have enough time to relax to their desired positions. Nevertheless, this problem would cancel out to some extent between the various (complexed and uncomplexed) simulations and some information may still be extracted about how the motions differ in TM regions and depend on the bound ligand. Finally, the recent realization that GPCRs form homo-oligomeric and hetero-oligomeric complexes has added a new dimension to rational drug design. However, this important aspect of GPCR biology remains largely unincorporated into schemes to search for new therapeutics.[282]

Despite these limitations, we have gained insights into the binding mode of agonists to the α_{1A}–AR and the subsequent conformational changes to the receptor structure in the production of an 'agonist-bound' receptor form. Receptor binding and activation tests, combined with a 3D model of the receptor, allow us to study the receptor in great detail, which improves our understanding of conformational changes that take place upon receptor activation. Our models agree with the mutagenesis data available to date, however, we await the production of an experimental crystal structure of an adrenoceptor in the inactive or active state for a thorough comparison.

Chapter VII: Antagonist / α_{1A}-AR Complex Study

"The outcome of any serious research can only be to make two questions grow where only one grew before."

Thorstein Veblen

7.0 Introduction

Rational development of novel antagonists requires an understanding of the interactions involved in antagonist binding to the α_{1A}-AR and the structural change induced. Antagonist binding is effected by Asp-106 mutations and so a number of antagonist docking studies have been performed in the ligand binding site (see Chapter VI).[99,191] Mutagenesis studies have also identified that three consecutive residues (Gln-177, Ile-178 and Asn-179) of EC-II, are involved in α_{1A}-AR subtype selectivity of the antagonist WB4101 (**14**).[100] A similar importance was placed on EC-II residues for the serotonin 5-HT_{1D} receptor[283] and the δ-opioid receptor.[284] Perez *et al.*,[285] also postulated that Phe-308 and Phe-312 of TM-VII were important sites for α_1-AR antagonism. The principal residues involved in ligand binding are indicated in Figure 7.1.

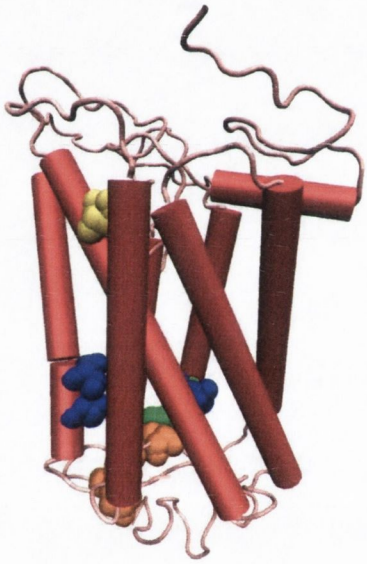


Figure 7.1: Residues of α_{1A}-AR which may be important for antagonist binding include the residues Asp-106, Ser-188 and Ser-192 in blue, the EC-II residues 177-179 in orange and Phe-308 and Phe-312 in green. The location of Asp-123 of the DRY motif is also indicated in yellow.

Bissantz *et al.*,[252] performed docking simulations on minimised Rhodopsin-derived GPCR structures, including the dopamine D_3, muscarinic M_1 and vasopressin V_{1a} receptors. Alternatively, Ishiguro *et al.*[286] manually expanded the binding site of α_1-ARs to facilitate binding of antagonists. Recently, Evers *et al.*,[287] developed an α_{1A}-AR model utilising a docked antagonist as an additional restraint in the modelling procedure. Evers *et al.*,[287] subsequently performed virtual screening of a library of ligands and concluded that rhodopsin based homology models may be used as the

structural basis for GPCR lead finding and compound optimization. However, we believe that the inactive crystal structure template is not directly suitable for docking purposes, as the modelled binding sites are often too narrow to accommodate larger antagonists. While, the study of Evers *et al.*[287] addressed this point, it did not address the issue of conformational change, which is the basis of receptor activation or inhibition.

Through our molecular modelling studies we aim to facilitate the integration of the available experimental observations and biophysical data into a scheme to examine receptor structure and functional blockade. In this work, we examined the use of homology models developed from the bovine Rhodopsin crystal structure in direct antagonist binding. We have developed a computational strategy to produce ligand induced receptor structures, which were further examined through redocking, crossdocking, virtual screening and subsequent MD simulations. We examined the binding modes for a variety of antagonist structural classes, both selective and non-selective for the α_{1A}-AR. Our computational approach involved the following four steps:

1) Docking studies of antagonists with the α_{1A}-AR homology models, which facilitate an examination of the initial interactions. The antagonist complexes were optimised through short MD simulations (200 ps).

2) Extended MD simulations (1 ns) of antagonists/α_{1A}-AR complexes, involving, the ligands doxazosin, tamsulosin and compound **6**, were performed in a $H_2O/CHCl_3/H_2O$ membrane mimic to simulate induced structural changes to the α_{1A}-AR upon antagonist binding.

3) A series of 23 antagonists were then docked with the 'antagonist-bound' receptor forms developed in step 2.

4) Further MD simulations (500 ps) were performed on the antagonist/receptor complexes to determine the interacting residues and the various binding modes adopted by the antagonists in the 'antagonist-induced' receptor forms.

7.1 Analysis of Antagonists in the α_{1A}-AR models

In this section, the two α_{1A}-AR homology models (developed in Chapter IV) were used for docking of all antagonists (reviewed in Chapter V). The binding site was taken as all residues within 10 Å of Asp-106, Ser-188 and Ser-192 and the EC-II residues 177-179 as suggested by site directed mutagenesis studies. The antagonist / α_{1A}-AR complexes were then refined through MD simulations (200 ps).

7.1.1 Interactions of the Clinical Antagonists with the α_{1A}-AR homology models

For the quinazoline series of antagonists, prazosin, doxazosin, terazosin and alfuzosin, the MM atom types were assigned from GAFF.[288] The quinazoline moiety requires two additional angular parameters, *c3-os-ca* and *nh-ca-na*[ii] (with corresponding force constants of 62.7 and 71.9 kcal mol^{-1}, respectively), which were obtained by analogy to similar GAFF parameters. The two angular force constants were fitted to best reproduce the DFT (B3LYP/6-31G*) determined vibrational data with little resultant change to the force constants at 63.9 and 74.3 kcal mol^{-1}. As a further validation of the parameters used, MD simulations (200 ps) were performed on the four antagonists in water boxes from which stable structures emerged.

When the quinazoline antagonists were complexed with Model I, no interactions occurred with the postulated binding site residues over the course of the MD simulations (200 ps). For the Model II complexes, Asp-106 interactions were maintained with doxazosin, alfuzosin and terazosin, while interactions with Ser-188 occurred only for alfuzosin and terazosin over the MD simulation (200 ps). Upon docking tamsulosin with both receptor models no initial interactions were formed with Asp-106. Hence, neither receptor model appears suitable for direct docking of this series of ligands.

7.1.2 Interactions of Antagonists 1-8 with the α_{1A}-AR homology models

In the refinement process of the MM parameters for the bis-imidazolidinium and bis-guanidinium class of compounds (**1-8**), we first considered the simplified imidazole

[ii] GAFF parameter assignment; the atom type *c3* – represents an sp3 Carbon; *os* - an ether or ester oxygen; *ca* - an sp2 aromatic carbon; *nh* - an amine N connected one or more aromatic rings and *na* - an sp2 N with three connected atoms.

and guanidine structures. The C-N (*c2-n2*)[iii] bonded parameter was refined by fitting to DFT optimised structures (to 420.6 kcal mol^{-1}, 1.38 Å). The force constants of the angular terms were also fitted to best reproduce the theoretical frequencies obtained from DFT calculations at the B3LYP/6-31G* level scaled by 0.9614.[179] This included *c2-n2-hn* (at 48.2 kcal mol^{-1}, 120.0 °) and *n2-c2-n2* (at 46.1 kcal mol^{-1}, 120.0°). In addition to structural and energetic properties, it is important to test the force fields ability to account for dynamical properties. This was accomplished by comparing the MM determined vibrational frequencies for imidazole (1634, 1691 and 1736 cm^{-1}) with experiment (1621, 1681 and 1767 cm^{-1}).[289]

For the compounds of interest (**1-8**), the MM parameters were further validated by comparing the experimental[290] and theoretical vibrational frequencies of the bis-2-imino-imidazolidinium compound **3** and the bis-guanidinium compound **7** (Table 7.1).[291,292] A small RMSD of 20.3 cm^{-1} for compound **3** and of 32.7 cm^{-1} for compound **7** were observed between the available experimental and MM determined frequencies with a RMSD. The similarity between the MM and QM derived IR frequencies was also quite high (RMSD was 20.5 cm^{-1} for compound **3** and 39.9 cm^{-1} for compound **7**). After the MD simulations (150 ps) of antagonists **1-8** in a water box all compounds maintained an appropriate average planar structure.

Table 7.1: Comparison of experimental, DFT and MM, IR frequencies for compounds **3** and **7**in cm^{-1}.

Compound	Assignment	Expt.	QM	MM
3	C-H aromatic	3100	3109	3099
3	C=N	1630	1642	1632
3	NH^{2+}	1567	1569	1538
3	(H_2)C-N(secondary)	1280	1284	1267
3	O=S=O	1135	1165	1148
3	C-H aromatic substituted in 1,4 positions	1083	1091	1119
7	C-H aromatic	3080	3085	3076
7	C=N	1655	1662	1669
7	NH^{2+}	1550	1543	1520
7	O=S=O	1130	1105	1121
7	C-H aromatic substituted in 1,4 positions	1085	1092	1093

[iii] GAFF parameter assignment; the atom type *c2* – represents an sp2 Carbon; n2 – an aliphatic sp2 N with two connected atoms and hn – a H bonded to a nitrogen atom.

Over the course of the 200 ps MD simulations of the bis-imino-imidazolidinium compounds (**1-4**) complexed with either Model I or Model II, no long-term interactions occurred with Asp-106, while some temporary interactions occurred with residues of EC-II. For the bis-guanidinium compounds (**5-8**), a large range of interactions occurred with different residues. A lasting Asp-106 interaction was maintained for compounds **6** and **7** and occasionally for compound **8**. Strong interactions were developed amongst the residues of EC-II (residues 177-180) over the four guanidinium / Model II complexes. There were also a number of interactions with TM-V, involving Ser-188 for compound **7** and Ser-192 for compounds **5**, **6** and **8**.

In summary, the guanidinium class of antagonists (**5-8**) interacted better than the imidazolidinium class (**1-4**), particularly with Model II. This analysis corresponds with the experimental activities, which were in general higher for the guanidinium compounds. However, these receptor structures still do not appear appropriate for direct screening of all antagonists. Hence, we aimed to develop antagonist-induced receptor structures to be utilised in subsequent structure based drug design (SBDD) studies.

7.2 Extended Molecular Dynamics simulations for α_{1A}-AR complexes

The Model II – doxazosin complex (II-Dox), Model II – tamsulosin complex (II-Tam) and Model II–**6** complex (II-**6**) were chosen for further MD simulations (1 ns) performed in a $H_2O/CHCl_3/H_2O$ cell. Model II was chosen for the receptor as better interactions were observed with the ligands in its binding site. For the ligands, doxazosin is an example of a non-selective α_1-AR antagonist, while tamsulosin is a selective α_{1A}-AR antagonist. Finally, compound **6** was chosen as this antagonist exhibited the best activity in pharmacological tests on our novel compounds.[61]

7.2.1 Hydrogen Bond Analysis of the three antagonist / α_{1A}-AR complexes

The initial docking of doxazosin with the α_{1A}-AR (complex II-Dox) resulted in the 4-amino group and the 6-methoxy group interact with Asp-106 (2.88 Å / 2.50 Å), while the 7-methoxy group is the closest to Ser-188 (4.28 Å). The EC-II residues are in close proximity to the binding site allowing for interactions with Gln-177, similar to that recently observed by Pedretti *et al.* for their α_{1A}-AR model[191]

Over the course of the MD simulation (1 ns), a long lasting interaction was maintained between the protonated nitrogen and Asp-106, while there were sporadic interactions of the neighbouring methoxy group also with Asp-106. No other HBs formed over the course of the simulation.

For the initial Model II – tamsulosin complex (II-Tam), the protonated phenethyl amine moiety is nearest to Ser-188, while the sulphonamide is close to residues 177-179 of EC-II. Monitoring the percentage of HB occurrence over the MD simulation (1 ns) it is evident that dynamics are essential to allow for reorientation of the ligand and optimisation of interactions. Indeed, interactions with the two carboxylate oxygens of Asp-106 formed with the sulphonamide moiety of tamsulosin over the equilibration period (see Asp-106a and Asp-106b in Figure 7.2). Regarding the EC-II residues, an interaction between the sulphonamide and Ile-178 occasionally occurs. Furthermore, there are interactions involving Glu-180 with the protonated nitrogen (Glu-180a) and the adjacent methyl linker (Glu-180b) over the simulation.

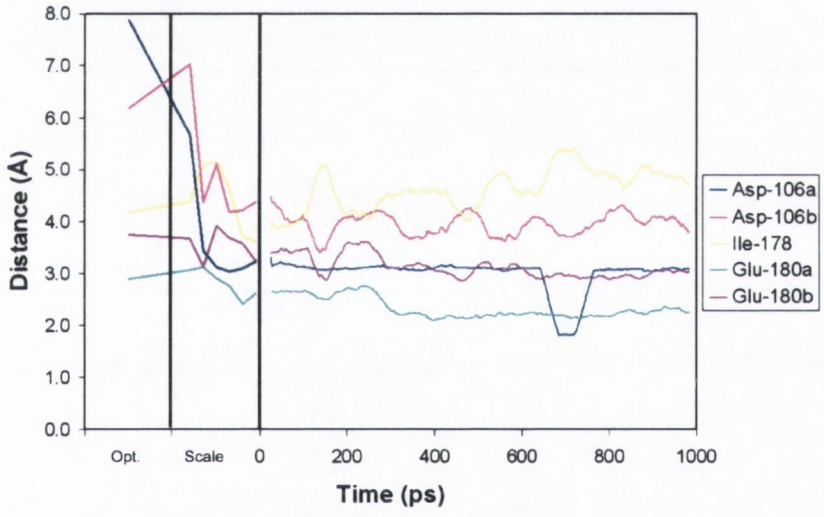


Figure 7.2: Hydrogen Bond analysis in Å for simulation II-Tam. Optimisation and scaling steps, followed by moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Asp-106 is d[N...O], for Ile-178 is d[N...O] and for Glu-180 is d[N...O].

For the docked complex II-**6**, a number of interacting nitrogens are present (see Figure 7.3). One guanidinium moiety forms a HB with Asp-106 and Ile-178. The other guanidinium moiety is close to Ser-192 of TM-V. Over the MD simulation (1 ns), a steady interaction remains with Asp-106, while the EC-II residues, Ile-178 and Glu-180 interact with an outer guanidinium nitrogen. The Glu-180 interaction involves both

oxygens (Glu-180a until 500 ps and with Glu-180b throughout). At the other guanidinium moiety there are steady interactions with Ser-192 and Trp-285.

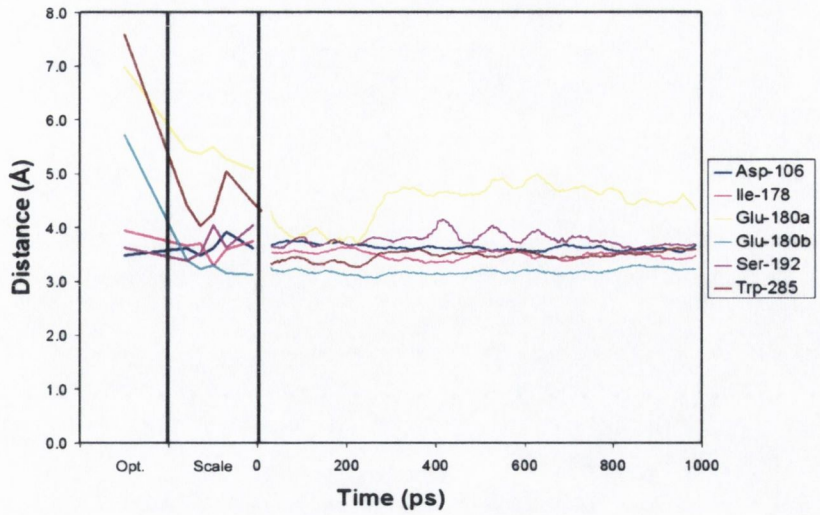


Figure 7.3: Hydrogen Bond analysis in Å for simulation II-**6**. Optimisation and scaling steps, followed by the moving average trendlines (period of 50 ps) over the production run (1 ns). The indicated distance for Asp-106 is d[N...O], for Ile-178 is d[N...O], for Glu-180 is d[N...O], for Ser-192 is d[N...O] and for Trp-285 is d[N...O].

In summary, the most notable changes in the HB occurrence for the three antagonist/α_{1A}-AR complexes were the development of an interaction between tamsulosin and Asp-106; and between compound **6** and Glu-180 and Trp-285 over the equilibration phase. Over the three production simulations, strong Asp-106 interactions were maintained with compound **6** and tamsulosin while it was less frequent for doxazosin. Simulations II-Tam and II-**6** exhibited strong interactions with Ile-178 and Glu-180, while only simulation II-**6** forms interactions with both Ser-192 and Trp-285.

7.2.2 Structural analysis over the three antagonist/α_{1A}-AR complex simulations

The effects of the interaction of the three antagonists, doxazosin, tamsulosin and compound **6**, on the conformation of α_{1A}-AR, were compared by monitoring the time dependent helical RMSDs over the MD simulations (see Figure 7.4). Again structural changes were observed over the equilibration periods, which stabilised over the production runs. Analysing the time dependent RMSD data for simulation II-Dox, TM-IV exhibited a rise in RMSD between 150 – 650 ps before undergoing a decrease in RMSD. The remaining helices exhibit structural stability throughout the production run.

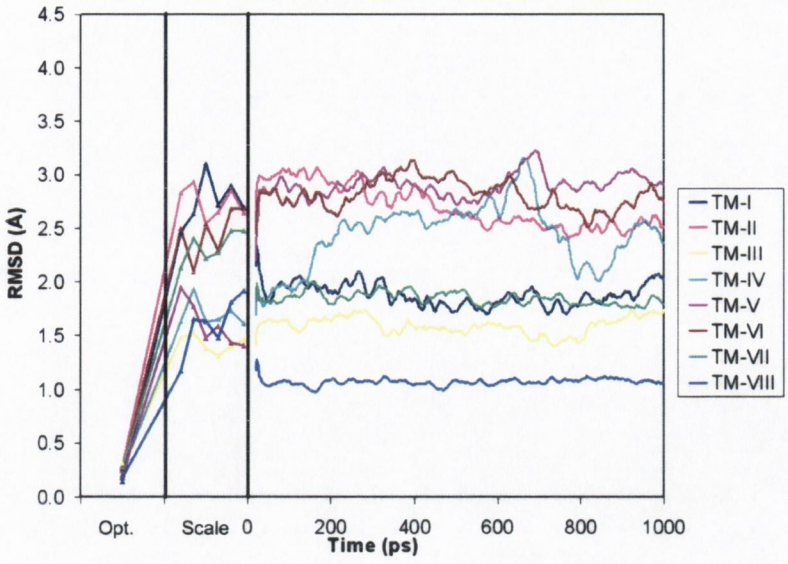


Figure 7.4: RMSDs in Å for simulation II-Dox. Optimisation and scaling steps, followed by moving average trendlines (period of 50 ps) over the production run (1 ns).

For simulation II-Tam, TM-II and TM-IV undergo a rise in RMSD over the equilibration period, which is followed by a gradual rise in RMSD, indicating a variety of structural rearrangements over the course of the simulation as it is perturbed by the antagonist (see Figure 7.5). Finally, considering the helical regions of simulation II-**6**, for TM-I, there is a gradual decrease in RMSD before structural stabilisation at ~700 ps. A small rise in RMSD is observed for TM-II and TM-VII at ~300 ps before structurally stabilising, while TM-VI experienced a rise in RMSD until ~800 ps. Finally, for TM-IV, a series of dips and rises are observed.

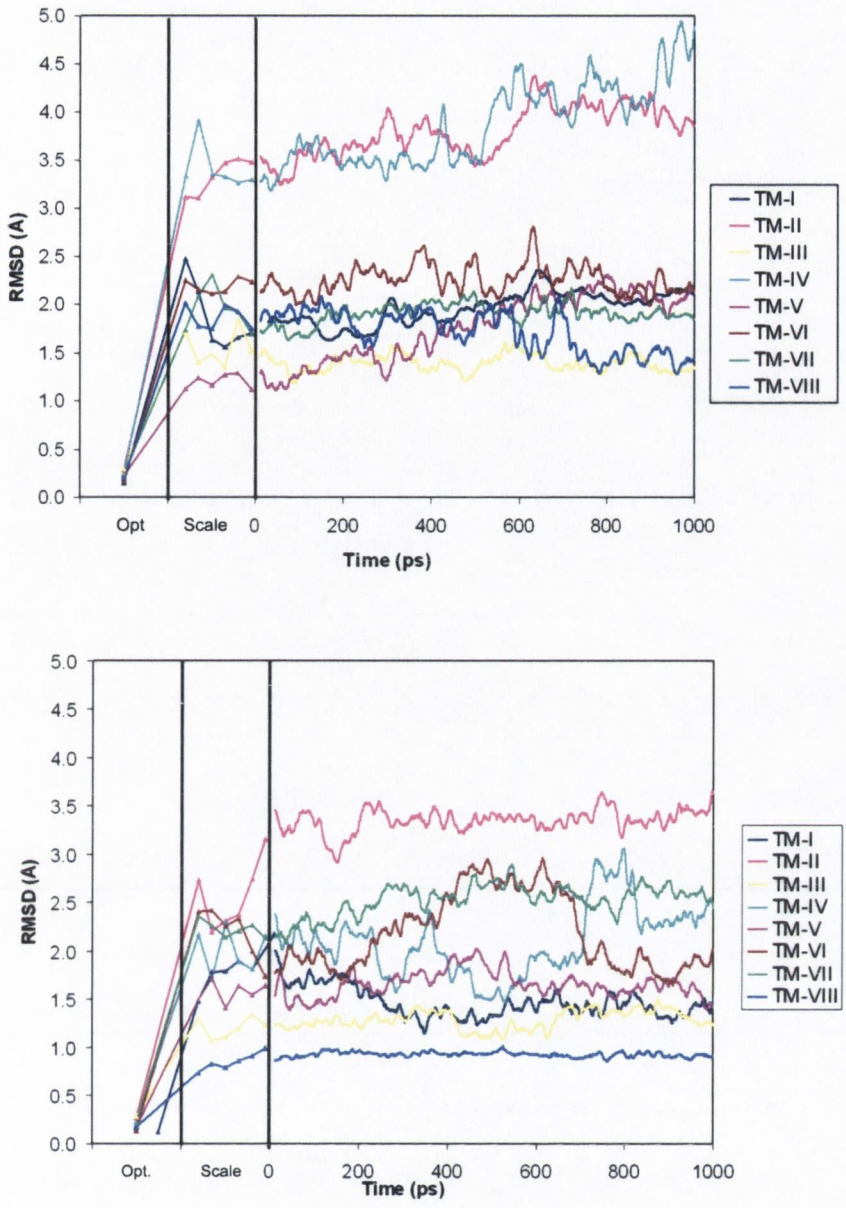


Figure 7.5: RMSDs in Å for simulations II-Tam and II-**6**. Optimisation and scaling steps, followed by moving average trendlines (period of 50 ps) over the production run (1 ns).

In summary, marked structural changes occurred for simulation II-Dox, with a non-selective antagonist and simulations II-Tam and II-**6** with selective antagonists over the equilibration period with smaller movements over the production runs. For all three simulations the largest structural change over the production run occurs for TM-IV. The complex structures were averaged over the last 200 ps of the MD simulations, optimised and termed FII-Dox, FII-Tam and FII-**6**.

7.2.3 Binding modes after antagonist/α_{1A}-AR Molecular Dynamics Simulations

An analysis was performed on the final binding modes of the complexes for which different interactions were observed. For the final binding mode of FII-Dox, the only interaction present between the protonated nitrogen with Asp-106 (3.14 Å), Figure 7.6.

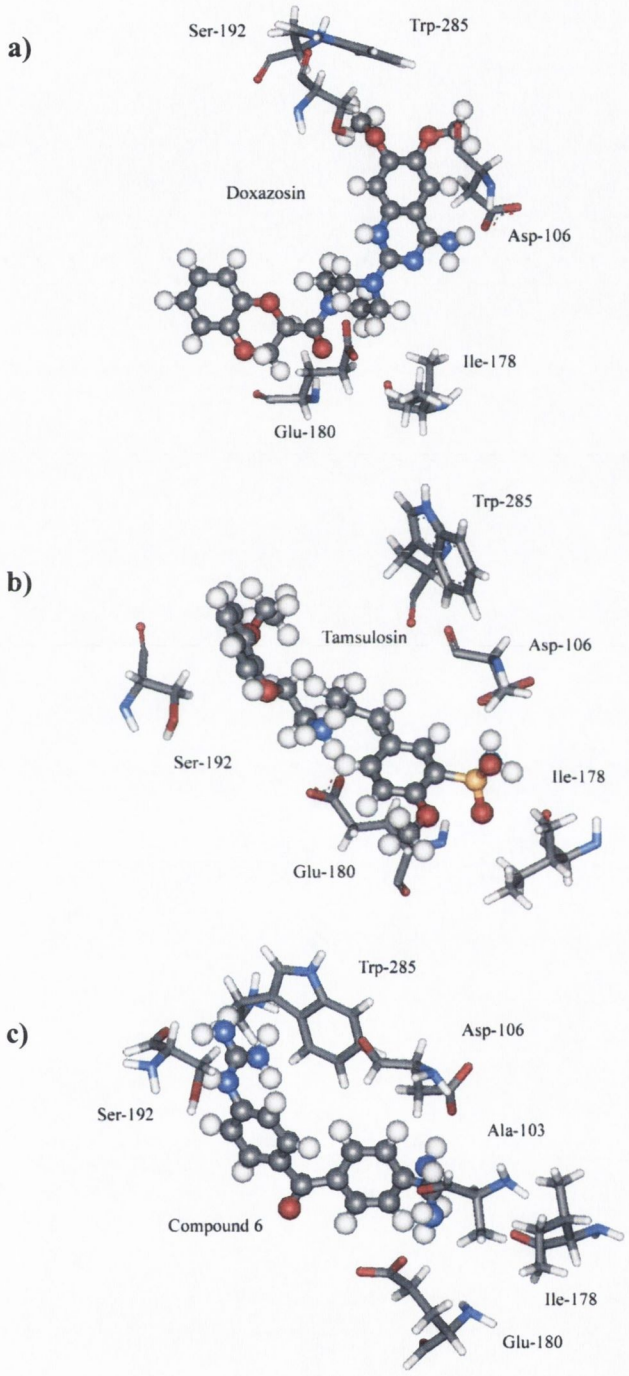


Figure 7.6: Final binding modes of (a) the FII-Dox complex, (b) the FII-Tam complex and (c) the FII-**6** complex.

The final optimised complex FII-Tam, again exhibits interactions with Asp-106 (3.45 Å) and Glu-180 (3.23 Å) coupled with a reorientation of the aromatic group of tamsulosin. The final binding mode of FII-**6** contains four HBs - with Asp-106 (3.26 Å); Ser-192 (3.45 Å); Glu-180 (3.23 Å); and Ala-103 (3.76 Å).

The final receptor structures with the ligands removed were termed II[Dox], II[Tam] and II[**6**]. The stereochemical analysis indicates that all three receptors are stereochemically acceptable, receptor II[Dox] having 312 residues (99.7 %) in allowed regions and only one tail residue in disallowed regions, while II[Tam] has three loop residues and II[**6**] has four loop residues in disallowed regions (see Figure 7.7).

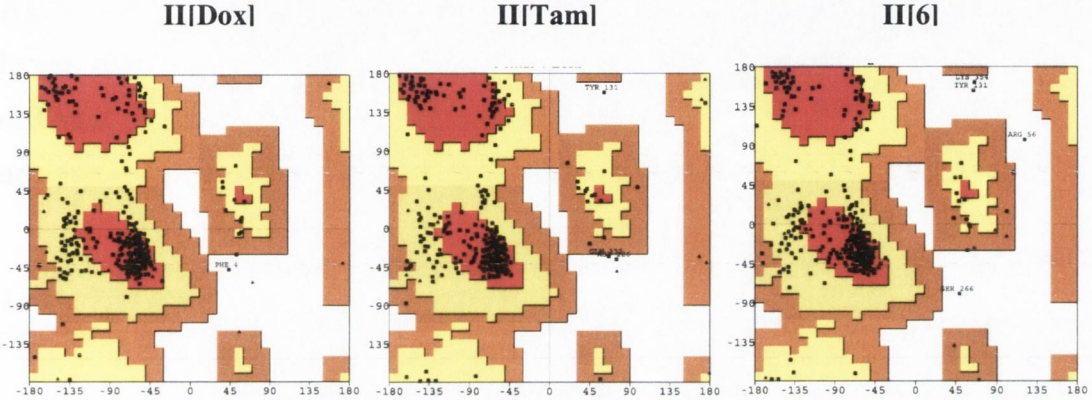


Figure 7.7: Stereochemical analysis of final antagonist complexed structures after 1ns of simulation, II[Dox], II[Tam] and II[**6**] complexes.

In summary, three 'antagonist-bound' receptor structures were produced, which after initial structural changes exhibited largely stable helical RMSDs over the last 200 ps of simulation. Furthermore, the final receptor structures were stereochemically satisfactory and interactions were formed between the ligands and residues in TM-III, EC-II and TM-V.

7.3 Comparison of the various α_{1A}-AR Models

By comparing the antagonist-induced receptor structures with the uncomplexed receptor (Final II from Chapter IV), we gain insights into the structural displacements of receptor domains induced upon antagonist binding (Section 7.3.1). A comparison was also made between the effect of functionally different ligands on the receptor structures, *i.e.* between an agonist-induced receptor conformation (FI-AD) and the three antagonist-induced receptor conformations produced in this chapter (section 7.3.2).

7.3.1 Structural Comparison of Uncomplexed and Antagonist Bound α_{1A}-AR

The structural movement from the uncomplexed to the antagonist induced receptor forms was monitored through a RMSD analysis. A large overall RMSD was observed for the Cα atoms in the range of 5.93 Å to 6.60 Å, which indicates structural differences between the two states (see Table 7.2). The helical regions deviate to a maximum of 3.84 Å in TM-VI for receptor II[**6**], indicating a structural difference between the uncomplexed and antagonist complexed receptor forms in the TM regions. For receptor structures II[Dox] and II[**6**], TM-VI exhibits the largest RMSD changes while for receptor structure II[Tam] it was TM-V at 2.88 Å. The RMSDs for TM-I (3.30 Å), TM-II (3.29 Å) and TM-IV (3.14 Å) for receptor structure II[**6**] are also notably larger than for the other two receptor models. Hence, the three antagonists exert different effects on the various helical regions of the receptor structure relative to the uncomplexed receptor form.

Table 7.2: RMSD differences in Å, for helices of final 'antagonist-bound' receptor structures from the uncomplexed receptor Final II.

Helix	II[Dox]	II[Tam]	II[6]
TM-I	1.42	2.55	3.30
TM-II	2.67	2.74	3.29
TM-III	2.07	1.75	1.69
TM-IV	2.21	1.78	3.14
TM-V	2.65	2.88	2.09
TM-VI	3.63	2.41	3.84
TM-VII	2.06	2.51	2.66
TM-VIII	1.01	1.86	1.78
C α (all)	6.60	5.93	6.44

Focusing on the helices involved in binding interactions, namely TM-III and TM-V, we observe noticeable differences between the uncomplexed receptor structure and the different final complexed structures (see Figure 7.8).

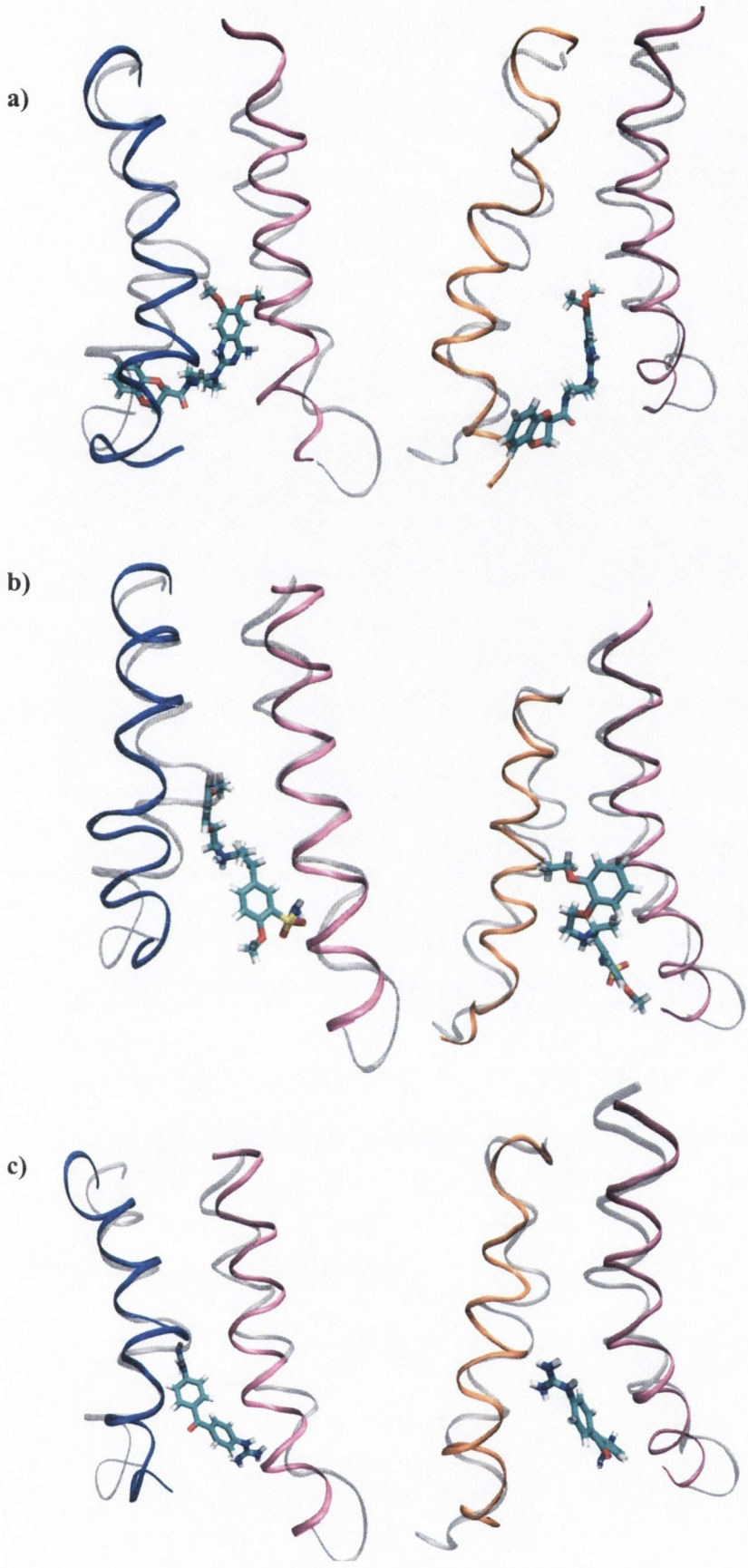


Figure 7.8: TM-III (mauve) and TM-V (blue) on left; TM-III (mauve) and TM-VI (orange) on right, for complex FII-Dox (a), for complex FII-Tam (b) and for complex FII-**6** (c). The corresponding uncomplexed helical structures are indicated in grey.

For the FII-Dox structure, the methoxy groups of the quinazoline portion of doxazosin are orientated towards the middle of TM-III and the Pro kink region of TM-V. The benzopyrane moiety of doxazosin is also orientated towards both the extracellular side of TM-V and TM-VI. For the FII-Tam and FII-**6** complexes, the antagonists contact with the extracellular end of TM-III. For the tamsulosin complex, there is a different ligand orientation in the binding site with no interactions occurring with the extracellular end of TM-V, which undergoes a helical straightening. Again, for the tamsulosin and compound **6** optimized complexes, additional space is created due to movement of the Pro kink region TM-V, allowing for an additional interaction.

The variations in stabilizing features over the course of the simulations were also monitored. The HB analysis indicated a large total number of 238 HBs for receptor structure II[Dox], of 231 for structure II[Tam] and of 227 for structure II[**6**], which is a decrease from the 245 HBs present in Final II. In terms of interhelical HBs there are 24 for structure II[Dox], 17 for structure II[Tam], while structure II[**6**] is equivalent to Final II with 15 interhelical HBs.

The Asp-123 residue of the DRY motif donates to nearby Ile-119 for the uncomplexed and three final antagonist structures, while an additional interaction has formed with Ile-120 in all three final antagonist structures. The fully conserved Arg-124, interacts with Ile-120, Ser-121 for the uncomplexed and antagonist induced structures, and additionally with Asp-123 and Glu-267 for all three antagonist structures and also with Ser-59 of IC-I for structure II[Dox]. Finally, Tyr-125 of DRY has a variety of interactions which vary over the three final antagonist complex structures and interacts with Tyr-130 and Ser-121 for structure II[Dox]; to Ser-121, Ile-122 and Val-202 for structure II[Tam] and to none for structure II[**6**], while in the uncomplexed receptor Tyr-125 only has an interaction with Ile-122.

Regarding the conserved Asn-322 residue of TM-VII, interactions are formed with Cys-280 and Asn-318 for the uncomplexed receptor structure, which has changed to interactions with Asp-72, Ser-113 and Asn-318 for receptor structure II[Dox]; Asn-67, Asn-318 and Ser-319 for receptor structure II[Tam]; and finally with Asn-67, Ser-113 and Ser-319 for receptor structure II[**6**]. The interactions of the DRY and NPxxY motifs have changed during the conformational change induced by the antagonist.

7.3.2 Structural Analysis between 'Agonist-bound' and 'Antagonist-bound' Receptor forms

In order to further investigate the structural component of activation of the human α_{1A}-AR, an analysis of the differences between the 'antagonist-bound' and the 'agonist-bound' receptor structures was performed. Taking the adrenaline-bound receptor structure (FI-AD) as a reference structure, the RMSD for the II[Dox], II[Tam] and II[**6**] receptor structures from it were determined (see Table 7.3). Across the receptor structures different helices exhibited the largest RMSD, with TM-II (3.32 Å) for II[Dox], TM-VI (3.09 Å) for II[Tam] and TM-IV (3.18 Å) for II[**6**]. The differences in the RMSDs for the various antagonists again indicate that the ligands induce different changes to the receptor conformation.

Table 7.3: RMSD differences in Å, for helices of selected antagonist complexes over 1ns MD from the agonist-bound structure FI-AD.

	II[Dox]	**II[Tam]**	**II[6]**
TM-I	2.56	2.83	3.08
TM-II	3.32	2.56	2.85
TM-III	1.66	1.61	1.30
TM-IV	3.37	2.47	3.18
TM-V	2.97	2.73	2.88
TM-VI	2.29	3.09	3.14
TM-VII	2.01	1.69	2.64
TM-VIII	0.89	1.87	1.84
C α (all)	6.46	5.41	5.65

There are a number of differences in the interactions of the DRY and NPxxY motifs as a consequence of agonist and antagonist binding. For the DRY motif in the agonist-complexed receptor, interactions were formed with Asp-123 donating to Ile-119 while all three antagonists additionally interacted with Ile-120. For Arg-124 in the agonist-complexed receptor, a series of donations to residues occur in TM-I, TM-III and IC-III. For all three antagonist complexes, Arg-124, loses interactions with TM-I and donates to Ile-120, Ser-121, Asp-123 and Glu-267 and additionally with Ser-59 for structure II[Dox].

Regarding the NPxxY motif for the AD complex, Asn-322 interacts solely with its neighbouring Ile-321. For the antagonist complexes, differing interactions from the agonist-bound receptor are formed. Interactions are formed with Asp-72, Ser-113 and Asn-318 for structure II[Dox]; Asn-67, Asn-318 and Ser-319 for receptor II[Tam]; and finally with Asn-67, Ser-113 and Ser-319 for receptor II[**6**]. Hence, all three antagonists form an interaction with a TM-II and TM-VII residue, while doxazosin and compound 6 form an additional interaction with a TM-III residue.

This analysis indicates that the uncomplexed, 'agonist-bound' and 'antagonist-bound' receptor models are structurally different and experience different interactions for the conserved motifs. These results suggest that binding of agonists and antagonists induces different types of conformational changes in the α_{1A}-AR, which is in agreement with the different responses of the receptor when interacting with each.

7.3.3 Comparative study of uncomplexed, agonist-bound and antagonist-bound receptors

Figure 7.9 indicates the RMSDs calculated for each helix of the AD complex and compound **6** complex compared with the ligand-free state, which for clarity were taken as reference agonist and antagonist compounds. Specifically, the agonist adrenaline, produces a larger effect on helices TM-I, TM-VI and TM-VII, whereas the antagonist, compound **6**, induces larger changes in TM-II and TM-IV. The previous finding of the importance of structural changes in TM-III and TM-VI for the activation of the receptor is consistent with deductions made from experimental spin labeling studies of light activation of rhodopsin[86] and with results previously found for the β_2-AR.[87,91]

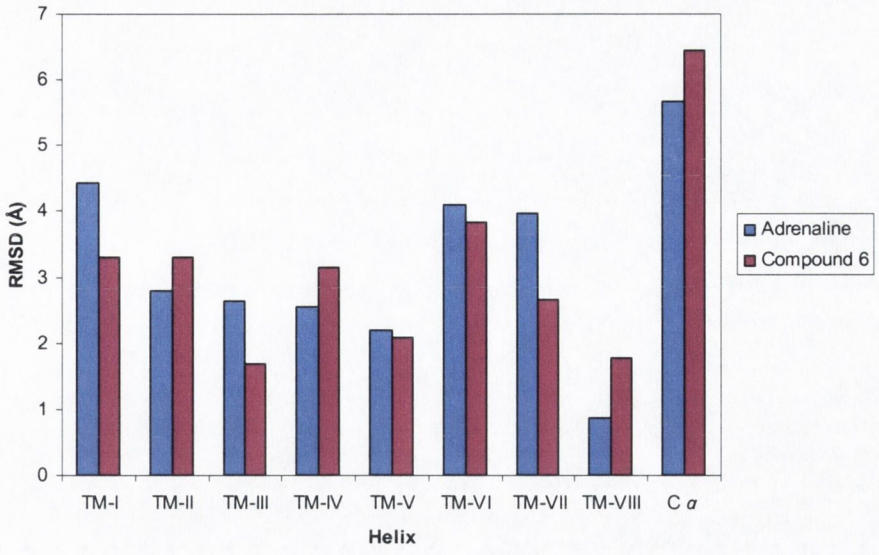


Figure 7.9: RMSD values of each helix in the adrenaline complex and the compound **6** complex compared with the corresponding helix in the ligand-free receptor.

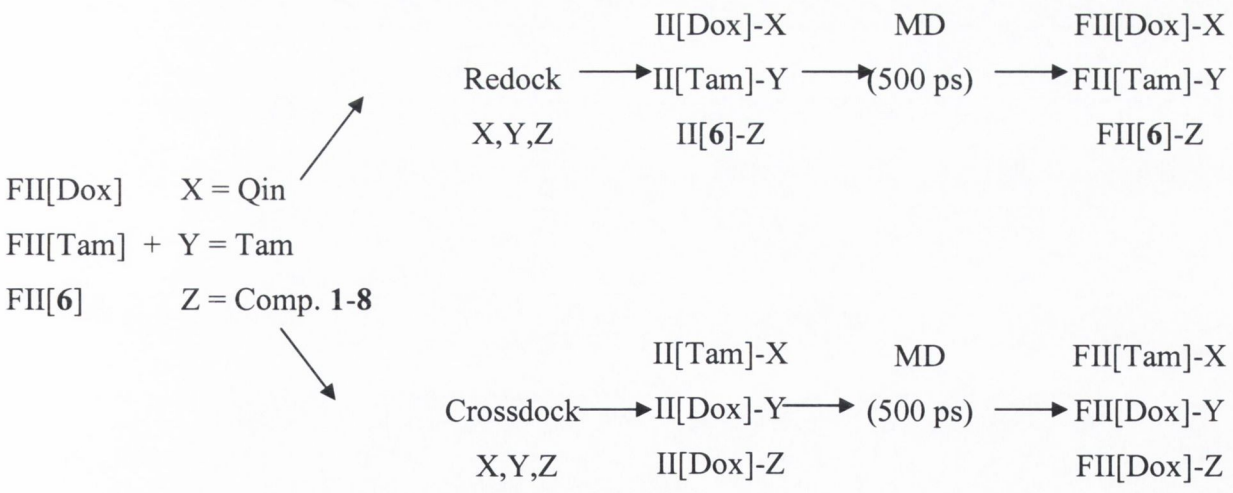
By examining the differences in the inactive ligand-free, the inactive antagonist-bound and the active agonist-bound complexes obtained, a number of sites can be proposed for future mutagenesis studies. Among the residues common to both agonists and antagonists, are Asp-106, Ser-188 and Ser-192. As shown, ligands that bind in the same binding pocket can nevertheless have both common and distinct receptor-residue interactions. The main variations occur in interactions with contact residues of EC-II. Specifically, Thr-174 is recommended for a mutation that could affect agonist activation, as it is involved in four of the agonist-bound simulations and none of the antagonist bound simulations. Among the residues involved in the interaction with our tested antagonists that are not present in the interactions of the agonists are Ile-178, Asn-179 and Glu-180 of EC-II. Two of these (Ile-178 and Asn-179), were previously identified by mutagenesis studies as being involved in subtype selectivity of the antagonist WB4101.[100] The additional residue, Glu-180, to our knowledge has not been considered in mutagenesis studies. Mutations at these sites would be particularly useful in providing additional insights into the validity of the models and the ligand binding pocket.

In summary, the different effects on receptor conformation of agonist and antagonist binding were examined by comparison of the ligand-receptor complexes with each other and with the ligand free receptor model. These complexes were used to probe the mechanism of receptor activation by identifying differences in receptor

conformation between the agonist and the antagonist complexes during unconstrained dynamics. They also allowed the selection of candidate sites for future mutagenesis experiments. The overall analysis indicates that the 'antagonist-bound' receptor is in a different conformation than the uncomplexed inactive state.

7.4 Redocking and Crossdocking of Antagonist Structural Classes with the 'antagonist-induced' α_{1A}-AR models

In the following sections we utilized the developed 'antagonist-bound' receptor conformations, II[Dox], II[Tam] and II[**6**] for redocking antagonists of the same structural class. The receptor structure II[Dox] was utilized for redocking the quinazoline class, receptor structure II[Tam] for redocking tamsulosin and receptor structure II[**6**] for redocking the bis-guanidinium and bis-imidazolidinium compounds (**1-8**), see Scheme 7.1. A series of crossdocking experiments were also performed to examine the use of the different 'antagonist-bound' receptor forms in general ligand screening. The quinazoline based antagonists were crossdocked with the II[Tam] receptor structure. For the selective antagonist tamsulosin we crossdocked with the II[Dox] receptor structure. Finally, for the bis-guanidinium and bis-imidazolidinium compounds (**1-8**) we crossdocked with the II[Dox] receptor structure.



Scheme 7.1: Computational pipeline for use of 'antagonist-bound' receptor structures in Chapter VII, where X, Y and Z indicate the various antagonists, Qin. for the quinazolines and Tam for Tamsulosin.

7.4.1 Binding Interactions of Current Clinical Antagonists – Quinazoline Class - with the ligand-induced α_{1A}-ARs

The receptor structure II[Dox] was redocked with the current clinical antagonists alfuzosin, terazosin, prazosin and doxazosin for which a large variety of interactions emerged over the MD simulations (500 ps). For the II[Dox]-dox complex, the 4-amino group interacts with both Asp-106 and Tyr-316, while an oxygen of the pyrane ring interacts occasionally with Met-292 over the simulation (Figure 7.10). For the alfuzosin complex, while the 4-amino group is close to Asp-106 it is not correctly orientated for binding. However, Glu-180 of EC-II interacts with the furan ring of alfuzosin and Ser-188 interacts with a methylene linker. For the terazosin complex, the 4-amino group interacts with Ile-157, Ser-158 and Ser-188; while Ser-192 interacts with one of the methylene linkers. For the structurally similar antagonist prazosin no interactions were observed.

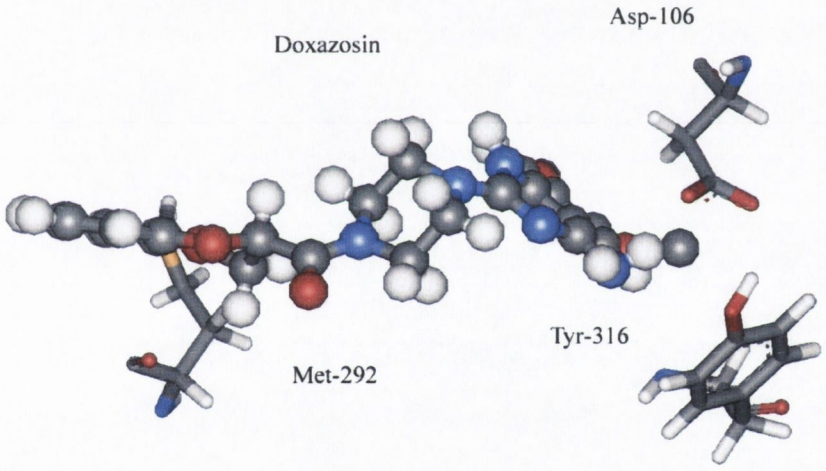


Figure 7.10: Binding mode of doxazosin in the FII[Dox]-dox complex.

When cross-docked with receptor structure II[Tam], the protonated nitrogen of doxazosin interacts with Asp-106, while a methoxy group interacts with Ile-178. For the prazosin and terazosin complexes, similar binding interactions were observed with the 4-amino group interacting with Asp-106. Secondary interactions involved a methoxy group and Ile-178 for terazosin and Gln-166 for prazosin. Ser-296 is also in close proximity to the carbonyl group connected to the piperazine ring of the three antagonists. For the crossdocked alfuzosin complex, the protonated 1-nitrogen of the

quinazoline ring interacts with Asp-106 while a neighbouring methylene linker interacts with Glu-180.

The analysis indicates that the two ligands have induced slightly different effects on the binding site and upon redocking into II[Dox] and crossdocking into II[Tam] different binding modes emerge for the quinazoline ligands. With receptor structure II[Dox], only doxazosin forms an interaction with Asp-106, while terazosin forms a series of interactions with Ser residues including Ser-158, Ser-188 and Ser-192. Stronger anchors are formed between the 4-amino group of the quinazoline antagonists and Asp-106 with receptor structure II[Tam], while secondary interactions involved the extracellular end of TM-IV and the neighbouring EC-II. For the antagonist alfuzosin, an interaction occurs with Glu-180 of EC-II and Ser-188 of TM-V for both receptors II[Dox] and II[Tam].

7.4.2 Examination of Interaction Energies, Binding Energies and Binding Affinities for the quinazoline antagonists with the ligand-induced α_{1A}-ARs

Through examining the interaction and binding energies, coupled with the predicted binding affinities an estimate of the fitness score of a ligand molecule to its target protein can be obtained for that given complex structure. Ligand interaction and binding energies are highly dependent on the correct prediction of the bound structure of a ligand and on an accurate prediction of all interactions involved. As the solvation contribution to the binding free energy is determined as the difference between solvation free energies of the complex, the receptor and the ligand (the first two usually being on the order of $10^3 - 10^4$ kcal mol^{-1}), small absolute errors in these solvation free energies lead to large (relative) error in the predicted binding free energy. Hence, estimating reliable ligand binding energetics is a challenging task. In this study, the complexes were averaged over the last 200 ps of simulation, were optimised under solvated conditions and the IE and BE values were obtained as static Amber MM calculations. Examining the IE and BE, doxazosin, which was utilized in developing the II[Dox] receptor structure had the strongest interaction when redocked with receptor structure II[Dox], followed by alfuzosin, terazosin and prazosin (see Table 7.4). For the II[Tam] structure, which was not developed utilising a quinazoline ligand, different orders were observed for IE and BE. In terms of IE, terazosin had the strongest interaction, followed by alfuzosin, doxazosin and prazosin. In terms of BE, doxazosin

followed by alfuzosin, prazosin and terazosin had the strongest interactions with receptor II[Tam], see Table 7.4.

An essential element in the ligand design process is to predict binding affinities (BAs) for candidate ligands. It provides a means to score compounds and screen virtual compound libraries in an attempt to enhance selection of those members, which are most likely to be active against the target of interest. The BAs of the final complexes were estimated using the XScore scoring function. In terms of BA, doxazosin scored best and alfuzosin lowest for both receptor structures. However, the predicted antagonist affinities range from 5.84 - 7.38 can be assigned as medium affinity drugs ($5.0 < -\log K_d < 8.0$).[293]

Table 7.4: Interactions Energies (IEs) and binding energies (BEs) (in kcal mol^{-1}) for the quinazoline antagonists with receptor structures II[Dox] and II[Tam]. Binding Affinities (BAs) were determined with Xscore.

Compounds	IE		BE		BA	
	II[Dox]	II[Tam]	II[Dox]	II[Tam]	II[Dox]	II[Tam]
Alfuzosin	-224.4	-251.2	-320.4	-206.1	5.84	5.87
Doxazosin	-246.3	-246.1	-310.6	-336.9	7.27	7.38
Terazosin	-169.3	-262.8	-669.8	-179.3	6.58	6.25
Prazosin	-72.9	-271.2	-286.8	-238.8	6.47	6.4

The predicted binding affinity scores and the experimental binding constants (K_i, nM)[73] for the α_1-ARs antagonists, prazosin, alfuzosin, doxazosin and terazosin follow a similar trend but a low correlation was found for receptor II[Dox] with an R^2 of 0.36 and an even smaller correlation for II[Tam] with an R^2 of 0.32. As prazosin appeared is an outlier the R^2 values were recalculated for the other three quinazoline antagonists, which resulted in a higher R^2 of 0.99 for receptor structure II[Dox] and an R^2 of 0.89 for receptor structure II[Tam]. This analysis indicates a high correlation between the predicted binding affinity scores and the experimental binding constants of the three quinazoline ligands with both receptor structures.

7.4.3 Binding Interactions of the Current Clinical Antagonist – Tamsulosin - with the ligand-induced α_{1A}-ARs

When tamsulosin was redocked with the II[Tam] receptor structure, over the production run the protonated phenethyl amine moiety interacts with Glu-180, while the sulphonamide group interacts with Asp-106 and Ile-178, see Figure 7.11.

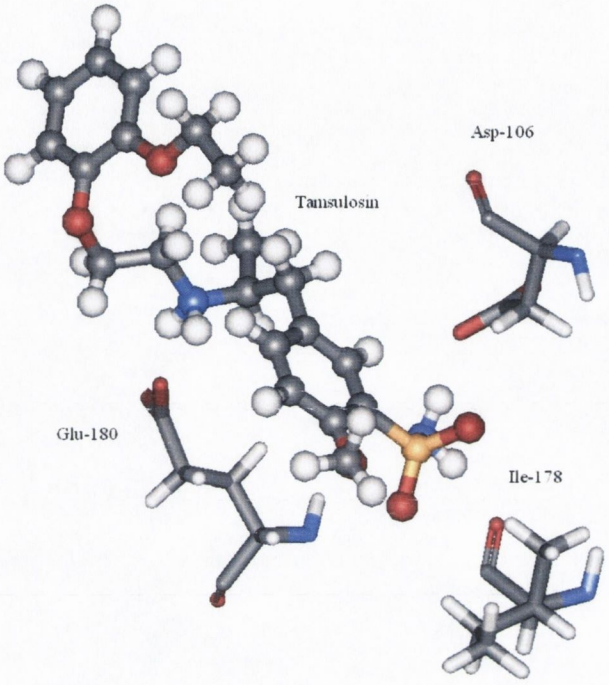


Figure 7.11: Binding mode of tamsulosin in the FII[Tam]-tam complex.

When tamsulosin was crossdocked with the II[Dox] receptor structure a different binding mode is observed, Figure 7.12. The ethanyl amine chain of tamsulosin, interacts with Asp-106 and Ser-188 while the Glu-180 interacts with the methoxy group. This is in agreement with the postulate of Ishiguro *et al.*[286] that the phenethyl amine moiety interacts with the carboxylic acid of Asp-106 in TM-III and the methoxy group with the hydroxyl group of Ser-188 in TM-V.

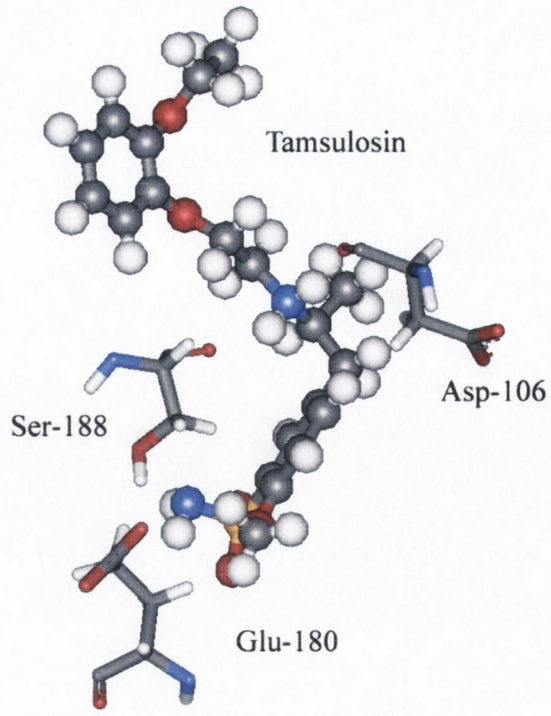


Figure 7.12: Binding mode of tamsulosin in the FII[Dox]-tam complex.

Lower IEs were determined for tamsulosin redocked into receptor structure II[Tam] at –255.1 kcal mol^{-1} than with receptor structure II[Dox] at -236.4 kcal mol^{-1}. Similarly, the BEs were -356.3 and -239.0 kcal mol^{-1} respectively, indicating that tamsulosin redocked better into the receptor structure induced by tamsulosin than into the quinazoline induced receptor structure. The computed BA results at 6.95 and 6.08 respectively, again correspond with a medium range affinity for tamsulosin with the α_{1A}-AR. In summary, it is the receptor structure developed with the ligand which gives rise to stronger ligand-receptor interactions, as determined by the IE, BE and BA.

7.4.4 Binding Modes of Antagonists 1-8 with the ligand-induced α_{1A}-ARs

The series of antagonists (**1-8**), were redocked into receptor structure II[**6**] for which poor interactions with the receptor were observed for all the complexed imidazolidinium derivatives (**1-4**). A better series of interactions were observed for the guanidinium compounds. For the II[**6**]-**5** complex the outer guanidinium nitrogens interact with Asp-106, Ile-178 and Glu-180 at one end and with Cys-110, Thr-111 and Ser-192 at the other guanidinium moiety. For the complexes of compounds **6-8** all three interact with Asp-106, Ile-178 and Glu-180 at one guanidinium end and only differ in their interactions at the other end of the molecule. All three interact with Cys-110 and

Ser-192, while the II[**6**]-**6** and II[**6**]-**8** complexes also interacts with Trp-285 (see Figure 7.13). The II[**6**]-**7** complex additionally interacts with Ser-188. Hence, the main interactions for compounds **5-8**, involve TM-III residue Asp-106, EC-II residues Ile-178 and Glu-180 at one guanidinium end and TM-III residue Cys-110, TM-V residues Ser-188 and Ser-192 and TM-VI residue Trp-285 at the other.

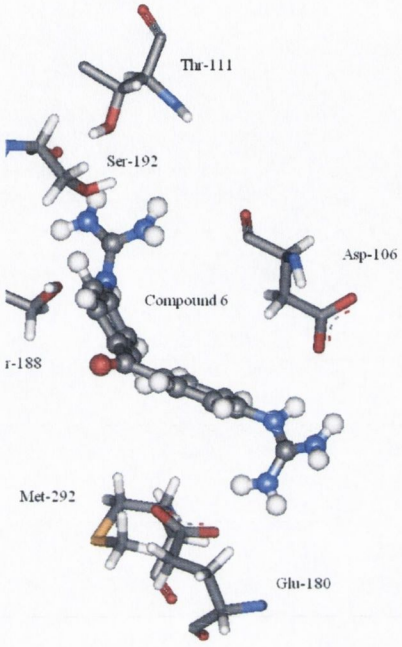


Figure 7.13: Binding mode of compound **6** in the II[**6**] receptor structure.

To further investigate the different behaviour of the imidazolidinium derivatives with respect to the guanidinium derivative, compound **6** was mutated to all the compounds **1-8** and further MD simulations (500 ps) were performed. For compounds **1-4**, interactions appear to again be sterically prohibited due to clashes with Ala-103 and no favorable interactions occurred. Similar interactions occurred with the guanidinium compounds as in the previous redocked analysis.

When antagonists **1-8** were crossdocked in receptor structure II[Dox], again there were no interactions formed for the imidazolidinium compounds **1-4**, as they appear too large for the site and are sterically hindered. The smaller bis-guanidinium compounds form a number of interactions. Compound **5**, only utilized one half of the molecule to form two interactions involving the outer guanidinium nitrogens with Asp-106 and Ser-188. For the remaining compounds, **6-8**, both guanidinium ends of the molecule were involved in forming interactions but they differed from the binding

modes observed in II[**6**]. For compound **6**, six interactions are observed involving Asp-106, Glu-180 and Met-292 orientated around one guanidinium end and Thr-111, Ser-188 and Ser-192 around the alternative guanidinium moiety (Figure 7.14). For the II[Dox]-**7** complex, a different orientation was observed, with one guanidinium moiety interacting with Asp-106 and Gly-315, while the alternative end interacts with Ile-178, Glu-180 and Met-292. For the II[Dox]-**8** complex, interactions occur with Asp-106, Ser-188 and Ser-192 at one end and Asn-179 and Met-292 at the other guanidinium end.

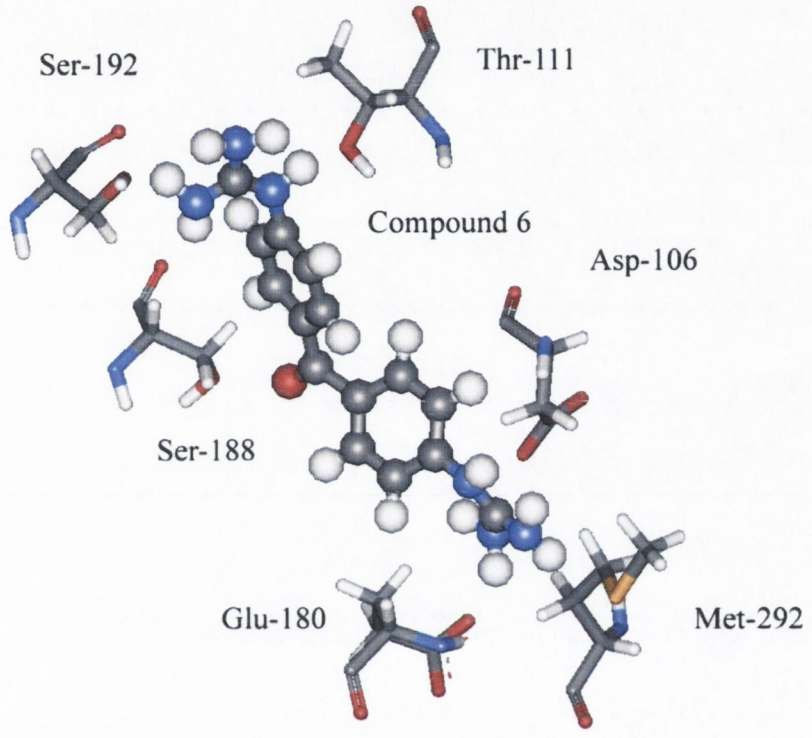


Figure 7.14: Binding mode of compound **6** in the II[Dox] receptor structure.

In summary, when the series of antagonists (**1-8**) were redocked into receptor structure II[**6**] and crossdocked into receptor structure II[Dox] poor interactions were observed for all the complexed imidazolidinium derivatives (**1-4**) while better interactions were observed for the guanidinium compounds (**5-8**). This finding correlates with the observed experimental data, for which the guanidinium compounds exhibited a higher percentage inhibition of the contraction of noradrenaline in tissue with BPH.[61] Again differences were observed in the binding modes of the various guanidinium compounds for the two ligand induced receptor models. The main interactions, occurring at both ends of the guanidinium derivatives, involved Asp-106, Cys-110 and Thr-111 in TM-III, a variety of EC-II residues including Ile-178, Asn-179 and Glu-180 and Ser-118 and Ser-192 in TM-V. Additionally, there was an interaction

formed with Trp-285 for complex II[**6**]-**6** and with Met-292 and Gly-315 for a number of crossdocked complexes. In this study, residues in EC-II appear particularly important in the binding of the guanidinium antagonists, including Ile-178, Asn-179 and Glu-180.

7.4.5 Examination of Interaction Energies, Binding Energies and Binding Affinities for antagonists 1-8 with the ligand-induced α_{1A}-ARs

A comparison of the IE, BE and BA values was performed for antagonists **1-8** with receptor structures II[Dox] and II[**6**], see Table 7.5. In general, for the II[Dox] receptor structure the guanidinium compounds have lower IEs than the imidazolidinium compounds. Compound **6** scores the lowest, which is in agreement with the available experimental data.[61] Furthermore, all the IEs are lower for receptor II[**6**], which was induced by an antagonist of this structural class, than with the II[Dox] receptor structure. In terms of BEs, compound **5** was ranked the lowest, followed by compound **2** and then compound **6**. Alternatively, compound **1**, **8** and then **6** exhibited the lowest BEs with receptor structure II[**6**]. Hence, in both instances, compound **6**, which performed best in biological studies[61] was ranked third lowest by BE. However, the BE are very poor for receptor II[**6**] even though the structure was generated with compound **6** bound.

As expected the predicted binding affinities of the guanidinium compounds with receptor II[**6**] are notably higher than with receptor structure II[Dox]. However, the imidazolidinium compounds have a higher predicted affinity than the guanidinium compounds in contrast with the experimental data[61] (see Table 7.5). This analysis calls into question the general applicability of this scoring function to accurately predict binding affinities particularly for this class of ligands.

Table 7.5: Binding Interactions (in kcal mol^{-1}) for the antagonists (**1-8**) with the receptor structures II[Dox] and II[**6**]. Binding Affinities were determined with Xscore.

	Interaction Energy		Binding Energy		Binding Affinity	
Compound	**II[Dox]**	**II[6]**	**II[Dox]**	**II[6]**	**II[Dox]**	**II[6]**
1	-386.5	-389.2	-194.2	-144.9	7.19	7.45
2	-386.1	-452.2	-497.4	33.1	7.57	7.41
3	-401.0	-430.7	-805.7	256.5	6.97	7.09
4	-397.6	-409.0	-647.9	-35.4	7.33	6.8
5	-328.3	-470.6	-460.1	222.8	6.11	7.04
6	-445.3	-445.8	-335.7	-38.2	6.64	7.18
7	-425.6	-437.5	-478.6	-39.8	6.94	7.09
8	-402.9	-442.2	-620.7	-128.5	6.84	7.32

In summary, in section 7.4, we have successfully utilised ligand induced receptor structures to redock ligands of a similar structural class, namely the quinazolines and the guanidinium ligands. We have also examined crossdocking of ligands with receptor structures developed utilising a ligand of a different structural class, which yielded slightly different binding interactions to the redocked studies. Our analysis highlights that different ligands exert different effects on the binding site, which influences the resultant binding modes in subsequent docking studies.

7.5 Examination of interactions of antagonists 9-18 with ligand induced α_{1A}-AR

'Antagonist-bound' receptor structures are of use in the design of novel drugs, if they can successfully be used for screening of possible antagonists and hence have an application in future SBDD or virtual high throughput screening (vHTS). Our three optimised 'antagonist-induced' receptor structures, with the ligands removed (II[Dox], II[Tam] and II[**6**]), were utilised in a virtual screening procedure. The optimised receptors were docked with a series of α_1-AR (**14** and **17**) and α_{1A}-AR (**9-13, 15-16** and **18**) ligands and MD simulations (500 ps) were performed, see Scheme 7.2.

II[Dox]	Dock	II[Dox]-X	MD	FII[Dox]-X
II[Tam]	Compounds (**9**-	II[Tam]-X	(500 ps)	FII[Tam]-X
II[**6**]	**18**)	II[**6**]-X	⟶	FII[**6**]-X

Scheme 7.2: Computational Pipeline for screening of antagonists with the 'antagonist-induced' receptor structures. The antagonists, X, were labelled **9-18** as before.

7.5.1 Examination of interactions of antagonists 9-18 with receptor II[Dox]

The subsequent analysis focuses on those antagonists that form stable HB anchors over the 500 ps simulations with Asp-106 of II[Dox]. This occurs for antagonists **10**, **11**, **14** and **17** with receptor structure II[Dox]. In complex II[Dox]-**10**, an interaction between Asp-106 and the ether oxygen is formed, while the piperazine moiety interacts with Ile-178, Glu-180 and Ser-188. For compound **11**, interactions formed with Asp-106 and Glu-180. In the case of complex II[Dox]-**14**, the protonated nitrogen of **14** interacts with Asp-106 and Met-292, while Glu-180 interacts with a methylene group neighbouring the amino group. Finally, for complex II[Dox]-**17**, interactions are formed between the protonated nitrogen of **17** and Asp-106 and Ser-188 of the receptor. For the remaining complexes, compound **9** formed interactions with Asn-179 and Ser-188, while compound **12** interacted with Asn-179, Glu-180 and Ser-188. Compound **13** formed interactions with Glu-180 and Met-292. Compound **15** formed no favorable interactions while compound **16** interacted solely with Ser-188 and compound **18** solely with Glu-180.

Receptor structure II[Dox], does not appear suitable for successful screening of our test set of active α_{1A}-AR ligands as only four of the ten test compounds formed a favorable interaction with Asp-106. Of those ligands which formed favorable interactions the main residues involved included Asp-106 of TM-III, Ile-178, Asn-179 and Glu-180 of EC-II, Ser-188 of TM-V and Met-292 of TM-VI.

7.5.2 Examination of interactions of antagonists 9-18 with receptor II[Tam]

For the II[Tam] receptor structure, interactions with Asp-106 occur with antagonists **14**, **16**, **17**. For the II[Tam]-**14** complex, Asp-106 interacts with an oxygen of the dioxane ring, while Glu-180 interacts with the methoxy group. In the case of II[Tam]-**16**, a methylene group of **16** interacts with Asp-106, Glu-180 and Ser-188, while Ile-178 interacts with the dimethoxy phenyl ring. Finally, in the complex

the remaining antagonist compounds, compounds **9** and **11** interacted with Glu-180 and Ser-188, while compound **9** formed an additional interaction with Ser-192. The remaining compounds interacted solely with either Glu-180 (compounds **12** and **13**) or with Ser-188 (compound **10**, **15** and **18**).

For receptor structure II[Tam], only three of the antagonists formed a favorable interaction with Asp-106 and this ligand-induced receptor again does not appear suitable for successful screening of our test set of α_{1A}-AR ligands. The main interacting residues included Asp-106 of TM-III, Ile-178 and Glu-180 of EC-II and Ser-188 of TM-V, while no interactions were formed with Met-292 of TM-VI unlike for the II[Dox] receptor structure.

7.5.3 Examination of interactions of antagonists 9-18 with receptor II[6]

For the II[**6**] receptor, five of the ten test compounds formed a favorable interaction with Asp-106, compounds, **10-12** and **15-16**. Compound **10** interacts with Asp-106 through the methyl group (on the pyrimidinedione ring). Compound **11** interacts with Asp-106 and Glu-180 through two carbons of the piperazine ring. Compound **12** interacts through the protonated nitrogen with Asp-106 and Glu-180 and through a piperazine nitrogen with Ser-188. Compound **15** has an interaction between the protonated nitrogen and Asp-106, while a secondary interaction occurred between the methyl linker and Asp-106. Finally, compound **16** interacts with Asp-106 through a methoxy group and with Glu-180 through a methylene linker.

Of the remaining complexes, compound **9** interacts solely with Glu-180, compound **14** with Asn-179, compound **18** with Ser-188 while compounds **13** and **17** form no interactions with the receptor model.

In summary, for receptor structure II[**6**], similar to II[Dox] and II[Tam] the main interacting residues included residues in TM-III, EC-II and TM-V. Only receptor structure II[Dox] formed an additional interaction with a residue of TM-VI. The various ligands exert differing conformations on the binding site and none of the ligand induced receptor structures appear to be suitable for successful screening of the structurally different, α_{1A}-AR ligands of our test set.

7.5.4 Examination of Interaction Energies, Binding Energies and Binding Affinities for antagonists 9-18 with the ligand-induced α_{1A}-ARs

The interaction energies, binding energies and binding affinities were again calculated and the resultant ranking of the ligands (**9-18**) determined for each receptor structure. In Table 7.6, the complexes are listed in order of decreasing stability, with the lowest ranks assigned to complexes with the lowest IE, BE and highest BA. The complex with the lowest interaction energy varied over the three receptor structures, ranging from compound **9** for II[Dox], compound **14** for receptor structure II[Tam] and compound **11** for receptor structure II[**6**]. All three of the top ranked antagonists are selective α_{1A}-AR antagonists. The non-consistency in the ranking when ordered by IE indicates that the results are heavily dependent on the utilized receptor structure.

Our analysis also indicates that the ranking of ligands varies both across receptor structure and across scoring method. The antagonists with the lowest BEs again differed across the three receptor structures, from compound **10** for II[Dox], compound **13** for II[Tam] and compound **17** for II[**6**]. Of these, compound **10** is a known selective antagonists, compound **13** is thought to be α_{1A}-AR selective while compound **17** is a known non-selective α_1-AR antagonist. Similarly, the antagonists with the highest binding affinities again differed across the three receptor structures, from compound **9** for II[Dox], compound **12** for II[Tam] and compound **11** for II[**6**]. Of these, compounds **9** and **11** are known selective antagonists while compound **12** is also though to be selective for the α_{1A}-AR.

Our analysis raises two important issues. Firstly, the same scoring function ranks the antagonists differently for the three ligand induced receptor structures. This discrepancy may be a consequence of not inducing an appropriate receptor structure for that class of ligand. This is consistent with our postulate that it is necessary to induce a receptor structure by a ligand of the structural class being studied. Secondly, the different methods (IE, BE and BA) give different results for the same ligand-induced receptor structure. This highlights problems with the quality of the available methods of evaluating ligand-receptor complexes and that our dataset may not be large enough to be able to distinguish clearly between ligands of high and low activity.

Table 7.6: Ranking of antagonists (**9-18**) for the various receptor structures, according to the interaction energies (IE) and binding energies (BE) determined using the Amber force field. Binding affinities (BA) were determined using the Xscore scoring function. Compounds are ranked in order of decreasing stability of the complex.

	Compound No. ranked by IE.			**Compound No. ranked by BE.**			**Compound No. ranked by BA.**		
Rank	**II[Dox]**	**II[Tam]**	**II[6]**	**II[Dox]**	**II[Tam]**	**II[6]**	**II[Dox]**	**II[Tam]**	**II[6]**
1st	**9**	**14**	**11**	**10**	**13**	**17**	**9**	**12**	**11**
2nd	**14**	**9**	**12**	**13**	**18**	**14**	**15**	**18**	**18**
3rd	**12**	**16**	**18**	**9**	**16**	**15**	**11**	**16**	**15**
4th	**10**	**17**	**10**	**12**	**15**	**18**	**17**	**11**	**12**
5th	**16**	**10**	**16**	**18**	**11**	**12**	**12**	**17**	**10**
6th	**11**	**12**	**9**	**17**	**12**	**9**	**16**	**13**	**9**
7th	**18**	**13**	**15**	**14**	**9**	**11**	**10**	**10**	**16**
8th	**15**	**11**	**14**	**15**	**10**	**13**	**18**	**9**	**17**
9th	**17**	**18**	**13**	**16**	**14**	**16**	**14**	**14**	**14**
10th	**13**	**15**	**17**	**11**	**17**	**10**	**13**	**15**	**13**

7.6 Discussion / Conclusions

To obtain antagonist induced α_{1A}-AR conformations, antagonists were docked into the initial α_{1A}-AR models and the resulting complexes were optimised through MD simulations in a membrane mimic. The structural movement of the antagonist-induced complexes were monitored by comparison to the inactive receptor forms after 1 ns MD simulations. Different conformations emerged for TM-V with receptor structures FII[Dox] and FII[**6**], having a marked kink at the Pro residue, while for structure FII[Tam] the helix has straightened relative to the uncomplexed form. This movement allows for a change in the interactions of the antagonists with the α_{1A}-AR and this is consistent with the idea that GPCRs undergo facile transitions between conformational states under normal conditions and that ligands stabilize different receptor conformations.[294,295] Our RMSD results also suggest that, in agreement with different functional responses observed that binding of agonists and antagonists induces different helical changes in the α_{1A}-AR.

Subsequent MD simulations of the antagonist induced α_{1A}-ARs redocked with antagonists of the same structural class and crossdocked with antagonists of different structural classes were pursued. For the current clinical antagonists doxazosin and tamsulosin have induced slightly different effects on the binding site. Upon redocking into II[Dox] and crossdocking into II[Tam] different binding modes emerge for the quinazoline ligands. Similarly, a different orientation arises for the α_{1A}-AR selective antagonist tamsulosin in the receptor structures II[Dox] and II[Tam], with different nitrogens interacting with Asp-106. The bis-imidazolidinium (**1-4**) compounds did not interact as strongly as the guanidinium compounds when redocked or crossdocked. For compounds **5-8** interactions occurred at both bis-guanidinium moieties involving Asp-106 in TM-III, Ser-188 and Ser-192 in TM-V and residues Ile-178 and Glu-180 of the EC-II of the α_{1A}-AR. Unlike the TM regions, the extracellular loops, including EC-II, are highly divergent among Rhodopsin-like GPCRs and may provide more ligand contact variety among the receptor subtypes than the TMs. Again there were subtle differences in the binding modes in the two receptor structures as the various ligands perturbed the binding site in slightly different ways when producing the ligand-induced receptor models.

Our initial results, together with earlier studies,[296] strongly support the idea that the current rhodopsin crystal structure is not an optimal structural template for direct homology modeling of GPCR antagonist complexes. Furthermore, our study indicates that energy minimizations are not sufficient for structural refinement of receptors and MD simulations are necessary to induce 'antagonist-bound' receptor forms. Our three ligand-induced receptor models (II[Dox], II[Tam] and II[**6**]) were used to test binding of a range of known α_1–AR antagonists, **9-18**, where favorable interactions were found with the EC-II residues, Ile-178, Asn-179 and Glu-180 and also with Ser-188 in TM-V, which is consistent with the available body of experimental work. Our binding site analysis also indicates that receptor structures induced by the interaction with one class of antagonist are not directly suitable for screening of another structural class. It appears necessary to produce an 'antagonist-bound' receptor form for each class of antagonist compound if we wish to evaluate their detailed interactions with the α_{1A}-AR and establish possible binding modes.

The work of Evers *et al.*,[195] on antagonist/α_{1A}-AR complexes concluded that rhodopsin based homology models may be used as the structural basis for GPCR lead finding and compound optimization. Such studies can provide valuable information for antagonist binding sites; however our study would indicate that this approach will favour antagonists of a similar structure to that used in developing the homology model. Indeed, those antagonist identified by Evers *et al.*,[195] using virtual screening have similar chemical structures to the antagonist used in developing there receptor model. To obtain novel antagonists a greater degree of flexibility within the receptor model will be required, which could be generated through the use of multiple ligand induced receptor structures obtained for related classes either through a dynamic approach, as used here, or possibly through homology modeling utilizing ligand restraints.[195] The advantage of our approach is that the binding site and conformational changes occur as part of the iterative cycle of docking and dynamics allowing new interactions to form and small but possibly vital conformational change to occur as the ligand is accommodated in the binding site. Such a procedure is a more rational approach than direct docking into a rigid rhodopsin based homology model as it allows increased flexibility of the receptor. Our approach can in principle be applied to any member of the GPCR family with known ligand information and site-directed mutagenesis data. The predicted 3D structures and antagonist binding modes could also be used in

designing mutagenesis experiments to validate the structure of the binding sites of the computational models.

A ranking of the complexes involving antagonists **9-18** and the three 'antagonist-induced' receptor structures was pursued using the Amber FF and the Xscore scoring function. Our analysis also indicates that the ranking of ligands varies both across receptor structure and across the employed scoring methods (IE, BE and BA) which further highlights the need to develop an 'antagonist-bound' receptor form for each class of antagonist compound to be studied and that current scoring functions do not appear developed enough to distinguish between compounds of similar affinity.

Chapter VIII: Discussion / Conclusions & Further Studies

"Experience is what you get when you don't get what you want."

Dan Stanford

8.0 Discussion

A detailed SBDD study of the α_{1A}-AR has been hampered due to the lack of a crystal structure of Class A GPCRs in the inactive (other than bovine rhodopsin) or active state. The aim of this work was to determine the binding interactions and conformational changes induced by structurally diverse ligands using computational biophysical approaches. A computational protocol was developed, which integrated α_{1A}-AR structure prediction, MD simulations and automated docking to investigate the α_{1A}-AR and its complexes

Comparative modelling was performed to predict the 3D structure of the α_{1A}-AR based on both a pairwise and a multiple alignment to the bovine Rhodopsin sequence. One particularly complex issue is modelling the effect on the GPCR structure of the multi-phase environment in which these receptors are naturally embedded. An explicit representation of the phospholipid environment is computationally expensive and in this work a biphasic bilayer was introduced to model the membrane. The use of a membrane mimetic has the additional advantage of significantly reducing the computational expense of equilibration when compared to simulations in a lipid bilayer media. Comparisons were made between simulations performed in a TIP3P water model, as well as $H_2O/CHCl_3/H_2O$ and $H_2O/CCl_4/H_2O$ bilayers. The analysis indicated that the backbone conformation and the dynamical stability of the receptor are within the limits established for reliable structures. After the MD simulations (1 ns), the final structures were both sterically and energetically satisfactory and represent inactive forms of the α_{1A}-AR. Stabilising structural features were examined including disulfide bridges, ion pairs and inter-helical HBs. Our developed homology based α_{1A}-AR models represent the first published study of the inactive form of the α_{1A}-AR, refined in an appropriate solvent environment.[190]

A series of ligands (agonists and antagonists), which act on α_1-ARs were studied to gain insights into their chemical state under physiological conditions (*i.e.* protonation states and conformations). The agonists, adrenaline and noradrenaline, have only one nitrogen, which can be protonated and participate in the receptor/ligand interactions, while the antagonists tend to be larger molecules and have four or five nitrogens present. Trends were observed in the PAs obtained for the different nitrogens of each of the families studied. The PA values determined for the first protonation of all the

compounds studied are similar energetically (~250 kcal mol^{-1}). The first protonated nitrogens determined included, the 1-nitrogen of the quinazoline ring for the current clinical antagonists, the inner nitrogens at both ends of the bis-imino-imidazolidinium and bis-guanidinium compounds (**1-8**), a nitrogen of the piperazine ring for compounds **9**, **10** and **13** at the heterogeneous bridge for compounds **11** and **12**, at the sole nitrogen for compounds **14-17** and at the indole ring for compound **18**. In addition, a good correlation was found between the PA values corresponding to the first protonation in the gas phase of some of the compounds studied and their corresponding affinity constants K_i, which measures the affinity of the compounds for the α_{1A}-AR. To replicate physiological conditions it is necessary to include solvation effects, which were examined utilising the Onsager and PCM solvation techniques. Our studies showed that the inclusion of solvation effects yielded stabilization in the PA values obtained, which were more marked when utilising the PCM solvation technique.

Proton transfer is a vital part of many chemical processes and is determined by the acid dissociation constants (pKa) of the compounds involved. A further goal of this thesis was to utilise existing quantum chemical methods to predict pKa values of the series of antagonists. The technique reproduced the experimental pKas of the clinical antagonists utilising the Onsager technique and a $\Delta G^0_{solv}(H^+)$ value of –255 kcal mol^{-1}. However, the theoretical technique failed to reproduce the experimental pKas of the catecholamines and the antagonists **1-8**, indicating that the technique is not universally applicable. The loss or gain of a proton inherently involves ionic species, so that accurate solvation effects must be included to cancel the very high gas-phase protonation and deprotonation energies. In addition, the use of a finite basis set introduces errors in the relative gas-phase energies of charged and uncharged molecules. Absolute pKa calculations are further hindered as there is no accurate determination, experimental or otherwise for the solvation free energy of the proton, $\Delta G^0_{solv}(H^+)$. These limitations make it difficult to accurately predict pKa values theoretically.

The global minimum energy conformation of a ligand does not necessarily correspond to the pharmacologically active one; therefore the conformational freedom of the ligands also needs to be considered. Pharmacophore models can be employed as a preliminary *in silico* virtual screening, as a theoretical test of the selectivity of different

antagonists for the various receptor subtypes. Over a selection of conformers, the fit of the ligands to the pharmacophore models of the α_1-AR subtypes were analysed. The current clinical antagonists were found not to selectively fulfil the α_{1A}-AR pharmacophore model, which is consistent with their non-selectivity observed in clinical studies. For compounds **1**, **2**, **5** and **6** the α_{1A}-AR pharmacophore requirements were fulfilled, which is in agreement with the α_{1A}-AR activity determined for these compounds. However, compounds **4** and **8** exhibited reasonable activity despite failing to meet the pharmacophore criteria as they lacked an appropriately positioned polar bridge. Alternatively, compounds **3** and **7** fulfilled the α_{1A}-AR pharmacophore criteria, while having poor activity. These results indicated that a fit to this pharmacophore is not sufficient to estimate the viability of the antagonist models (**1-8**), with **4** and **8** being false negatives and **3** and **7** being false positives under these criteria. Of the remaining antagonists (**9-18**), compound **9** may prove an interesting agent as antagonist inhibition at the α_{1A}-AR subtype is useful in alleviating obstructive symptoms in cases of BPH.

Binding of an agonist, adrenaline or noradrenaline, to the adrenoceptor causes conformational changes that act as a switch transferring the signal to the trimeric G protein thus stimulating the production of intracellular secondary messengers (e.g. Ca^{2+}). Experimental knowledge of the mechanisms controlling ligand binding, activation, signal transduction and regulation of GPCRs is limited. In this work, we have developed a stepwise protocol to examine the interactions between the catecholamine agonists and our developed α_{1A}-AR models. The homology models were perturbed toward different functional states by placing appropriate ligands in the binding site and running separate MD simulations to relax the structure towards the appropriate state for the bound molecule. Further docking and MD studies of the ligands with 'ligand-induced' receptor forms were pursued. One such study involved mutating adrenaline to noradrenaline in one complex (FI-AD) and running MD simulations over the two structures. Such a study illustrates the developed differences over MD simulations for the two agonists commencing from the same binding mode. A larger structural movement was observed over the noradrenaline mutated simulation as it rearranged to accommodate a different ligand. Further redocked and crossdocked agonist simulations modelled the reorganisation of the native receptor structure induced by an agonist.

The analysis of the interactions of the catecholamine complexes is consistent with the results of mutagenesis experiments and suggests new residues involved in the agonist-binding site. The most critical interactions for the binding of the agonists exist between the ligands and residues in TM-III, TM-V, TM-VI and TM-VII, which is consistent with many earlier reports. The molecular modelling results show that the contact residues for the protonated amine is Asp-106 in TM-III, for the catechol hydroxy groups are Ser residues in TM-V, while for the β-OH it is Thr-174 in TM-III. The contact site for the extra methyl group of adrenaline is less clear but may involve Asp-106 in TM-III or Thr-174 in EC-II.

Some progress has been made toward an understanding of the conformational changes that are associated with GPCR activation. Our observed separation in the intracellular end of TM-V and TM-VI from TM-III is consistent with conformational changes in activation that have been detected by other computational and experimental studies. The role of conserved residues, such as the DRY and NPxxY motifs, in activation, were also examined. Upon examining the interactions of the conserved residues in GPCRs, activation may occur due to a change in the environment of Asp-123 of the DRY motif that causes Arg-124 to move out of the TM helical bundle and change the orientation of residues in IC-II and IC-III that allows for increased affinity of coupling to the G-protein. A combination of theory and experiment will remain the approach of choice, not only to discover the basis for receptor function but also to successfully design therapeutic agents to treat human disease.[297] We await the production of an experimental crystal structure of an adrenoceptor in the active state for a thorough comparison of our simulated results with experiment.

As the development of α_{1A}-AR selective antagonists should alleviate the symptoms of BPH with fewer side effects than non-selective agents, we undertook a detailed study of a large series of antagonists complexed with our α_{1A}-AR models. Our analysis illustrated the difficulties of using homology models for antagonist docking studies and the importance of a suitable ligand-induced receptor structure. Three initial antagonist/α_{1A}-AR complexes were used as starting structures for MD simulations, allowing for receptor flexibility and optimisation of the binding pocket for that ligand. After 1 ns MD simulations, different conformations emerged for TM-V with receptor structures II[Dox] and II[**6**], having a marked kink at the Pro residue, while for structure II[Tam], TM-V has straightened relative to the uncomplexed form. This movement

allowed for a change in the interactions of the antagonists with the α_{1A}-AR and may play an important role in antagonist binding.

Through redocking with our ligand-induced receptor forms, we have succeeded in obtaining binding modes that are consistent with the available mutagenesis data for α_1-AR antagonists and largely involving interactions with Asp-106, the Ser residues of TM-V and residues of EC-II. For the quinazoline-based antagonists, different binding modes occurred with receptor structures II[Dox] and II[Tam]. Similarly, a different orientation of the α_{1A}-AR selective antagonist tamsulosin occurred in the two receptor structures II[Dox] and II[Tam], with differing nitrogens interacting with Asp-106. The bis-imino-imidazolidinium compounds **1-4**, did not interact as well as the bis-guanidinium compounds **5-8** with either receptor model, II[Dox] or II[**6**]. The main interactions for the bis-guanidinium compounds **5-8** involved Asp-106, Ile-178 and Glu-180 at one guanidinium end and Cys-110, Ser-188, Ser-192 and Trp-285 at the other. Diverse binding modes were observed when utilizing different 'ligand-induced' receptor structures, which highlights that the various ligands perturb the binding site of the receptor in different ways. An appropriate choice of 'ligand-induced' receptor structure appears necessary to determine accurate binding modes.

For the screening of the remaining antagonists, **9-18**, only a number of interactions occurred with Asp-106 of TM-III. Further interactions involved the EC-II residues, Ile-178, Asn-179 and Glu-180 and also with Ser-188. We therefore propose a critical role of the EC-II loop in forming the binding site for small molecules. However, a number of α_{1A}-AR selective antagonists did not form favorable interactions with the three 'antagonist-induced' receptors. Hence, our binding site analysis indicates that receptor structures induced by the interaction with one class of antagonist are not directly suitable for screening of another structural class. It appears necessary to produce an 'antagonist-bound' receptor form for each class of antagonist compound if we wish to evaluate their detailed interactions with the α_{1A}-AR and establish possible binding modes. We have also examined the use of scoring methods (IE, BE and BA) to rank the bound complexes. Our analysis indicated that the same scoring function ranks the antagonists differently for the three induced receptor structures, which would further highlight the need for appropriate 'ligand-induced' receptor structures for each structural class of ligand to be studied. In addition the different methods of evaluation of ligand binding (IE, BE and BA) yield different results for the same ligand-induced

receptor structure, which indicates outstanding issues with the quality of the available scoring functions.

Finally, by examining the differences in the inactive ligand-free, the inactive antagonist-bound and the active agonist-bound complexes obtained, a number of sites can be proposed for future mutagenesis studies. Residues common to both agonist and antagonist interactions include Asp-106, Ser-188 and Ser-192. The main variations occur in interactions with contact residues of EC-II. Specifically, Thr-174 is recommended for a mutation that could affect agonist activation, as it is involved in four of the agonist-bound simulations and none of the antagonist bound simulations. Among the residues involved in the interaction with our tested antagonists that are not present in the interactions of the agonists are Ile-178, Asn-179 and Glu-180 of EC-II. Mutations at these sites would be particularly useful in providing additional insight into the validity of the models and the ligand binding pocket.

In conclusion, a structure-based approach to GPCR drug discovery in the absence of the real structures requires a multi-disciplinary approach, where molecular models represent a structural context to efficiently integrate experimental data.[298] In this work, a variety of *in silico* modelling approaches have been applied to α_{1A}-AR, thereby adding to the understanding of the α_{1A}-AR structure and function in a way that is complementary to experimental studies. Model structures for the α_{1A}-AR were developed through homology modelling on which MD studies were performed to examine the inactive receptor and the ligand induced receptor structures. Ultimately, it may require structures of multiple receptors bound to multiple ligands, including agonists and antagonists to determine the details of the conformational changes associated with receptor activation.

8.1 Further Studies

Knowledge of the 3D structure of GPCRs is essential for the understanding of their function, for the docking of existing ligands into their binding sites and for the rational design and/or optimisation of drugs. Over the course of this work a number of further areas arose that could be pursued in future studies.

8.1.1 Further adrenoceptor refinement

Most drugs targeting GPCRs act on multiple receptors simultaneously, requiring a multi-receptor approach to lead optimization. A further potential application of our generated α_{1A}-AR homology model would be to serve as a structural template for homology modelling of further members of the biogenic amine receptor family. Considering the high sequence (and probably structural) similarity of the ligand binding site of the biogenic amine binding receptors, this family represents a challenge for ligand design with respect to the problem of selectivity. Studies in our group are currently being undertaken to model the α_{1B} and α_{1D} adrenoceptor subtypes.[299]

To accurately model a GPCR one must take into consideration the environmental influence on the structure. The use of further membrane mimics could be examined, such as utilising hexane or octane as the hydrophobic moiety. Such studies could be coupled with the development of Amber parameters for phospholipids. A direct comparison of MD simulations of an adrenoceptor in a membrane mimic and a phospholipid environment could then be performed.

Different mechanisms of GPCR dimer formation have been observed for various receptors including the β_2-AR,[300] the δ-opioid receptor[301] and the metabotropic glutamate receptor[302] suggesting that receptor dimerization is not essential for G protein activation but may play a role in other receptor functions such as subtype-specific receptor regulation. The relevance of GPCR dimerization is currently one of the most actively studies areas of GPCR structure and function.[303] Studies could include MD simulations on receptor dimers in the absence of ligand, in the presence of an agonist and in the presence of an antagonist.[62]

The absence of fine resolution structural data on any GPCR makes it especially difficult to propose molecular mechanisms to explain the effects of mutations on receptor function. Computational and theoretical methods could provide a molecular view of the structural and energetic consequences of mutations. With a working model for some of the receptor-antagonist complexes, it should be possible to design mutational studies to test the interactions proposed and to design other molecules that may fit in the binding pocket and act as agonists and antagonists. Mutation studies could be performed on those residues found important in agonist and antagonist binding.

8.1.2 Ligand Approach and Interaction at the Binding Site

There are two options for the approach of a ligand to the binding site; Firstly, to squeeze through the space between TM-VII, TM-I and TM-II. Alternatively, the motion of the IV-V hairpin may leave the entry to the binding cavity open long enough for a ligand to enter in a flip-flop action. Although it is currently difficult to envisage the entrance route of ligands into the binding site crevice and the potential associated conformational rearrangement of EC-II, there may be a direct role of residues in EC-II in ligand binding (Figure 8.1). Steered Molecular Dynamics (SMD) involves the acceleration of conformational changes in biomolecular systems through the application of external forces and can speed up the approach of the ligand to the binding site.[304] The force applied in SMD simulations was initially inspired by Atomic force microscopy (AFM) experiments. Further examination of ligand approach to the α_{1A}–AR could be pursued using the SMD technique.

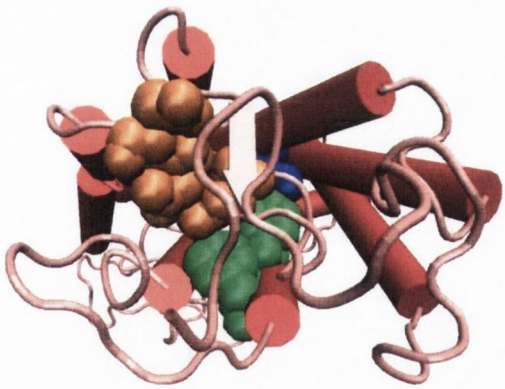


Figure 8.1: The approach path that a ligand would take to the α_{1A}–AR binding site. The relevant residues from experimental information include the residues 106, 188 and 192 in blue, the residues 177, 178 and 179 in orange, Phe 308 and 312 in green and Asp-123 of the DRY motif in yellow.

An investigation of large molecular systems is limited by the computational effort required. Quantum Mechanics/Molecular Mechanics (QM/MM) is a combination of methods that treat different parts of the system at different levels of theory; QM methods, being able to treat chemical reactions while MM methods can provide the biochemical environment for these processes. QM/MM structural optimisations could be performed on the bound complexes, thus allowing for a detailed study of the receptor-binding site.

Publications List:

- Kinsella, G.K., Rozas, I., Watson, G.W., Computational Development of an α_{1A} Adrenoceptor Model., *BBRC*, **2004**, *324*, 916.

- Kinsella, G.K., Watson, G.W., Rozas, I., Calculated proton affinity of α_{1A} adrenoceptor ligands and pharmacophore searches, (submitted to *Biorg. Med. Chem.*, Feb. 2005).

- Kinsella, G.K., Rozas, I., Watson, G.W., Molecular mechanism of interaction of catecholamines with the α_{1A} adrenoceptor – towards a ligand induced receptor conformation (accepted by J. Comput. Aided Mol. Des., May 2005).

- Kinsella, G.K., Rozas, I., Watson, G.W., Computational studies of a selection of antagonists with the α_{1A} adrenoceptor (submitted to J. Med. Chem, Apr. 2005).

- Kinsella, G.K., Rozas, I., Watson, G.W., Comparative Molecular Dynamics Simulations of uncomplexed, 'agonist-bound' and 'antagonist-bound' α_{1A} Adrenoceptor Models, *BBRC*, **2005**, *333*, 3, 737.

- Kinsella, G.K., Rodriguez, F., Watson, G.W., Rozas, I., Evaluation of the pKa for a series of ligands of the alpha1A adrenoceptor (to be submitted).

References

1 *The Economist*, The Economist Newspaper Ltd., 17th October, **1998**.

2 Cai, C.; Li, Z.; Wang, W.; Chen, Y. *Acta Pharmacol Sin.* **2004**, *25*, 1, 1.

3 http://nobelprize.org/chemistry/laureates/1962/, May **2005**.

4 http://www.rcsb.org/pdb/, May **2005**.

5 http://nobelprize.org/chemistry/laureates/2003, May **2005**.

6 Lee, M.R.; Baker, D.; Kollman, P.A. *J. Am. Chem. Soc.* **2001**, *123*, 1040.

7 Brandon, C.; Tooze, J. *Introduction to Protein Structure*, Garland, 2nd Edt., **1998**.

8 Hansson, T.; Oostenbrink, C.; van Gunsteren, W. *Curr. Opin. Struct. Biol.* **2002**, *12*, 2, 190.

9 http://adam.about.com/reports/000033.htm, A.D.A.M. Healthcare Center, 17th Dec. **2004**.

10 Emberton, M. *Eur. Urol. Supp.* **2003**, *199*, *2*, 2.

11 Kirby, R.; McConnell, J. D.; FitzPatrick J. M.; Roehrborn, C.; Boyle, P. *Textbook of Benign Prostatic Hyperplasia*, Oxford, UK, Isis Medical Media Ltd., **1996**.

12 Kirby, R.S.; McConnell, J.D. *Fast Facts – Benign Prostatic Hyperplasia* Health Press, 4th Ed., **2002**.

13 Matyus, P. *Med. Res. Rev.* **1997**, *17*, 6, 523.

14 Kirby R. S.; Christmas, J.T. Benign Prostatic Hyperplasia, Mosby International, 2nd Edt., **1997**.

15 Dowding, D.; Swanson, V.; Bland, R.; Thomson, P.; Mari, C.; Morison, A.; Beechey, C.; Simpson, R.; Niven, K. *Patient. Ed. Counsel.*, **2004**, *52*, 2, 201.

16 Beduschi, M. C.; Beduschi, R.; Oesterling, J. E. *Urology* **1998**, *51*, 861.

17 Nickel, J.C. *Urology* **2003**, *62* (Supplment 3A).

18 Chapple, C.R. *Eur. Urol. Supp.* **2003**, *2*, 1.

19 Souverin, P.C.; Erkens, J.A.; de la Rosette, J.J.M.C.H.; Leufkens, H.G.M.; Herings, R.M.C. *Eur. Urol.* **2003**, *43*, 2528.

20 Carson III, C.; Rittmaster, R. *Urology Supp.* **2003**, *61*, 4A.

21 Narayan P.; Tewari, A. *Urology* **1998**, *51*, (Supplement 4A).

22 Ahlquist R.P. *Am. J. Physiol.* **1948**, *153*, 586.

23 Linder, M.E.; Gilman, A.G. *Scient. Amer.*, **1992**, *36*.

24 http://employees.csbsju.edu/hjakubowski/classes/ch331/signaltrans/Gproteincyclase.gif, 30th Jul. **2005**.

25 Wess, J. *Pharmacol. Ther.* **1998**, *80*, 3, 231.

26 Michelotti, G.A.; Price, D.T.; Schwinn, D.A. *Pharmacol. Ther.* **2000**, *88*, 281.

27 Minneman, K.P.; Esbenshade, T.A. *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 117.

28 Price, D.T.; Schwinn, D. A.; Lomasney, J. W.; Allen, L. F.; Caron, M. G.; Kowitz, R. J. *J. Urol.* **1993**, *150*, 546, 551.

29 Bremner, J.B.; Griffith, R.; Coban, B. *Curr. Med. Chem.* **2001**, *8* 607.

30 Beduschi, M.C.; Beduschi, R.; Oesterling, J.E. *Urology* **1998**, *51*, 6, 861.

31 Kenny, B.A.; Naylor, A.M.; Wyllie, M.G. *Prostatic Adrenoceptors. In* Kirby, R.; McConnell, J. D.; FitzPatrick J. M.; Roehrborn, C.; Boyle, P. Editors. *Textbook of Benign Prostatic Hyperplasia*, Oxford, UK, Isis Medical Media Ltd., **1996**.

32 Arkin, I.T.; Brunger, A.T.; Engelman, D.M. *Proteins: Struct. Funct. Genet.* **1997**, *28*, 465.

33 Wallin, E.; von Heijne, G. *Protein Sci.* **1998**, *7*, 1029.

34 Stevens, T.J.; Arkin, I.T. *Proteins* **2000**, *39*, 417.

35 Bockaert, J.; Pin, J.P. *EMBO J.* **1999**, *7*, 1723.

36 Okada, T.; Palczewski, K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 420.

37 Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.A.; Motoshima, H.; Fox, B.A.; Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E.; Yananoto, M.; Miyano, M. *Science* **2000**, *289*, 739.

38 www.upstate.com/features/ gprotein.q./Purified, Feb. **2003**.

39 Klabunde, T.; Hessler, G. *ChemBioChem* **2002**, *3*, 928.

40 Honig, B. *J. Mol. Biol.* **1999**, *293*, 283.

41 Tramontano, A.; Morea, V. Proteins, **2003**, *53*, 352.

42 Browne W.J.; North A.C; Phillips D.C.; Drew, K.; Vanaman, T.C.; Hill, R.L. *J. Mol. Biol.* **1969**, *42*, 65.

43 Sali, A.; Blundell, T.L. *J. Mol. Biol.* **1993**, *234*, 779.

44 Sanchez, R.; Sali, A. *Curr. Opin. Struct. Biol.* **1997**, *7*, 206.

45 Okada, T.; Ernst, O. P.; Palczewski, K.; Hofmann, K. P. *TiBS,* **2001**, *26*, 5, 318.

46 Forster, M.J. *Micron* **2002**, *33*, 365.

47 Thompson, J.D.; Higgins, D.G.; Gibson, T.J. *Nuc. Acids Res.* **1994**, *22*, 4673.

48 McCammon, J.A.; Gelin, B.R.; Karplus, M. *Nature* **1977**, *267*, 585.

49 Hansson, T.; Oostenbrink, C.; van Gensteren, W.F. *Curr. Opin. Struct. Biol.* **2002**, *12*, 190.

50 Shacham, S.; Topf, M.; Avisarm N.; Glaser, F.; Marantz, Y.; Bar-Haim, S.; Noiman, S.; Naor, Z.; Becker, O.M. *Med. Res. Rev.* **2001**, *21*, 5, 472.

51 Zhong, Q.; Moore, P.B.; Newns, D.M.; Klein, M.L. *FEBS Lett.* **1998**, *427*, 267.

52 Zhong, Q.; Husslein, T.; Moore, P.B.; Newns, D.M.; Pattnaik, P.; Klein, M.L. *FEBS Lett.* **1998**, *434*, 265.

53 Guba, W.; Haessner, G.; Breipohl, S.; Henke, S.; Knolle, J.; Santagada, V.; Kessler, H. *J. Am. Chem. Soc.* **1994,** *116*, 7532.

54 Pellegrini, M.; Mierke, D.F. *J. Med. Chem.* **1997**, *40*, 99.

55 Troxler, L.; Baaden, M.; Bohmer, V.; Wipff, G. *Supramo. Chem.* **2000**, *12*, 27.

56 Wymore, T.; Wong, T.C., 45th Annual Meeting of the Biophysical Society, Boston, MA, **2001**.

57 Chess-Williams, R. *Expert Opin. Pharmacother.* **2002**, *3*, 2.

58 Kortmann, B.B.M.; Floratos, D.L.; Kiemeney, L.A.L.M.; Wijkstra, H.; de la Rosette, J.J.M.C.H. *Urology* **2003**, *62*, 1.

59 Souverin, P.C.; Herings, R.M.C.; Man in 't Veld, A.J.; de la Rosette, J.J.M.C.H.; Farmer, R.D.T.; Leufkens, H.G.M. *Eur. Urol.* **2002**, *146*, 1.

60 Dardonville, C.; Goya, P.; Rozas, I.; Alasua, A.; Martin, I.; Borrego, J. *Bioorg. Med. Chem.* **2000**, *8*, 1567.

61 Dannert, M.T.; Alsasua, A.; Borrego, MJ.; Dardonville, C.; Sanchez, M.; Martin, M.I., XXIV Congreso de la Sociedad Española de Farmacología. *Methods and Findings in Experimental and Clinical Pharmacology*, **2002**, *24*, 158.

62 Kuo, G.; Prouty, C.; Murray, W.V.; Pulito, V.; Jolliffe, L.; Cheung, P.; Varga, S.; Evangelisto, M.; Wang, J. *J. Med. Chem.* **2000**, *43*, 2183.

63 Meyer, M.D.; Altenbach, R.J.; Basha, F.Z.; Carroll, W.A.; Condon, S.; Elmore, S.W.; Kerwin, J.F.; Sippy, K.B.; Tietje, K.; Wendt, M.D.; Hancock, A.A.; Brune, M.E.; Buckner, S.A.; Drizin, I. *J. Med. Chem.* **2000**, *43*, 1586.

64 Chern, J.W.; Tao, P.; Wang, K.; Gutcait, A.; Liu, S.; Yen, M.; Chien, S.; Rong, J. *J. Med. Chem.* **1998**, *41*, 3128.

65 Quaglia,W.; Pigini, M.; Piergentili, A.; Giannella, M.; Gentili, F.; Marucci, G.; Carrieri, A.; Carotti, A.; Poggesi, E.; Leocardi, A.; Melchiorre, C. *J. Med. Chem.* **2002**, *45*, 1633.

66 Barbaro, R.; Betti, L.; Botta, M.; Corelli, F.; Giannacini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. *J. Med. Chem.* **2001**, *44*, 2118.

67 Wong, W.C.; Chiu, G.; Wetzl, J.M.; Marzabadi, M.R.; Nagarathnam, D.; Wang, D.; Fang, J.; Miao, S.W.; Hong, X.; Forray, C.; Vaysse, P.J.J.; Branchek, T.A.; Gluchowski, C. *J. Med. Chem.* **1998**, *41*, 2643.

68 Lopez, F. J.; Arias, L.; Chan, R.; Clarke, D. E.; Elworthy, T. R.; Ford, A. P. D. W.; Guzman, A.; Jaime-Figueroa, S.; Jasper, J. R.; Morgans, D. J.; Padilla, F.; Perez-Medrano, A.; Quintero, C.; Romero, M.; Sandoval, L.; Smith, S. A.; Williams, T. J.; Blue, D. R. *Bioorg. Med. Chem. Lett.* **2002**, *13*, 1873.

69 Corsano, S.; Strappaghetti, G.; Barbaro, R.; Giannacini, G.; Betti, L.; Lucacchini, A. *Bioorg. Med. Chem.* **1999**, *7*, 5, 933.

70 Jang, Y.H.; Sowers, L.C.; Cagin, T.; Goddard III, W.A. *J. Phys. Chem. A*, **2001**, *105*, 274.

71 Silva, C.O.; da Silva, E.C.; Nascimento, M. A. C. *J. Phys. Chem. A.* **2000**, *104*, 2402.

72 Li, H.; Hains, A.; Everts, J.; Robertson, A.D.; Jensen, J.H. *J. Phys. Chem. B* **2002**, *106*, 3486.

73 Bremner, J.B.; Coban, B.; Griffith, R. *J. Comput. Aid. Mol. Des.* **1996**, *10*, 545.

74 Bremner, J. B.; Coban, B.; Griffith, R.; Groenewoud, K. M.; Yates, B. F. *Bioorg. Med. Chem.*, **2000**, *8*, 201.

75 Li, M.; Tsai, K.; Xia, L. *Bioorg. Med. Chem. Lett.*, **2005**, *15*, *3*, 657.

76 Strader, C. D.; Sieal. I. S.; Candelore. M. R.; Rands. E.; Hill, W. S.; Dixon, R. A. F. *J. Biol. Chem.* **1988**, *263*, 10267.

77 Porter, J.E.; Hwa, J.; Perez D.M. *J. Biol. Chem.* **1996**, *271*, 28318.

78 Wang, C.D; Buck, M.; Fraser, C.M. *Mol. Pharmacol.* **1991**, *40*, 168.

79 Cavalli, A.; Fanelli, F.; Taddei, C.; De Benedetti, P.G.; Cotecchia, S. *FEBS Lett.* **1996**, *399*, 9.

80 Piascik, M. T.; Perez, D. M. *JPET* **2001**, *298*, 403.

81 Ambrosio, C.; Molinari, P.; Cotecchia, S.; Costa, T. *Mol. Pharm.* **2000**, *57*, 198.

82 Del Castillo, J.; Katz, B. *Proc. R. Soc. Lond. [Biol]* **1957**, *146*, 369.

83 Gether, U.; Lin, S.; Ghanouni, P.; Ballesteros, J.A.; Weinstein, H.; Kobika, B.K. *EMBO J.* **1997**, *18*, 22, 6737.

84 Vaidehi, N.; Floriano, W.B.; Trabanino, R.; Hall, S.E.; Freddolino, P.; Choi, E.J.; Zamanakos, G.; Goddard III, W.A. *PNAS*, **2002**, *99*, 20, 12622.

85 Luo, X.; Zhang, D.; Weinstein, H. *Protein Eng.* **1994**, *7*, 1441.

86 Farrens, D.L.; Altenbach, C.; Yang, K.; Hubbell, W.L; Khorana, G. *Science* **1996**, *274*, 768.

87 Gether, U; Sanan, L.; Ghanouni, P.; Ballesteros, J.A.; Weinstein, H.; Kobilka, B.K. *EMBO, J.* **1997**, *16*, 22, 6737.

88 Oliveira, L.; Paiva, P.B.; Paiva, A.C.M.; Vriend, G. *Proteins* **2002**, *52*, 4, 553.

89 Ballesteros, J.; Palczewski, K. *Curr. Opin. Drug Disc. Devel.* **2001**, *4*, 561.

90 Greasley, P.J.; Fanelli, F.; Scheer, A.; Abuin, L.; Nenniger-Tosato, M.; DeBenedetti, P.G.; Cotecchia, S. *J. Biol. Chem.* **2001**, *276*, 49, 46485.

91 Gether, U.; Asmar, F.; Meinild, A.K.; Rasmusen, S.G.F. *Pharm. Tox.* **2002**, *91*, 304.

92 Seeber, M.; de Benedetti, P.G.; Fanelli, F. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1520.

93 Nyrönen, T.; Pihlavisto, M.; Peltonen, J.M.; Hoffrén, A.; Varis, M.; Salminen, T.; Wurster, S.; Marjamäki, A.; Kanerva, L.; Katainen, E.; Laaksonen, L.; Savola, J.; Scheinin, M.; Johnson M.S. *Mol. Pharm.* **2001**, *59*, 5, 1343.

94 Kobilka, B. *Mol. Pharm.* **2004**, *65*, 1060.

95 Gether, U.; Kobilka, B.K. *J. Biol. Chem.* **1998**, *273*, 29, 17979.

96 Javitch, J.A.; Fu, D.; Liapakis, G.; Chen, J. *J. Biol. Chem.* **1997**, *272*, 18546.

97 Trumpp-Kallmeyer, S.; Joflack, J.; Bruinvels, A.; Hibert, M. *J. Med. Chem.* **1992**, *35*, 3448.

98 Schwartzm T.W.; Rosenkilde, M.M. *Trends Pharmacol. Sci.* **1996**, *17*, 213.

99 Ishiguro, M.; Futabayashi, Y.; Ohnuki, T.; Ahmed, M.; Muramatsu, I.; Nagatomo, T. *Life Sci.* **2002**, *71*, 2531.

100 Zhao, M.; Hwa, J.; Perez, D. M. *Mol. Pharm.* **1996**, *50*, 1118.

101 Hamaguchi, N.; True, T.; Saussy, D. L.; Jeffs, P. W. *Biochem.* **1996**, *35*, 14312.

102 Waugh, D.; Gaivin, R.; Zuscik, M.; Gonzalez-Cabrera, P.; Ross, S.; Yun, J.; Perez, D. *J. Biol. Chem.* **2001**, *27*, 25366.

103 Meng, E.C.; Shiochet, B.K.; Kuntz, I.D. *J. Comput. Chem.* **1992**, *13*, 505.

104 Eldridge, M.D.; Murray, C.W.; Auton, T.R.; Paolini, G.V.; Mee, R.P. *J. Am. Chem. Soc.* **1997**, *11*, 425.

105 Boehm, II-J *J. Comput. Aided. Mol. Des.* **1991**, *8*, 243.

106 Muegge, I.; Martin, Y.C. *J. Med. Chem.* **1999**, *42*, 5, 791.

107 Lazaridis, T., *Curr. Organic Chem.* **2002**, *6*, 1319.

108 Sybyl Molecular Modelling System, version 6.8: Tripos Associates: St. Louis, MO.

109 Hardy, L.W.; Malikayil, A. *Curr. Drug Disc.* **2003**, 15.

110 Leach, A.R., *Molecular Modelling: Principles and Applications*, Prentice Hall, 2nd Ed., **1996**.

111 Fock, V. *Z. Physik.* **1930**, *61*, 126.

112 Pauli, W., *Z. Physik.* **1925**, *31*, 765.

113 Slater, J.C. *Phys. Rev.* **1929**, *34*, 1293.

114 Slater, J.C. *Phys. Rev.* **1930**, *35*, 509.

115 Jensen F., *Introduction to Computational Chemistry*, John Wiley & Sons, New York, **1998**.

116 Roothan, C.C.J. *Rev. Mod. Phys.* **1951**, *23*, 69.

117 Hall, G.G., *Proceedings of the Royal Society* (London), 1951, A205, 541.

118 Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, 864.

119 Kohn, W.; Sham, L. *J. Phys. Rev. Lett* **1985**, *55*, 2471.

120 Koch, W.; Holthausen, M.C. *A Chemists guide to Density Functional Theory*, Wiley-VCH, Weinheim, Germany, **2001**.

121 Becke, A.D. *Phys. Rev.* **1988**, *A38*, 3098.

122 Lee, C.; Yang W.; Parr, R.G. *Phys. Rev.* **1988**, *B37*, 785.

123 Becke, A.D. *J. Chem. Phys.* **1993**, *98*, 1372.

124 Becke, A.D. *J. Chem. Phys.* **1993**, *98*, 5648.

125 Stevens, P.J.; Devlin, F.J.; Chabalowski C.F.; Frisch, M.J. *J. Phys. Chem.* **1994**, *98*, 11623.

126 Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B.* **1988**, *37*, 785.

127 Vosko, S.J.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

128 Cornell, W. D.; Cieplak, P.; Bayly, I.; Gould, I. R.; Merz, K. M.; Ferguson. D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

129 MacKerell, A.D. *J. Phys. Chem. B* **1998**, *102*, 3586.

130 Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

131 Allinger, N.L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

132 Allinger, N.L.; Li, F.; Yan, L.; Tai, J.C. *J. Comp. Chem.* **1990**, *11*, 848.

133 Ewald, P. *Annalen der Physok* **1921**, *64*, 253.

134 Mulliken, R.S. *J. Chem. Phys.* **1955**, *23*, 1833.

135 Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, 3219.

136 Chirlian, L.E.; Francl, M.M. *J. Comput. Chem.* **1987**, *8*, 894.

137 Breneman, C.M.; Wiberg, K.B. *J. Comput. Chem.* **1990**, *11*, 361.

138 Besler, B.H.; Merz Jr., K.M.; Kollman, P.A. *J. Comp. Chem.* **1990**, *11*, 431.

139 Singh, U.C.; Kollman, P.A. *J. Comp. Chem.* **1984**, *5*, 129.

140 Honig, B.; Sharp, K.; Yang, A. *J. Phys. Chem.* **1993**, *97*, 1101.

141 Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. *J. Chem. Phys.* **1983**, *79*, 926.

142 Berendsen, H.C.; Postma, J..P.M.; van Gunsteren, W.F.; Hermans, J. *Interaction Models for Water in Relation to protein hydration.* Pullman B., Dordrecht, Reidel, 331, **1981**.

143 Levitt, M.; Hirshberg, M.; Sharon, R.; Laidig, K.E.; Daggett, V. *J. Phys. Chem. B* **1997**, *101*, 5051.

144 Born, M. *Zeitschrift fur Physik* **1920**, *1*, 45.

145 Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1586.

146 Wong, M.W.; Wiberg, K.B.; Frisch, M.J. *J. Am. Chem. Soc.* **1992**, *114*, 1645.

147 Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

148 Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

149 Carloni, P.; Alber, F. *Quantum Medicinal Chemistry* WILEY-VCH, Weinheim, Germany, **2003**.

150 Needleman, S.B.; Wuncsh, C.D. *J Mol Biol.* **1970**, *48*, 443.

151 Smith, T.F.; Waterman, M.S. *J Mol Biol.* **1981**, *147*, 195.

152 http://www.cryst.bbk.ac.uk/PPS95/course/9_quaternary/3_geometry/torsion.html, J. Cooper, May 7th 2005.

153 Ramachandran, G.N.; Sassiekharan, V. *Adv. Prot. Chem.* **1968**, *28*, 283.

154 Press, W.H.; Flannery, B.P.; Teukelsky, S.A.; Vetterling, W.T., *Numerical Recipes in Fortran*, Cambridge University Press, **1992**.

155 Schlegel, H.B. *J. Comp. Chem.* **1982**, *3*, 214.

156 Beusen, D.D.; Berkley Shands, E.F.; Karasek, S.F.; Marshall, G. R.; Dammkoehler, R.A. *J. Mol. Struct.* **1996**, *370*, 157.

157 Dardonville, C.; Goya, P.; Rozas, I.; Alasua, A.; Martin, I.; Borrego, J. *Bioorg. Med. Chem.* **2000**, *8*, 1567.

158 Haile, J.M. *Molecular Dynamics simulations* Wiley Interscience, **1992**.

159 Swope, W.C.; Anderson, P.H.; Berens, P.H.; Wilson, K.R. *J. Chem. Phys.* **1982**, *76*, 637.

160 Feller, S.E.; Zhang, Y.; Pastor, R.W.; Brooks, B.R. *J. Chem. Phys.* **1995**, *103*, 4613.

161 Wang, J.; Cieplak, P.; Kollman, P.A., *J. Comput. Chem.*, **2000**, *21*, 1049.

162 Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A., *J. Comput. Chem.* **2004**, *25*, 1157.

163 Verlet, L. *Phys. Rev.* **1967**, *165*, 201.

164 McQuarrie, D.A.; Simon, J.D. *Molecular Thermodynamics*, University Science Books, 55D Gate Five Road, Sausalito CA 94965, USA, **1999**.

165 Halls, M.D.; Velkovski, J.; Schlegel, H. B. *Theor. Chem. Acc.* **2001**, *105*, 413.

166 a) Reiss, H; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207. b) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610. c) Gilson, M.K.; Honig, B. *Proteins: Struct. Funct. Genet.* **1988**, *4*, 7.

167 a) Tissandier, M.D.; Cowen, K. A.; Feng, W.Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V. *J. Phys. Chem. A* **1998**, *102*, 7787. b) Tuňon, I.; Silla, E.; Bertrán, J. *J. Phys. Chem.* **1993**, *97*, 5547. c) Tawa, G.J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. *J. Chem. Phys.* **1998**, *109*, 4852. d) Mejias, J. A.; Lago, S. *J. Chem. Phys.* **2000**, *113*, 7306. e) Zhan, C.-G.; Dixon, D.A. *J. Phys. Chem. A* **2001**, *105*, 11534.

168 Fischer, E. *Ber. Dtsch. Chem.* Ges. **1894**, *27*, 2985.

169 Brady Jr., G.P.; Stouten, P.F.W. *J. Comput. Aid. Mol. Des.* **2000**, *14*, 4, 383.

170 Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R; Ferrin, T.E. *J. Mol. Biol.* **1982**, *161*, 269.

171 Miller, M.D.; Kearsky, S.K.; Underwood, D.J.; Sheridan, R.P. *J. Comput. Aid. Mol. Des.* **1994**, *8*, 153.

172 Welch, W.; Ruppert, J.; Jain, A.N. *Chem. Biol.* **1996**, *3*, 449.

173 Leach, A.R.; Kuntz, I.D. *J. Comp. Chem.* **1992**, *13*, 730.

174 Friedman, S.H.; Ganapathi, P.S.; Rubin, Y.; Kenyon, G.L. *J. Med. Chem.* **1998**, *41*, 2424.

175 Gschwend, D.A.; Sirawaraporn, W.; Santi, D.V.; Kuntz, I.D. *Proteins* **1997**, *29*, 59.

176 Hoffman, L.R.; Kuntz, I.D.; White, J.M. *J. Virol.* **1997**, *77*, 8808.

177 Sanner, M.; Olson, A.J.; Spehner, J.C., 11th ACM Symp. Comp. Geom, C6-C7, **1995**.

178 Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Zakrzewski, V.G.; Montgomery Jr. J.A.; Stratmann, R.E.; Burant, J.C.; Dapprich, S.; Millam, J.M.; Daniels, A.D., Kudin, K.N.; Strain, M.C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G.A.; Ayala, P.Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J.J.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Cioslowski, J.; Ortiz, J.V.; Baboul, A.G.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Challacombe, M.; Gill, P.M.W.; Johnson, B.; Chen, W.; Wong, M.W.; Andres, J.L.; Gonzalez, C.;

Head-Gordon, M.; Replogle, E.S.; Pople, J.A. Gaussian 98 (Rev A.11), Gaussian, Inc., Pittsburgh PA, 2001.

179 Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

180 Wang, R.; Lai, L.; Wang, S., *J Comput. Aided Mol. Des.*, **2002**, *16*, 11.

181 http://www.gpcr.org/7tm/, Vriend, G.; Horn, F., 28th Dec. **2004**.

182 Trumpp-Kallmeyer, S.; Hoflack, J.; Bruinvels, A.; Hibert, M. *J. Med. Chem.* **1992**, *35*, 3448.

183 Gether, U. *Endocrine Rev.* **2000**, *21*, 1, 90.

184 Baldwin, J.M.; Schertler, G.F.; Unger, V.M. *J. Mol. Biol.* **1997**, *272*, 144.

185 Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I. N.; Bourne, P.E. *Nucleic Acids Res.* **2000**, *28*, 235.

186 Teller, D.C.; Okada, T.; Behnke, C.A.; Palczewski, K.; Stenkamp, R.E. *Biochem.* **2001**, *40*, 7761.

187 Okada, T; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E.M.; Shichida, Y. *PNAS* **2002** *99*, 9, 5982.

188 Becker, O.M.; Shacham, S.; Marantz, Y.; Noiman, S. *Curr. Opin. Drug Disc. Devel.* **2003**, *6*, 3, 353.

189 Okada, T.; Palczewski, K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 420.

190 Kinsella, G.K.; Rozas, I.; Watson, G.W. *BBRC* **2004**, *324*, 916.

191 Pedretti, A.; Silva, M.E.; Villa, L.; Vistoli, G. *BBRC* **2004**, *319*, 493.

192 http://www.gpcr.org/7tm/models/vriend/index.html#001_001, Vriend, G.; Horn, F., 28th Dec. **2004**.

193 http://www.gpcr.org/7tm/models/vriend2/index.html, Vriend, G.; Horn, F., 28th Dec. **2004**.

194 http://www.gpcr.org/7tm/models/vriend3/index.html#001_001, Vriend, G.; Horn, F., 28th Dec. **2004**.

195 Evers, A.; Klabunde, T., *J. Med. Chem.* **2005**, *48*, 1088.

196 http://www.ebi.ac.uk/emboss/align/, EBI, 28th Dec. **2004**.

197 Sali A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.

198 Weber, I.T. *Proteins* **1990**, *7*, 172.

199 Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. *Ann. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.

200 Sansom, S.P.; Weinstein, H. *TiPS* **2000**, *21*, 445.

201 von Heijne, G. *J. Mol. Biol.* **1991**, *218*, 499.

202 Williams, K.A.; Deber, C. M. *Biochem.* **1991**, *30*, 8919.

203 Laskowski, R.A.; MacArthur, M.W.; Moss, D. S.; Thornton, J.M. *J. Appl. Crystallogr.* **1993**, *26*, 283.

204 Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.

205 Smondyrev , A.M.; Berkowitz, M.L. *J. Comput. Chem.* **1999**, *20*, 545.

206 Lin, J.; Baker, N.A.; McCammon, J.A. *Biophys. J.* **2002**, *83*, 1374.

207 Moore, P.B.; Lopez, C.F.; Klein, M.L. *Biophys. J.* **2001**, *81*, 2484.

208 a) Huang, K-S.; Bayley H.; Liao, M-J.; London, E.; Khorana, HG. *J. Biol. Chem.* **1981**, *256*, 3802. b) Popot, J-L.; Gerchman, S-E.; Engelmann, D. M. *J. Mol. Biol.* **1987**, *198*, 655. c) Sturgis, J. N.; Robert, B. *J. Mol. Biol.* **1994,** *238*, 445. d) Surrey, T.; Schmid, A.; Jahnig, F. *Biochem.* **1996**, *35*, 2283. e) Tribet, C.; Audebert, R.; Popot, J-L. *Proc. Natl. Acad. Sci.* USA, **1996**, *93*, 15047. f) Kikuchi, J.; Asakura, T.; Loach, P. A.; Parkes-Loach, P.S.; Shimada, K.; Hunter, C.N.; Conroy, M.J.; Williamson, M.P. *Biopolymers* **1999**, *49*, 361.

209 Weast R.C. *Handbook of Chemistry and Physics*, 60th Ed., CRC Press, **1980**.

210 Zhong, Q.F.; Jiang, Q.; Moore, P.B.; Newns, D.M.; Klein, M.L. *Biophys., J.* **1998,** *74*, 3.

211 Nagarathnam, D.; Wetzel, J.M.; Miao, S.W.; Marzabadi, M.R.; Chiu, G.; Wong, W.C.; Hong, X.; Fang, J.; Forray, C.; Branchek, T.A.; Heydorn, W.E.; Chang, R.S.L.; Broten, T.; Schorn, T.W.; Gluchowski, C. *J. Med. Chem.* **1998**, *41*, 5320.

212 Gross, G.; Hanft, G.; Rugevics, C. *Eur. J. Pharmacol.* **1988**, *151*, 333.

213 Blue, D.R.; Ford, A.P.D.W.; Morgans, D.J.; Padilla, F.; Clarke, D.E. *Neurourol. Urodyn.* **1996**, *14*, 345.

214 Barbaro, R; Betti, L.; Botta, M.; Corelli, F.; Giannaccini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. *Bioorg. Med.Chem.* **2002**, *10*, 361.

215 Kava, M.S.; Blue Jr, D.R.; Vimont, R.L.; Clarke, D.E.; Ford, A.P. *Br. J. Pharmacol.* **1998**, *123*, 1359.

216 Williams, T.J.; Blue, D.R.; Daniels, D.V.; Davis, B.; Elworthy, T.; Gever, J.R.; Kava, M.S.; Morgans, D.; Padilla, F.; Tassa, S.; Vimont, R.L.; Chapple, C.R.; Chess-Williams, R.; Eglen, R.M.; Clarke, D.E.; Ford, A.P.D.W. *Br. J. Pharmacol.* **1999**, *127*, 252.

217 Schroeder O.E.; Andriole, E.J.; Carver, K.L.; Colyer, K.E.; Poutsma, J.C. *J. Phys. Chem. A* **2004**, *108*, 326.

218 Kinser, R.D.; Nicol, G.; Ridge, D.P. *J. Phys. Chem. A* **2002**, *106*, 9925.

219 Curtiss, L.A.; Raghavachari, K.; Refern, P.C; Rassolov, V.; Pople, J.A. *J. Chem. Phys.* **1998**, *109*, 7764.

220 Pokon, E.K.; Lipak, M.D.; Feldgus, S.; Shields, G.C. *J. Phys. Chem. A* **2001**, *105*, 10483.

221 Burk, P.; Koppel, I.A; Koppel, I.; Leito, I.; Travnikova, O. *Chem. Phys. Lett.* **2002**, *323*, 482.

222 Dmitrenko, O.; Thorpe, C.; Bach, R.D., *J. Phys. Chem. B*, **2003**, *107*, 13229.

223 Bach, R.D.; Thorpe, C.; Dmitrenko, O. *J. Phys. Chem. B*, **2002**, *106*, 16, 4325.

224 Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.

225 Steel, M.C., Ph.D. *Thesis, Explorations in Solvation Models of Solvation within a quantum chemical framework*, Trinity College Dublin, March **2000**.

226 Wong, M.W.; Wiberg, K.B.; Frisch, M.J. *J. Am. Chem. Soc.* **1992,** *114*, 523.

227 Wong, M.W.; Wiberg, K.B.; Frisch, M.J. *J. Am. Chem. Soc.* **1992,** *114*, 1645.

228 De Benedetti, P. G.; Fanelli, F.; Menziani, M. C.; Cocchi, M.; Testa R.; Leonardi, A. *J. Struct. Chem.* **1991**, *233*, 343.

229 a) Tuňon, I.; Silla, E.; Bertrán, J. *J. Phys. Chem.* **1993**, *97*, 5547. b) Tawa, G.J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. *J. Chem. Phys.* **1998**, *109*, 4852. c) Mejias, J. A.; Lago, S. *J. Chem. Phys.* **2000**, *113*, 7306. d) Zhan, C.-G.; Dixon, D.A. *J. Phys. Chem. A* **2001** *105*, 11534. e) Tissandier, M.D.; Cowen, K.A.; Feng, W.

Y.; Gundlach, E.; Cohne, M.H.; Earhart, A. D.; Coe, J. V. *J. Phys. Chem. A* **1998**, *102*, 7787.

230 Strange P. G., *TiPs*, **1996**, 17.

231 Hansch, C.; Leo A.J., Mechem Project Issue No.26, Pomona College, Claremont, CA.

232 Perrin, D.D. *Dissociation Constants of Organic Bases in Aqueous Solution* (compiled for IUPAC) Butterworths, London, **1965**.

233 Data on file, Abbott Laboratories (scientific Report PPRd/86/238, 10.

234 Pfizer Global Pharmaceuticals, (Internal Communication).

235 Merck Index – an Encyclopaedia of chemicals, drugs and biologicals, 13th Edt., O'Neil, M.J., Merck & Co., Inc., Whitehouse Station, NJ, **2001**.

236 Kinsella, G.K.; Rodriguez, F.; Watson, G.W.; Rozas, I., unpublished data, **2005**.

237 Aggarwal, V.K.; Emme, I.; Fulford, S.Y. *J. Org. Chem.* **2003**, *68*, 692.

238 Magill, A.M.; Yates, B.F., *Aust. J. Chem.*, **2004**, *57*, 1205.

239 Chen, J.L.; Noodleman, L.; Case, D.A.; Bashford, D. *J. Phys. Chem.* **1994,** *98*, 11059.

240 Barbaro, R.; Betti, L.; Botta, M.; Corelli, F.; Giannacini, G.; Maccari, L.; Manetti, F.; Strappaghetti, G.; Corsano, S. *J. Med. Chem.* **2001**, *44*, 2118.

241 Li, M.-Y.; Tsai, K.-C.; Xia, L. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 657.

242 Kinsella, G.K.; Rozas, I.; Watson, G.W., *Bioorg. Med. Chem.*, submitted Apr. **2005**.

243 Chipman, D.M. *J. Phys. Chem. A* **2002**, *106*, 7413.

244 Bea, I.; Jaime, C.; Kollman, P. *Theor. Chem. Acta.* **2002**, *108*, 286.

245 Strader, C.D.; Sigal, I.S.; Register, R.B.; Candelore, M.R.; Rands, E.; Dixon, R.A. *Proc. Natl. Acad. Sci.* **1987**, *84*, 4384.

246 Strader, C.D. *Mol. Pharmacol.* **1991**, *40*, 168.

247 Greasley, P.J.; Fanelli, F.; Scheer, A.; Abuin, L.; Nenniger-Tosato, M.; DeBenedetti, P.G.; Cotecchia, S. *J. Biol. Chem.* **2001**, *276*, 49, 46485.

248 Scheer, A.; Fanelli, F.; Costa, T.; de Benedetti, P.G.; Cotecchia, S. *Proc. Natl. Acad. Sci.* USA **1997**, *94*, 808.

249 Kobilka, B. *Mol. Pharmacol.* **2004**, *65*, 5, 1060.

250 Gether, U.; Kobilka, B.K. *J. Biol. Chem.* **1998**, *273*, 29, 17979.

251 Javitch, J.A.; Fu, D.; Liapakis, G.; Chen, J. *J. Biol. Chem.* **1997**, *272*, 18546.

252 Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. *Proteins* **2003**, *50*, 5.

253 Chambers, J.J.; Nichols, D.E. *J. Comput. Aided Mol. Des.* **2002**, *16*, 511.

254 Carmine, R.; Molinari, P.; Sbraccia, M.; Ambrosio, C.; Costa, T. *Mol. Pharmacol.* **2004**, *66*, 2.

255 Wang, J.; Morin, P.; Wang, W.; Kollman, P.A. *J. Am. Chem. Soc.* **2001**, *123*, 5221.

256 Cambridge Structural Database, http://www.ccdc.cam.ac.uk/.

257 Aldrich Library of FT-IR Spectra, 1 (1), 1296A (ir).

258 Wang, J.; Cieplak, P.; Kollman, P.A., *J. Comput. Chem.*, **2000,** *21*, 1049.

259 Javitch, J.A.; Ballesteros, J.A.; Weinstein, H.; Chen, J., *Biochem.*, **1997**, *37*, 998.

260 Visiers, I.; Ballesteros, J.A.; Weinstein, H. *Methods Enzymol.* **2002**, *343*, 329.

261 Ambrosio, C.; Molinari, P.; Cotecchia, S.; Costa, T. *Mol. Pharm.* **2000**, *57*, 198.

262 Liapakis, G.; Ballesteros, J. A.; Papachristou, S.; Chan, W. C.; Chen, X.; Javitch, J. A. *J. Biol. Chem.* **2000**, *275*, 37779.

263 Strader, C.D.; Fong, T.M.; Tota, M.R.; Underwood, D.; Dixon, R.A., *Annu Rev. Biochem.*, **1994,** *63*, 101.

264 Wess, J., *Pharmacol. Therap.*, **1998,** *80*, 231.

265 Dunham, T.D.; Farrens, D.L. *J. Biol. Chem.* **1999**, *274*, 1683.

266 Farrens, D.L.; Altenbach, C.; Yang, K.; Hubbell, W.L.; Khorana, H.G. *Science* **1996,** *274*, 768.

267 Sheikh, S.P.; Vilardarga, J.P.; Baranski, T.J.; Lichtarge O.; Iiri, T.; Meng, E.C.; Nissenson, R.A.; Bourne, H.R. *J. Biol. Chem.* **1999**, *274*, 17033.

268 Jensen, A.D.; Guarnieri, F.; Rasmussen, S.G.; Asmar, F.; Ballesteros, J.A.; Gether, U. *J. Biol. Chem.* **2001**, *276*, 9279.

269 Gether, U.; Asmar, F.; Meinild, A.K.; Rasmussen, G.F. *Pharmacol. Toxicol.* **2002**, *91*, 304.

270 Gershengorn, M.C.; Osman, R. *Endocrinology* **2001**, *142*, 2.

271 Zhou, W.; Flanagan, C.; Ballesteros, J.A.; Konvocka, K.; Davidson, J.S.; Weinstein, H.; Millar, R.P.; Sealfon, S.C. *Mol. Pharmacol.* **1994**, *45*, 165.

272 Sealfon, S.C.; Chi,L.; Ebersole, B.J.; Rodic, V.; Zhang, D.; Ballesteros, J.A.; Weinstein, H. *J. Biol. Chem.* **1995**, *270*, 16683.

273 Jongejan, A.; Bruysters, M.; Pardo, L.; Leurs, R., Oral Presentation, XVIIth Symposium of Medicinal Chemistry, **2004**.

274 Furse, K., E.; Lybrand, T. P. *J. Med. Chem.* **2003**, *46*, 21, 4450.

275 Luo, X.; Zhang, D.; Weinstein, H., *Protein Eng.*, **1994**, *7*, 1441.

276 Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P.G.; Cotecchia, S., *Embo J.*, **1996**, *15,* 3566.

277 Zhang, D.; Weinstein, H., *J. Med. Chem.*, **1993**, *36*, 934.

278 Farrens, D.L.; Altenbach, C.; Yang, K.; Hubbell. W.L.; Khorana, H.G., *Science*, **1996**, *274,* 768.

279 Gether, U.; Ballesteros, J.A.; Seifert, R.; Sanders-Bush, E.; Weinstein, H.; Kobilka, B.K., *J. Biol. Chem.*, **1997**, *272,* 2587.

280 Gether, U.; Lin, S.; Ghanouni, P.; Ballesteros, J.A.; Weinstein, H.; Kobilka, B.K., *Embo J.*, **1997**, *16*, 6737.

281 Gether, U.; Lin, S.; Kobilka, B.K., *J. Biol. Chem.*, **1995,** *270*, 28268.

282 George, S.R.; O'Dowd, B.F.; Lee, S.P. *Nat. Rev. Drug Disc.*, **2002**, *1*, 808.

283 Wurch, T.; Colpaert, F.C.; Pauwels, P.J. *Mol. Pharmacol.* **1998**, *1*, *54,* 6, 1088.

284 Varga, E.; V., Li, X.; Stropova, D.; Zalewska, T.; Landsman, R. S.; Knapp, R. J.; Malatynska, E.; Kawai, K.; Mizusura, A.; Nagase, H. *Mol. Pharmacol.* **1996**, *50*, 1619.

285 Waugh, D. J. J.; Gaivin, R.J.; Zuscik, M.J.; Gonzalez-Cabrera, P.; Ross, S.A.; Yun, J.; Perez, D.M. *J. Biol. Chem.* **2001**, *6*, 276, 27, 25366.

286 Ishiguro, M.; Futabayashi, Y.; Ohnuki, T.; Ahmed, M.; Muramatsu, I.; Nagatomo, T. *Life Sciences* **2002**, *71*, 2531.

287 Evers, A.; Klabunde, T., *J. Med. Chem.*, **2005**, *48*, 1088.

288 Wang, C.D.; Buck, M.; Fraser, C.M. *Mol. Pharmacol.* **1991**, *40*, 168.

289 Pretsch, E.; Clerc, T.; Seibl, J.; Simon, W. *Tables of Spectral Data for Structure Determination of Organic Compounds*, 2nd Edt., Springer-Verlay, **1989**.

290 Dardonville, C. Bis(2-aminoimidazolinium) and bis-guanidinium derivatives with analgesic and alpha adrenergic activities and affinity towards imidazoline receptors", (in Spanish), Ph.D. Thesis by Christophe Dardonville, Universidad Complutense de Madrid (Spain), June **2000**.

291 Silverstein, R.M.; Bassler, C.; Morrill, T.C. *Spectrometric Identification of Organic Compounds* 4th Edt., Wiley, NY, **1981**.

292 Kemp W. *Organic Spectroscopy* W.H. Freeman, New York, 3rd Ed., **1991**.

293 Wang, R.; Lu, Y.; Fang, X.; Wang, S. *J. Chem. Inf. Comput.* Sci. **2004**, *44*, 2114.

294 Kenakin, T. *Trends Pharm. Sci.* **1997**, *18*, 416.

295 Salamon, Z.; Hruby, V.J.; Tollin, G.; Cowell, S. *J. Pept. Res.*, **2002**, *60*, 322.

296 Furse, K.E.; Lybrand, T.P. *J. Med. Chem.* **2003**, *46*, 4450.

297 Gerchengorn, M. C.; Osman, R. *Endocrinology* **2001**, *142*, 1.

298 Muller, G. *Curr. Med. Chem.* **2000**, *7*, 861.

299 Matijssen, B.C.AG.; Rozas, I.; Watson, G.W., unpublished results, 2005.

300 Hebert, T.E.; Moffett, S.; Morello, J.P.; Loisel, T.P.; Bichet, D.G.; Barret, C.; Bouvier, M. *J. Biol. Chem.* **1996**, *271*, 16384.

301 Cvejic, S.; Devi, L.A. *J. Biol. Chem.* **1997**, *272*, 26959.

302 Romano, C.; Yang, W.L.; O'Malley, K.L. *J. Biol. Chem.* **1996**, *271*, 28612.

303 Milligan, G.; Pediani, J.; Fidock, M.; Lopez-Gimenez. J.F. *Biochem. Soc. Trans.* **2004**, *32*, 5, 847.

304 Isralewitz, B.; Baudry, J.; Gullingsrud, J.; Kosztin, D.; Schulten, K. *J. Mol. Graphics Modell.* **2001**, *19*, 13.